Risk management of patients with atrial fibrillation in general practice

Sander van Doorn

Risk management of patients with atrial fibrillation in general practice
Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht.
PhD thesis, Utrecht University, the Netherlands.

ISBN/EAN: 978-94-6233-717-6
Author: Sander van Doorn
Cover design: Ilse Vis (ilsevis.design@gmail.com)
Layout & printing: Gildeprint, Enschede

The printing of this thesis was financially supported by SBOH, employer of GP trainees in the Netherlands.
Part of the research reported in this thesis was funded by Boehringer Ingelheim.

Risk management of patients with atrial fibrillation in general practice

Risico management van patiënten met atriumfibrilleren in de eerste lijn

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op donderdag 28 september 2017 des middags te 2.30 uur

door

Sander van Doorn

geboren op 22 mei 1984
te Amsterdam

Promotoren:	Prof.dr. K.G.M. Moons
	Prof.dr. A.W. Hoes
Copromotoren:	Dr. G.J. Geersing
	Dr. F.H. Rutten

CONTENTS

Introduction

THE CASE OF MRS. ROBERTS

Mrs. Roberts is a 73 years old lady who was a primary school teacher during her active working life. She visits her general practitioner (GP) infrequently, most often for osteoarthritis related complaints of knees and hips.

Now, she consults her GP for coughing along with a sense of headache and lightheadedness, a combination of complaints she never experienced before. Upon physical examination, the GP notices a fast and irregular pulse and a subsequent electrocardiogram (ECG) confirms his suspicion: Mrs. Roberts has atrial fibrillation.

As recommended in the Dutch College of General Practitioners' guidelines on the management of atrial fibrillation, the GP considers initiating oral anticoagulation for stroke prevention after balancing the individual risk of stroke against the risk of bleeding associated with anticoagulation. Based on her CHA2DS2-VASc score of 2, the GP estimates the stroke risk of Mrs. Roberts to be around 2% each year, but what about her risk of bleeding?

The health record of Mrs. Roberts mentions three falls, and the GP knows she sometimes uses over-the-counter non-steroidal anti-inflammatory drugs (NSAIDs) as painkillers for her osteoarthritis. In addition, the daughter of Mrs. Roberts recently consulted the GP because she feared her mother suffered some cognitive decline: she showed increasing clumsiness and more difficulties in remembering things. How would this affect her medication compliance? Could her coughing be the result of heart failure as a concomitant problem with her atrial fibrillation? How frail is Mrs. Roberts?

The GP considers the risk of stroke high enough to warrant anticoagulation, despite the associated risk of bleeding. He prescribes a vitamin K antagonist and plans to carefully monitor Mrs. Roberts for her medication use and possible comorbid conditions.

ATRIAL FIBRILLATION

Atrial fibrillation (AF) is the most common cardiac arrhythmia. Its prevalence in the population at large is 1–2%,[1] and one in every four adults will eventually develop AF.[2] In the Netherlands, the healthcare costs associated with AF well exceed €500 million each year, mostly spent on inpatient hospital care and interventions.[3] Many factors can contribute to the development of AF, including genetic predisposition, structural heart disease and hypertension. Structural remodelling of the atria of the heart advances with increasing age, making AF predominantly a disease of the elderly. Its prevalence rises from 4% in those aged 65 years to over 15% in elderly over 85 years of age.[4] As a consequence of the aging population, and in addition to improved detection, the number of patients with AF is expected to increase in the coming decades to over half a million patients in the Netherlands by 2060.[5]

RISKS IN ATRIAL FIBRILLATION

Patients with atrial fibrillation are at high risk of a multitude of adverse outcomes. Most importantly, the risk of ischaemic stroke is on average increased five-fold.[6] In the treatment of AF, preventing of stroke therefore is a primary goal.

Oral anticoagulants can effectively prevent stroke with a relative risk reduction of 66%,[7–11] but inherently carry the risk of bleeding complications. Consequently, they should only be prescribed to patients whose risk of stroke is high enough to outweigh risk of bleeding. To estimate this stroke risk prior to anticoagulant treatment, prediction models have been developed of which the CHA2DS2-VASc is the most commonly used.[12] In this model, several patient characteristics are summed to result in a score, shown in Table 1. Subsequently, for patients with a high CHA2DS2-VASc score (corresponding to a high risk of stroke) anticoagulants are indicated, while for those with a low score (i.e. low risk of stroke) the benefits are considered not to outweigh the risk of bleeding.

TABLE 1. The original CHA2DS2-VASc score[12]

Risk Factor	Score
Congestive heart failure/LV dysfunction	1
Hypertension	1
Age ≥ 75 years	2
Diabetes mellitus	1
Stroke/TIA/TE	2
Vascular disease (prior myocardial infarction, peripheral artery disease, or aortic plaque)	1
Age 65–74 years	1
Sex category (i.e. female sex)	1

TE = thromboembolism

Studies have shown, however, that many AF patients at high risk of stroke do not receive anticoagulants.[13] Furthermore, there is currently a heavy debate[14–18] about whether in patients with a low CHA2DS2-VASc score (a score of 1 or 2) such as Mrs. Roberts anticoagulants are indicated. Is stroke risk in females with a score 2 truly 2% per year, and does the 66% relative risk reduction of anticoagulants outweighs the annual bleeding risk?

Moreover, not only stroke prevention is important in AF, but also dealing with the increased risk of heart failure,[19] hospitalisation,[20] mortality[21] and multiple other adverse events, such as renal failure[22] or cognitive decline.[23]

ATRIAL FIBRILLATION IN GENERAL PRACTICE

Many patients with AF are managed in general practice. They may present with their very first symptoms here, and treatment may be initiated. General practitioners need to monitor the disease course, and timely refer to secondary care in case specialist care is needed. Furthermore, GPs are often responsible for managing any co-occurring (chronic) conditions, each with potential complications and treatment interactions.

Altogether, atrial fibrillation poses tremendous challenges to general practitioners, patients and the healthcare system. This thesis aims to address major elements of these challenges by focussing on the following research domains relevant for general practice:

- The predictive ability of the CHA2DS2-VASc model in different patient settings;
- Improving anticoagulant treatment of patients with atrial fibrillation in general practice;
- Concomitant events in atrial fibrillation other than stroke, including heart failure, hospitalisation and mortality.

OUTLINE OF THIS THESIS

In **Chapter 1**, we present the results of a systematic review and meta-analysis of studies that validated the CHA2DS2-VASc prediction model, stratifying studies into general population and hospital care settings. In **Chapter 2** we addressed the reasons for non-adherence to the prevailing practice guideline on the management of atrial fibrillation in general practice. To improve the adherence to guidelines, we performed a large cluster randomised trial in general practice evaluating the effectiveness of an automatic decision support system for anticoagulant treatment recommendations, the results of which are described in **Chapter 3**. In **Chapter 4**, we describe the risk of concomitant events other than ischaemic stroke for patients with atrial fibrillation and whether these events can be predicted using existing stroke prediction models. In **Chapter 5**, we describe the prevalence of heart failure in high-risk community-dwelling patients with atrial fibrillation and study whether the biomarker NTproBNP can efficiently be used for heart failure screening. In **Chapter 6** we present the results of a case study addressing misclassification in data from routine healthcare and describe its influence on the validation of the CHA2DS2-VASc prediction model. Finally, the possible implications of our findings on future research on improving stroke risk prediction in patients with atrial fibrillation are discussed.

REFERENCES

1 Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, Singer DE. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* 2001; 285: 2370–5.

2 Lloyd-Jones DM, Wang TJ, Leip EP, Larson MG, Levy D, Vasan RS, D'Agostino RB, Massaro JM, Beiser A, Wolf PA, Benjamin EJ. Lifetime Risk for Development of Atrial Fibrillation. *Circulation* 2004; 110: 1042–6.

3 Ringborg A, Nieuwlaat R, Lindgren P, Jönsson B, Fidan D, Maggioni AP, Lopez-Sendon J, Stepinska J, Cokkinos DV, Crijns HJGM. Costs of atrial fibrillation in five European countries: results from the Euro Heart Survey on atrial fibrillation. *Europace* 2008; 10: 403–11.

4 Heeringa J, van der Kuip DAM, Hofman A, Kors JA, van Herpen G, Stricker BHC, Stijnen T, Lip GYH, Witteman JCM. Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study. *Eur Heart J* 2006; 27: 949–53.

5 Krijthe BP, Kunst A, Benjamin EJ, Lip GYH, Franco OH, Hofman A, Witteman JCM, Stricker BH, Heeringa J. Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. *Eur Heart J* 2013; 34: 2746–51.

6 Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991; 22: 983–8.

7 Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, Pogue J, Reilly PA, Themeles E, Varrone J, Wang S, Alings M, Xavier D, Zhu J, Diaz R, Lewis BS, Darius H, Diener H-C, Joyner CD, Wallentin L, et al. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2009; 361: 1139–51.

8 Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, Breithardt G, Halperin JL, Hankey GJ, Piccini JP, Becker RC, Nessel CC, Paolini JF, Berkowitz SD, Fox KAA, Califf RM, ROCKET AF Investigators. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med* 2011; 365: 883–91.

9 Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M, Al-Khalidi HR, Ansell J, Atar D, Avezum A, Bahit MC, Diaz R, Easton JD, Ezekowitz JA, Flaker G, Garcia D, Geraldes M, Gersh BJ, Golitsyn S, Goto S, et al. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011; 365: 981–92.

10 Giugliano RP, Ruff CT, Braunwald E, Murphy SA, Wiviott SD, Halperin JL, Waldo AL, Ezekowitz MD, Weitz JI, Špinar J, Ruzyllo W, Ruda M, Koretsune Y, Betcher J, Shi M, Grip LT, Patel SP, Patel I, Hanyok JJ, Mercuri M, et al. Edoxaban versus Warfarin in Patients with Atrial Fibrillation. *N Engl J Med* 2013; 369: 2093–104.

11 Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Ann Intern Med* 2007; 146: 857–67.

12 Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. 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.

13 Ogilvie IM, Newton N, Welner SA, Cowell W, Lip GYH. Underuse of oral anticoagulants in atrial fibrillation: a systematic review. *Am J Med* 2010; 123: 638–645e4.

14 Lip GYH, Nielsen PB. Should Patients With Atrial Fibrillation and 1 Stroke Risk Factor (CHA2DS2-VASc Score 1 in Men, 2 in Women) Be Anticoagulated?Response to Lip and Nielsen. *Circulation* 2016; 133: 1498–503.

15 Huisman MV. Patients With Atrial Fibrillation and a CHA2DS2-VASc Score of 1: Are They at Low or High Stroke Risk? *J Am Coll Cardiol* 2015; 65: 1395–7.

16 Savino JA, Halperin JL. Should Patients With Atrial Fibrillation and 1 Stroke Risk Factor (CHA2DS2-VASc Score 1 in Men, 2 in Women) Be Anticoagulated?Response to Savino and Halperin. *Circulation* 2016; 133: 1504–11.

17 Olesen JB, Torp-Pedersen C. Stroke risk in atrial fibrillation: Do we anticoagulate CHADS2 or CHA2DS2-VASc ≥ 1, or higher? *Thromb Haemost* 2015; 113: 1165–9.

18 Friberg L, Skeppholm M, Terént A. Benefit of Anticoagulation Unlikely in Patients With Atrial Fibrillation and a CHA2DS2-VASc Score of 1. *J Am Coll Cardiol* 2015; 65: 225–32.

19 Maisel WH, Stevenson LW. Atrial fibrillation in heart failure: epidemiology, pathophysiology, and rationale for therapy. *Am J Cardiol* 2003.

20 Steinberg BA, Kim S, Fonarow GC, Thomas L, Ansell J, Kowey PR, Mahaffey KW, Gersh BJ, Hylek E, Naccarelli G, Go AS, Reiffel J, Chang P, Peterson ED, Piccini JP. Drivers of hospitalization for patients with atrial fibrillation: Results from the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF). *Am Heart J* 2014; 167: 735–42. e2.

21 Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H. Impact of atrial fibrillation on the risk of death the Framingham Heart Study. *Circulation* 1998.

22 Bansal N, Fan D, Hsu C-Y, Ordonez JD, Marcus GM, Go AS. Incident Atrial Fibrillation and Risk of End-Stage Renal Disease in Adults with Chronic Kidney Disease. *Circulation* 2012; 127: CIRCULATIONAHA.112.123992–574.

23 Ott A, Breteler MMB, de Bruyne MC, van Harskamp F, Grobbee DE, Hofman A. Atrial Fibrillation and Dementia in a Population-Based Study: The Rotterdam Study. *Stroke* 1997; 28: 316–21.

1

Predictive performance of the CHA2DS2-VASc rule in atrial fibrillation: a systematic review and meta-analysis

Sander van Doorn, Thomas P.A. Debray, Femke Kaasenbrood, Arno W. Hoes, Frans H. Rutten, Karel G.M. Moons, Geert-Jan Geersing

J Thromb Haemost, 15(6), 1065–1077

ABSTRACT

BACKGROUND: The CHA2DS2-VASc decision rule is widely recommended for estimating stroke risk in patients with atrial fibrillation (AF) though validation studies show ambiguous and conflicting results.

OBJECTIVES: We aimed to (1) review existing studies validating CHA2DS2-VASc in AF patients not (yet) anticoagulated, 2) meta-analyse estimates of stroke risk per score, and 3) explore sources of heterogeneity across the validation studies.

METHODS: We performed a systematic literature review and random effects meta-analysis of studies externally validating CHA2DS2-VASc in AF patients not on anticoagulants. To explore between-study heterogeneity in stroke risk, we stratified studies to the clinical setting in which patient enrolment started, and performed meta-regression.

RESULTS: In total 19 studies were evaluated with over two million person-years of follow-up. In studies recruiting AF patients in hospitals, stroke risk for a score of zero, one and two were 0.4% (approximate 95% prediction interval (PI) 0.2 to 3.2%), 1.2% (95% PI 0.1 – 3.8%) and 2.2% (95% PI 0.03 – 7.8%), respectively. This was consistently higher than studies recruiting patients from the open general population, with risks of 0.2% (95% PI 0.0 – 0.9%), 0.7% (0.3 – 1.2%) and 1.5% (95% PI 0.4 – 3.3%) for score zero to two respectively. Heterogeneity as reflected by the wide prediction intervals could not be fully explained by meta-regression.

CONCLUSIONS: Studies validating CHA2DS2-VASc demonstrate high heterogeneity in predicted stroke risks for different scores.

INTRODUCTION

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia[1]and a major risk factor for ischaemic stroke.[2] Anticoagulants – such as vitamin K antagonists (VKA) or direct oral anticoagulants (DOACs) – can effectively reduce stroke risk,[3,4] but their relative benefits and harms depend on the absolute risk of stroke while off treatment, given that they inherently carry a risk of (major) bleeding complications. Hereto, clinical decision rules have been developed to estimate stroke risk in AF patients, with the CHA2DS2-VASc rule as the most well-known example.[5] Published as an update to the CHADS2 rule[6], the CHA2DS2-VASc was first recommended in the 2010 ESC practice guideline.[7] A swift uptake in clinical practice was followed, but subsequent validation studies showed ambiguous and conflicting results. This is for instance exemplified by the ongoing debate on the optimal threshold below which stroke risk is low enough to omit anticoagulation.[8–12]

Therefore, the aim of the present study was to evaluate the current evidence-base of using CHA2DS2-VASc for predicting stroke in AF patients. Hereto, we performed the following steps: 1) review existing studies validating CHA2DS2-VASc for AF patients not (yet) anticoagulated, 2) meta-analyse estimates of the c-statistic and stroke risk per score, and 3) explore sources of heterogeneity across the validation studies.

METHODS

Throughout the planning and conducting of this systematic review we followed the CHARMS recommendations for framing the review question, critical appraisal, and data extraction for systematic reviews of prediction modelling studies.[13] See Table 1 for details.

The CHA2DS2-VASc score

The CHA2DS2-VASc clinical decision rule was developed in 2010 by Lip et al. as an update to the original CHADS2 rule[6] by including additional predictors for stroke. Patients were assigned points for congestive heart failure (1 point), hypertension (1 point), age above 75 years (2 points), diabetes (1 point) and prior stroke (2 points), age above 65 (1 point), vascular disease (1 point) and female sex (1 point). Risk categories were defined using the total sum of scored points, and consisted of 'low' (0 points), 'intermediate' (1 point) and 'high (≥ 2 points). Using these categories, the c-statistic was 0.61 (0.51–0.70) in the derivation cohort.[5] See Supplemental Material 1. No efforts were made to adjust the c-statistic for potential over-optimism.[14]

TABLE 1: Framing the review aim using CHARMS key items

Item	Comment
1. Type of model	CHA2DS2-VASc rule
2. Intended scope of review	1) review existing studies validating CHA2DS2-VASc for AF patients not (yet) anticoagulated, 2) meta-analyse estimates of c-statistic and stroke risk per score, and 3) explore sources of heterogeneity across the validation studies.
3. Type of modelling studies	External validation studies of CHA2DS2-VASc
4. Target population to whom the model applies	Patients with non-valvular AF not already treated with oral anticoagulants.
5. Outcomes to be predicted	Any of ischaemic stroke and/or TIA; all types of stroke; systemic thromboembolism or a combination thereof.
6. Time span	One year risk of the outcome
7. Intended moment of using the model	At the time of diagnosis of AF, and annually when revising anticoagulant treatment indication.

Data sources and search strategy

We performed a systematic search to identify all studies that validated the CHA2DS2-VASc rule in patients with non-valvular AF. Medline and Embase were searched from January 1st 2001 till March 1st 2017. The search syntax was based on the broad Ingui search filter for identifying prediction studies,[15] and augmented with the filter by Geersing et al.[16] and the term 'Atrial Fibrillation' with its MeSH heading (Supplemental Material 2). Cross-reference checks were performed using the reference lists of each selected article.

Study selection

As CHA2DS2-VASc was specifically developed to guide anticoagulant decision making, notably for selecting AF patients in whom anticoagulant therapy can be safely withheld, we focused on studies validating this decision rule in AF patients not already treated with anticoagulants. To identify articles eligible for this review, the following inclusion criteria were used:

- Original research articles on the external validation of CHA2DS2-VASc (i.e. validation in patients not used for the derivation of the score);
- Including adults aged > 18 years with non-valvular AF;
- AF patients not yet treated with anticoagulation, or data presented separately for those not anticoagulated. Treatment with antiplatelet therapy was allowed;
- Allowing for extraction of c-statistic and/or absolute stroke or thrombo-embolic risks at different risk scores of CHA2DS2-VASc.

Studies including patient populations that dictate specific treatment decisions regardless of the score on a clinical decision rule, e.g. those after cardiac surgery, with

mechanical heart valves or mitral valve stenosis, after ablation or left appendage closure, were excluded.

A single reviewer (SvD) performed the study selection and included all eligible articles after consensus with a second reviewer (GJG).

If different articles used subsets of the same data source, the article studying the patient population most representative for our study domain was included after consensus between reviewers, or after consultation with the corresponding authors where needed.

Critical appraisal and risk of bias assessment

We critically appraised the selected studies according to the Checklist for critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) guidelines.[13] From the checklist we identified twelve items relevant for external validation studies (see Supplemental Material 3 for an overview). Two independent reviewers (SvD, FK) scored the risk of bias for each item (no risk of bias, risk of bias or unclear) and decided on a summary risk of bias estimate where studies without high risk on any item were considered at low risk of bias. Any disagreements were resolved by consensus with a third reviewer (GJG).

Data extraction and quantitative synthesis

Data extraction was independently performed by two reviewers (SvD, FK), disagreements were resolved by discussion. We stratified studies in those recruiting patients from an unselected general population setting – e.g. primary care databases or healthcare insurance data – and studies enrolling a selected subsample of AF patients, e.g. recruited during a hospital admission or during a visit to an outpatient cardiology department. We subsequently extracted the following information from each validation study, if reported:

- Setting: setting (e.g. general population or hospital care setting), locations and periods of recruitment;
- Study characteristics, i.e. the study design and the source of the data, the number of patients and total duration of follow-up, geographic region;
- Outcomes: type of outcomes studied (ischaemic stroke; all strokes including hemorrhagic stroke; or thromboembolism, commonly defined as ischaemic stroke, TIA, systemic embolism or a combination thereof);
- Population characteristics: the annual incidence of the main outcome, mean patient age, proportion of patients with (congestive) heart failure, hypertension, diabetes, a history of prior stroke or TIA, vascular disease, and proportion of females; the distribution across individual scores of CHA2DS2-VASc rule; and the proportion of patients using a platelet inhibitor;

– Validation study results: the c-statistic(s), the annual outcome risk per score, and corresponding estimates of uncertainty.

Data preparation

In accordance with previous recommendations, we rescaled the extracted c-statistic by applying the logit transformation.[17] If more than one c-statistic was reported, e.g. when calculated using aforementioned risk categories or one for a continuous score, the highest c-statistic was used. The error variance of the logit c-statistic was estimated from the reported confidence interval[18] or standard error (Delta method). If no information on uncertainty was reported, we used the approximation as reported by Debray et al.[17]

Furthermore, we rescaled annual stroke risk estimates by applying the square rooted transformation.[19] The corresponding variance was estimated using Poisson approximations and, again, applying the Delta method.[17]

Data analysis

We applied random effects meta-analysis using restricted maximum likelihood estimation (REML) to summarise estimates of model discrimination (logit c-statistic) and annual risk per score (square root risks).[17,20] In accordance with recent guidelines, confidence intervals were calculated using the Hartung-Knapp-Sidik-Jonkman method.[21] We calculated approximate 95% prediction intervals (95% PI) to ascertain the potential impact of between-study heterogeneity. These intervals indicate the range of performances (e.g. c-statistic or stroke rates per CHA2DS2-VASc score) that can be expected in future validation studies with similar characteristics as the ones included in our review. Additionally, we calculated the probability that the annual stroke risk was below a certain threshold if in 'real life' practice the CHA2DS2-VASc rule assigned an AF patient with a score 0, 1 or 2.[22]

Finally, we performed random effects meta-regression to investigate potential sources of heterogeneity. For study characteristics, we included the outcome under study, the number of person years of follow-up and the dichotomised risk of bias as covariates. For summarised patient characteristics, covariates of interest were mean age of the study population, proportion of females, mean CHA2DS2-VASc score, prevalence of heart failure and prevalence of platelet inhibitor use. See Supplemental Material 7. All analyses were performed with the package *metafor* (univariate meta-analysis) *1.9–8* in *R 3.3.0.*

RESULTS

Included studies

The process of study selection is shown in Figure 1. The initial search yielded 17,667 results, from which 8096 duplicates were discarded. After reading title and abstract we excluded 19,245 articles, primarily because these included patients outside the domain of interest (~60%), did not address risk prediction (~30%) or did not externally validate CHA2DS2-VASc (~5%).

In total, 126 studies were subjected to full text evaluation. Of these, inclusion criteria were not met in 107 studies, resulting in a final selection of 19 validation studies.

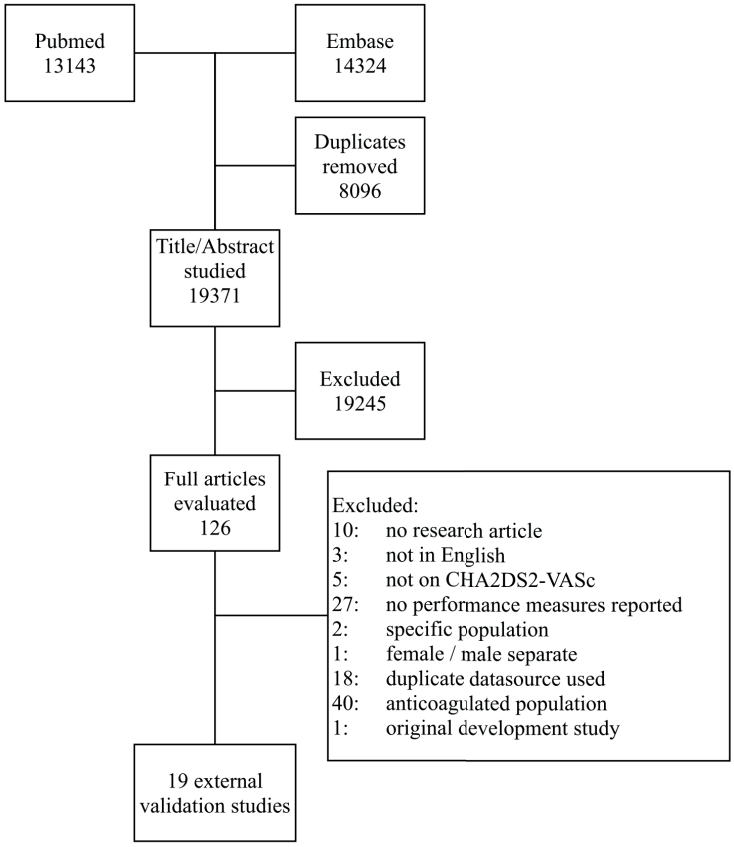


FIGURE 1. Process of study selection

The key characteristics of each study are presented in Table 2. Seven studies (in total 163,610 AF patients with in total 365,501 person years of follow-up) were performed in AF patients recruited from the general (unselected) population, and twelve studies (in total 683,138 AF patients and 1,738,930 person years follow-up) included a subsample of patients from the hospital care setting.

The outcomes under study consisted of 1) ischaemic stroke (ten studies), or 2) all thromboembolic events in eight studies, that is, ischaemic strokes and systemic thromboembolism (defined as peripheral embolism in six studies and peripheral embolism and/or pulmonary embolism in two studies), or 3) all types of stroke (ischaemic plus haemorrhagic stroke) in one study. Most studies originated from Europe and North America, and five were performed in East Asia. The number of included patients ranged from 154 to 198,697 and the follow-up time from 11 to 53 months (Table 2).

TABLE 2. Overview of the included studies

Author	Year	Location	Setting	Outcome	N	% on antiplatelet therapy	Months of follow-up	Incidence per 100 person-years
Hobbs [23]	2011	UK	Non-selected	ischaemic stroke	665	55.3	26.4	3.6
Olesen [24]	2011	Denmark	Hospital	TE	73538	34.7	12.0	7.6
Sandhu [25]	2011	Canada	Hospital	all stroke incl. ICH	4476	n.r.	12.0	6.2
Abu-Assi [26]	2013	Spain	Non-selected	TE	186	81.6	42.7	2.2
Guo [27]	2013	China	Hospital	TE	885	79.0	22.8	3.7
Singer [28]	2013	USA	Non-selected	TE	10927	n.r.	35.8	2.1
Forslund [29]	2014	Sweden	Non-selected	ischaemic stroke	9959	0.0	12.0	2.0
Komatsu [30]	2014	Japan	Hospital	TE	332	33.0	53.0	2.1
Siu [31]	2014	China	Hospital	ischaemic stroke	3881	0.0	38.4	9.3
Abumuaileq [32]	2015	Spain	Hospital	TE	154	97.4	11.0	5.8
Saliba [33]	2015	Israel	Non-selected	ischaemic stroke	41140	n.r.	11.1	4.5
Suzuki [34]	2015	Japan	Hospital	ischaemic stroke	3588	41.8	16.8	1.3
van den Ham [35]	2015	UK	Non-selected	ischaemic stroke	60594	n.r.	25.2	3.0
Aspberg [36]	2016	Sweden	Hospital	ischaemic stroke	152153	n.r.	26.4	3.3
Chao [37]	2016	Taiwan	Hospital	ischaemic stroke	186570	0.0	40.8	3.7
Nielsen [38]	2016	Denmark	Hospital	ischaemic stroke	198697	n.r.	29.0	3.2
Xing [39]	2016	China	Hospital	TE	413	68.3	23.9	14.3
Allan [40]	2017	UK	Non-selected	ischaemic stroke	40139	n.r.	26.4	3.8
McAlister [41]	2017	Canada	Hospital	TE	58451	n.r.	31.0	4.2

TE=thromboembolism; ICH=intra-cranial haemorrhage; n.r.=not reported

Risk of bias

The risk of bias of the included studies is summarised in Supplemental Material 4. For details on individual studies, see Supplemental Material 5. Overall, two of seven studies performed in the general population were considered at risk of bias, as was the majority of studies enrolling patients from a hospital care setting. In general, the source of the data and the eligibility criteria caused no concern for bias. Some studies did not provide information on the use of antiplatelet therapy. This could induce biased results on the predictive accuracy of CHA2DS2-VASc since, albeit to a limited extent, antiplatelet therapy may reduce the occurrence of stroke and thus underestimate the predictive accuracy of the rules.[3] The definition and measurement of the predictors – i.e. the variables included in the CHA2DS2-VASc rule – and the outcome under study frequently differed across studies. Mostly, these were clearly defined. Predictors were mostly assessed blinded for the outcome. No study explicitly reported whether the outcomes were assessed blinded for the initial CHA2DS2-VASc score, potentially introducing bias for outcomes requiring subjective interpretation such as TIA.[13]

In six studies (two in unselected and four in selected patients) the number of outcome events was lower than the generally recommended ~100 events for validation of a decision rule.[42] In addition, the amount and handling of missing data was unclear in the majority of studies. As data are seldom missing completely at random, inadequate handling of missing data could introduce bias.[43–45]

Meta-analysis of discriminative ability

In both populations, there was substantial between-study heterogeneity. In studies enrolling patients from the general population, we found an average c-statistic of 0.64 (95% CI 0.56 to 0.71). The variation in discriminative performance of CHA2DS2-VASc across studies is indicated in Supplemental Material 6 and reflected by the wide approximate 95% prediction interval (95% PI) that ranged from 0.45 to 0.79.

In studies recruiting from a hospital care setting, we found a somewhat higher average c-statistic of 0.71 (95% CI 0.62 to 0.79). Again, the 95% PI was wide, ranging from 0.40 to 0.90. Based on 'eye-balling, we identified three outlying studies. When we excluded two studies with high c-statistics,[24,30] summary estimates were similar to the results of studies in general population settings. After excluding one study with a low c-statistic[31] or all three outlying studies, discrimination remained highest in studies recruiting from hospitals, with a lower point estimate but more narrow 95% prediction intervals (data not shown).

Meta-analysis of stroke risk per score

Figure 2 shows the forest plots with the annual stroke risks and/or systemic thromboembolism for the scores 0 to 3.

For every score on CHA2DS2-VASc, there was substantial heterogeneity in both settings of care with wide approximate 95% prediction intervals. In studies enrolling patients from the general population for example, the stroke risk for a CHA2DS2-VASc score of zero in a new validation study could lie between 0.0% and 0.9%. For score one and two, these were 0.3% and 1.2%, and 0.4 and 3.3%, respectively. See Figure 2a.

Studies recruiting from a hospital care setting showed a more diverse distribution of risks per score, with higher pooled annual risks for all scores. For instance, the annual risk for a CHA2DS2-VASc score of 1 was 1.4% (approximate 95% PI 0.04 to 6.5%) in these studies, compared to 0.7% (approximate 95% PI 0.3 to 1.2%) in studies enrolling AF patients from the general population. See Figure 2b. Although excluding one outlier[31] of the hospital-based studies in a sensitivity analysis resulted in a lower risk for CHA2DS2-VASc score 1 of 1.2%, this was still nearly twice as high as compared to the pooled estimate of studies enrolling patients from the general population. Furthermore, the prediction interval remained wide, ranging from 0.06 to 3.8. The differences between study populations did not sufficiently explain heterogeneity. See Figure 2c.

Excluding one study that recruited patients all 75 years of age and older[23] in an additional sensitivity analysis did not change the results. Two studies sampled from the same CPRD data source and including the same patients multiple times in our meta-analysis cannot be ruled-out. Excluding either the study by van den Ham[35] or by Allan[40] did not change our results (data not shown).

To further illustrate the interpretation of pooled stroke risks and their uncertainty (due to estimation error and heterogeneity), we calculated the probability that patients with a certain CHA2DS2-VASc score have an annual stroke risk below 1%. For patients recruited from the general population, this probability was 98% (score 0), 91% (score ≤ 1) and 19% (score ≤ 2). For patients recruited from a hospital care setting, these probabilities dropped to 71%, 39% and, respectively, 17%.

Meta-regression and best available evidence

To further explore sources of heterogeneity in both discrimination and the stroke risk per score, we performed several meta-regression analyses. These demonstrate that it was difficult to identify any relevant sources, as regression coefficients for risk of bias, study characteristics or summarised patient characteristics were not statistically significant. Furthermore, we explored heterogeneity between studies considered as 'best available evidence'. See Supplemental Material 7.

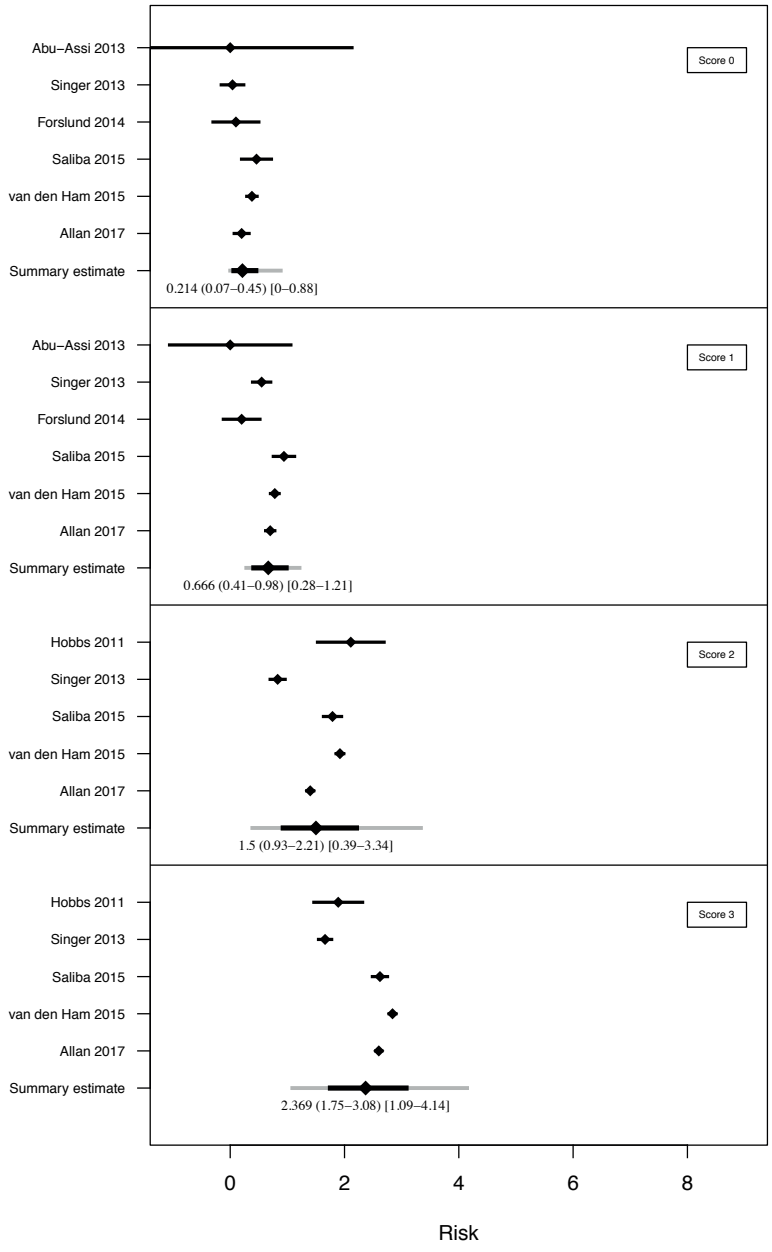


FIGURE 2a. Stroke risk (events per 100 person-years) per score in studies recruiting from the general population
Solid bars represent 95% confidence intervals
Dashed bars represent 95% confidence intervals, estimated
Grey bars represent 95% prediction intervals
Summary estimate is c-statistic (95% CI) [95% PI]

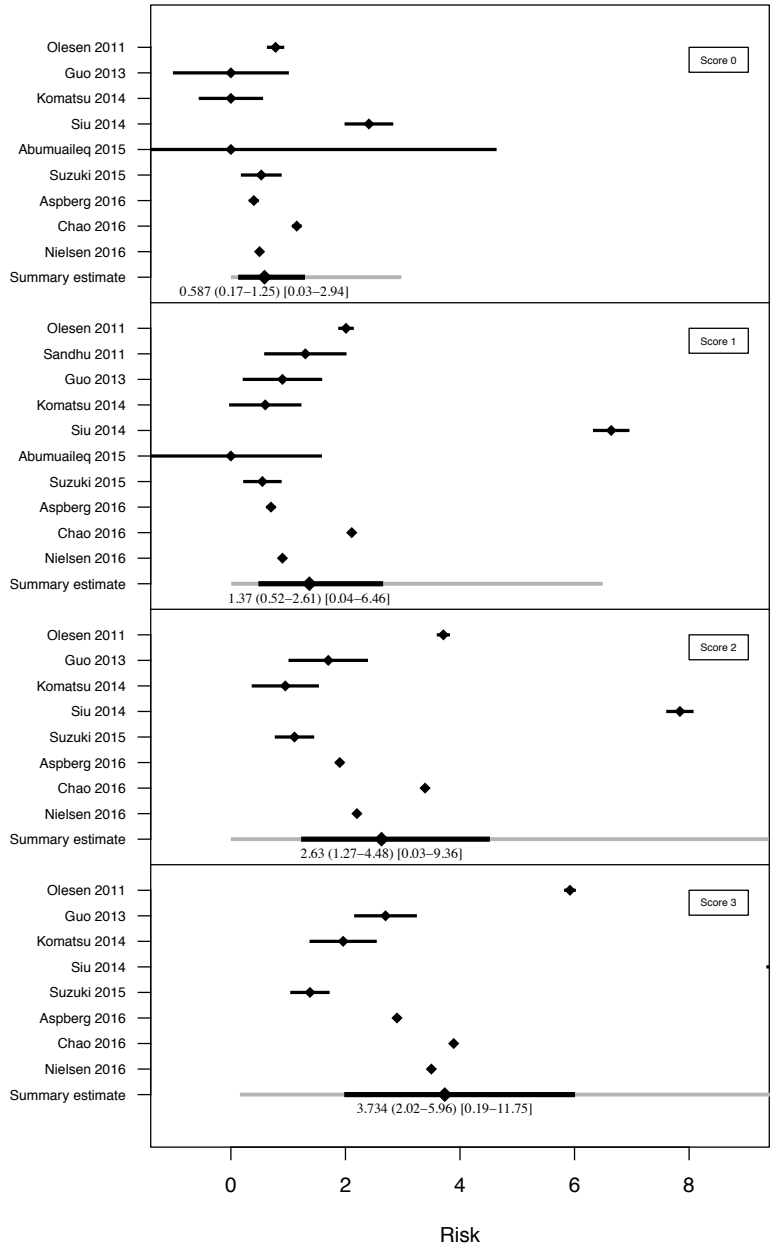


FIGURE 2b. Stroke risk (events per 100 person-years) per score in studies recruiting from hospitals
Solid bars represent 95% confidence intervals
Dashed bars represent 95% confidence intervals, estimated
Grey bars represent 95% prediction intervals
Summary estimate is c-statistic (95% CI) [95% PI]

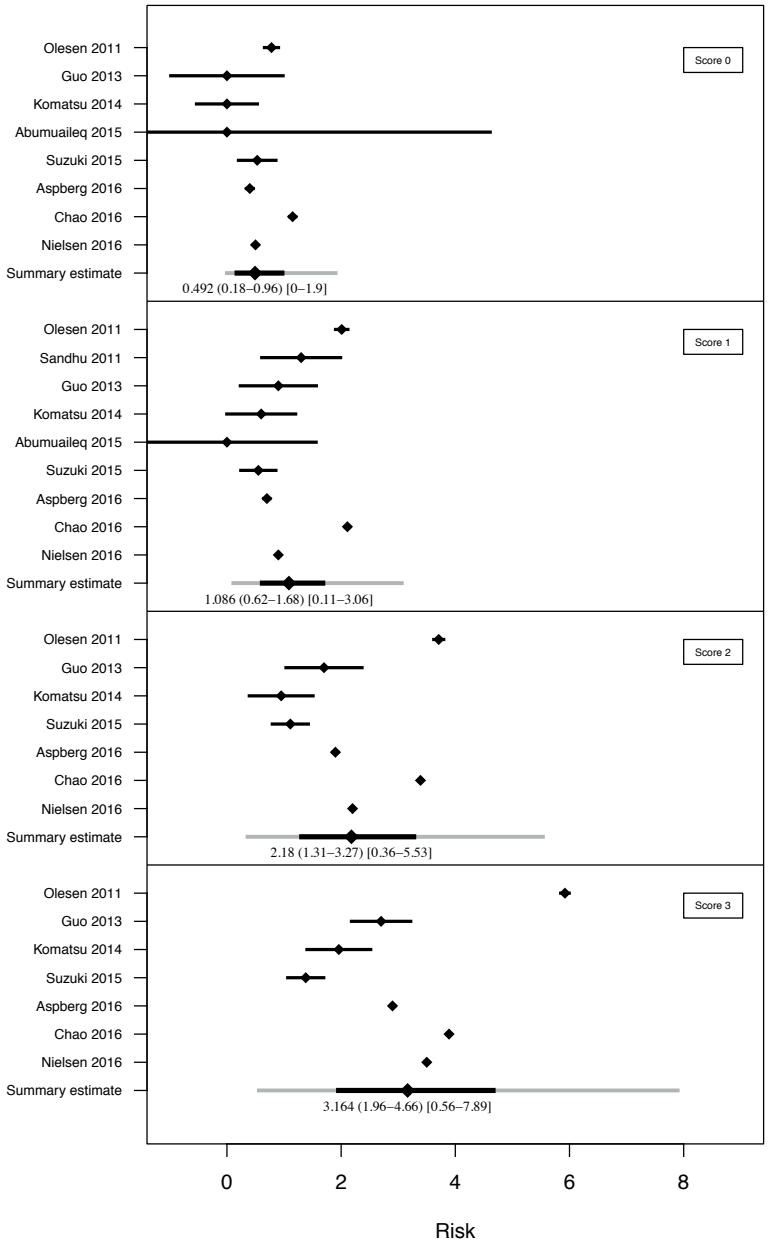


FIGURE 2c. Stroke risk (events per 100 person-years) per score in studies recruiting from hospitals, excluding outliers
Solid bars represent 95% confidence intervals
Dashed bars represent 95% confidence intervals, estimated
Grey bars represent 95% prediction intervals
Summary estimate is c-statistic (95% C.I.) [95% P.I.]

DISCUSSION

This systematic review and meta-analysis thoroughly explores heterogeneity in the results of all currently available validation studies of CHA2DS2-VASc. Our analysis confirms that most validation studies of CHA2DS2-VASc yield conflicting results, with highly variable estimates for stroke risk per score. This heterogeneity partly appears to arise from population or case-mix differences across the validation studies, as stratified analyses showed lower stroke risk estimates for studies enrolling patients from the open general population as compared to studies using hospital-based recruitment strategies. Yet, substantial between-study heterogeneity remained and could not be resolved by adjusting for differences in study characteristics, differences in risk of bias, or other differences in population characteristics.

Strengths and limitations

A major strength of this study is that we applied rigorous and state-of-the art systematic review and quantitative synthesis. Previous studies[46,47] found a similar modest discrimination of CHA2DS2-VASc, but did not provide stroke risk for each score or explored potential sources for heterogeneity such as case-mix differences. Another recent publication[48] also found substantial variation in stroke risks, but could not identify any source of this heterogeneity. Our systematic literature review and meta-analysis adds the following additional inferences. First, the study by Quinn does not provide a summarised estimate per CHA2DS2-VASc score, and the observed heterogeneity was summarised in indices of an I^2-value or the Q-statistic. We did provide such summarised estimates per CHA2DS2-VASc-score and report the heterogeneity around these point estimates with 95% prediction intervals that are easily appreciated in clinical practice. Furthermore, we used the CHARMS checklist specifically designed to appraise prediction modelling studies and assess their risk of bias. Lastly, in contrast to previous studies, we additionally calculated a summarised estimate of the c-statistic as a measure of discrimination.

Also, the NICE guidelines[49] provide a formal evaluation of CHA2DS2-VASc, and the score provided in a cost-effective method to estimate stroke risk and indicate anticoagulant treatment. However, stroke risks were based on a single study,[50] no summary estimate was provided and subsequently a measure of uncertainty was lacking.

Nevertheless, for full appreciation of our results several issues should be considered. The CHA2DS2-VASc rule has been advocated as superior to its precursor CHADS2 in identifying AF patients at 'truly low' risk of stroke, in particular for low scores on CHA2DS2-VASc (0 or 1). However, prompted by validation studies of CHA2DS2-VASc showing widely varying results, there has been much debate

on what score (in particular 0 or 1) truly defines low risk. This is reflected in our meta-analysis. Indeed, we found that future validation studies where patient enrolment starts in hospital care may observe a very low stroke risk for CHA2DS2-VASc score zero (well below 1%; the threshold above which it is often advocated that the benefits of anticoagulants outweigh the bleeding risk).[51,52] However, patients with a score of one, two or even three may also found to have such a low stroke risk. Conversely, future studies may also find an in fact high stroke risk (e.g. above 3%) already in patients with score zero, which we believe explains the confusion and recent debate on what score denotes a low risk.[8,9,53–55]

We undertook many efforts to explore sources of this large extent of between-study heterogeneity, and several issues require further inspection. First, our results suggest that differences in risk of bias do not play an important role when summarizing estimates of prediction model performance. Although several validation studies showed shortcomings, differences were small and even the most homogeneous group of studies at low risk of bias showed conflicting results in stroke risk per score.

Second, some heterogeneity could be explained by differences in case-mix across the validation studies. We observed a clinically relevant higher stroke risk in studies recruiting AF patients from a hospital care setting compared to those enrolling from the general population. It is possible that AF patients recruited from the hospital care setting represent more diseased patients with a higher baseline risk of stroke, independent of their overall CHA2DS2-VASc score. As an example, the type and burden of AF (paroxysmal, persistent or permanent) may have an association with stroke risk,[56] as may the severity and duration of a patient's individual CHA2DS2-VASc risk factors, or risk factors not included in the CHA2DS2-VASc score such as renal failure.[57] Case-mix differences in such risk factors between patients sampled from hospital care and from the general population could, at least partly, explain the variation in validation studies and the observed difference in stroke risk in our results.

However, third, although we stratified studies to the clinical setting where patient enrolment started, this strategy did not sufficiently help to explain all heterogeneity across the validation studies. Adjustment for differences in study characteristics through meta-regression did not much affect the extent of between-study heterogeneity. For instance, the definition of the outcome under study – only ischaemic stroke; ischaemic stroke and thromboembolism or indeed even any type of stroke including intra-cranial haemorrhage – could potentially influence the risks for each CHA2DS2-VASc score but including the different outcomes as a covariate in the meta-regression model did not affect the results. Similarly, additional summarised population characteristics such as mean CHA2DS2-VASc score or use

of platelet inhibitors did not account for the heterogeneous results. Importantly, we could not evaluate which combination of predictors contributed to a patient's CHA2DS2-VASc score. It is believed that not all predictors in the CHA2DS2-VASc decision rule carry the same stroke risk[58,59] although this is not acknowledged in the rule as almost all its predictors contribute 1 point. Likewise, females with a score 1 (i.e. no other risk factor) are likely to be at lower risk than males with 1 risk factor. However, the included validation studies often do not report stratified analyses for males and females and thus in this meta-analysis of aggregated data we were unable to account for this. Individual patient data meta-analysis would be needed to fully clarify issues such as stroke risk in females with no additional stroke risk factors.

Fourth, the age categories of the CHA2DS2-VASc are broad and thus a patient aged 65 will receive the same score as a 74 years old person for the 'Age' category, though stroke risk will likely not be equal. This also results in heterogeneity and variation found in the results of validation studies

Fifth, it has previously been shown that ethnicity has an effect on stroke risk[60,61] and stroke mortality[62]. Unfortunately we did not have enough data on ethnicity to consider this variable in our meta-analysis. We did, however, include the geographic region as a covariate in the meta-regression model (data not shown), and in line with the findings by Quinn et al.[48] this could not sufficiently explain the large heterogeneity.

Sixth, we were only able to meta-analyse the stroke risk per score and the c-statistic, as these were the measures most often and consistently reported in current validation studies. Other measures such as decision curve analysis have been proposed to better investigate the clinical value of using a decision rule.[63,64] Unfortunately, these aspects of model performance are rarely consistently quantified. This renders meta-analysing such measures difficult, if not (yet) impossible, and can be considered a limitation to this current systematic review and meta-analysis of pooled c-statistics and strokes risks per score.

Furthermore, it is important to consider that we only included validation studies where patients did not (yet) use anticoagulants. By itself, the choice of including only patients not on anticoagulation could result in a selected subtype of AF patients as this may have led to confounding 'by contra-indication'[65] when anticoagulation is withheld due to (for instance) severe illness or bleeding risk. We however deliberately chose not to include populations already on anticoagulants as CHA2DS2-VASc is intended to be used for stroke risk prediction *prior* to anticoagulant treatment decisions in AF patients.

Finally, it should be noted though that, essentially, heterogeneity is not uncommon for studies validating diagnostic or prognostic decision rules, also in the field

of thrombotic disorders. Clinical decision rules like CHA2DS2-VASc are popular in daily clinical practice because they are helpful and easy-to-apply methods to tailor subsequent treatment decisions. Indeed, for younger patients without additional stroke risk factors (score of 0) and for more 'high-risk' patients (score of around 2–3 or above), CHA2DS2-VASc clearly facilitates in anticoagulant treatment recommendations. However, based on synthesis of the current evidence in the literature, tailoring treatment for patients at 'low to intermediate risk' (i.e. CHA2DS2-VASc scores 1 to 2) remains ambiguous, as is also reflected by the discussion in literature on the optimal threshold for initiating anticoagulant treatment.[8–12]

Clinical implications and future research questions

In the treatment of atrial fibrillation, adequate identification of different stroke risk groups is pivotal for anticoagulant treatment decisions. The main inference of our meta-analysis is that – albeit a simple, effective and easy-to-use tool at truly low and truly high risk patients – CHA2DS2-VASc may have difficulties in tailoring anticoagulant treatment adequately in AF patients at intermediate risk of stroke (roughly those with a score of 1 or 2). Differences in stroke risks between studies recruiting from hospitals or from the general population indicate that possible case-mix differences between populations should be taken into account in clinical decision-making but further uncertainty remains. Future research should focus on this issue, with further model revision, and considering additional co-morbidity items (e.g. renal impairment), (novel) biomarkers or imaging such as left atrial wall remodelling patterns in addition to existing prediction models for quantifying the thrombotic risk in AF patients.[66–68] These additional tests may ultimately result in a better clinical decision. Whether this can be achieved should be the focus of further investigation.

CONCLUSIONS

Studies validating CHA2DS2-VASc demonstrate high heterogeneity in predicted stroke risks for different scores.

ACKNOWLEDGEMENTS

SD had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. SD performed the literature search. SD, FK and GJG performed critical appraisal. SD and FK performed data

extraction and risk of bias assessment. TPAD and SD performed data analysis. All authors interpreted the data. SD, TPAD, AWH, FHR, KGMM and GJG wrote and critically reviewed the manuscript.

DISCLOSURES

GJG and FHR received an unrestricted institutional grant from Boehringer Ingelheim. KGMM received a grant from The Netherlands Organization for Scientific Research (ZONMW 918.10.615 and 91208004). GJG and TPAD are supported by a VENI grant from The Netherlands Organization for Scientific Research (ZONMW 016.166.030 and, respectively, 916.17.050). All funding sources had no role in the design, conduct, analyses, or reporting of the study or in the decision to submit the manuscript for publication. All other authors declared no conflict of interest.

REFERENCES

1 Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, Singer DE. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* 2001; 285: 2370–5.

2 Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991; 22: 983–8.

3 Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Ann Intern Med* 2007; 146: 857–67.

4 Lin L, Lim WS, Zhou HJ, Khoo AL, Tan KT, Chew AP, Foo D, Chin JJ, Lim BP. Clinical and Safety Outcomes of Oral Antithrombotics for Stroke Prevention in Atrial Fibrillation: A Systematic Review and Network Meta-analysis. *J Am Med Dir Assoc* 2015; 16: 1103.e1–19.

5 Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. 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.

6 Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of Clinical Classification Schemes for Predicting Stroke: Results From the National Registry of Atrial Fibrillation. *JAMA* 2001; 285: 2864–70.

7 European Heart Rhythm Association, European Association for Cardio-Thoracic Surgery, Camm AJ, Kirchhof P, Lip GYH, Schotten U, Savelieva I, Ernst S, Van Gelder IC, Al-Attar N, Hindricks G, Prendergast B, Heidbuchel H, Alfieri O, Angelini A, Atar D, Colonna P, De Caterina R, De Sutter J, Goette A, et al. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). *Eur Heart J* 2010; 31: 2369–429.

8 Savino JA, Halperin JL. Should Patients With Atrial Fibrillation and 1 Stroke Risk Factor (CHA2DS2-VASc Score 1 in Men, 2 in Women) Be Anticoagulated?Response to Savino and Halperin. *Circulation* 2016; 133: 1504–11.

9 Lip GYH, Nielsen PB. Should Patients With Atrial Fibrillation and 1 Stroke Risk Factor (CHA2DS2-VASc Score 1 in Men, 2 in Women) Be Anticoagulated?Response to Lip and Nielsen. *Circulation* 2016; 133: 1498–503.

10 Friberg L, Skeppholm M, Terént A. Benefit of Anticoagulation Unlikely in Patients With Atrial Fibrillation and a CHA2DS2-VASc Score of 1. *J Am Coll Cardiol* 2015; 65: 225–32.

11 Huisman MV. Patients With Atrial Fibrillation and a CHA2DS2-VASc Score of 1: Are They at Low or High Stroke Risk? *J Am Coll Cardiol* 2015; 65: 1395–7.

12 Olesen JB, Torp-Pedersen C. Stroke risk in atrial fibrillation: Do we anticoagulate CHADS2 or CHA2DS2-VASc ≥ 1, or higher? *Thromb Haemost* 2015; 113: 1165–9.

13 Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, Reitsma JB, Collins GS. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. *PLoS Med* 2014; 11: e1001744.

14 Steyerberg E. Clinical Prediction Models. 1st ed. Springer Science & Business Media; 2008.

15 Ingui BJ, Rogers MA. Searching for clinical prediction rules in MEDLINE. *J Am Med Inform Assoc* 2001; 8: 391–7.

16 Geersing G-J, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons KGM. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. *PLoS One* 2012; 7: e32844.

17 Debray TPA, Damen JAAG, Snell KIE, Ensor J, Hooft L, Reitsma JB, Riley RD, Moons KGM. A guide to systematic review and meta-analysis of prediction model performance. *BMJ* 2017; 356: i6460.
18 Altman DG, Bland JM. How to obtain the P value from a confidence interval. *BMJ* 2011; 343: d2304–4.
19 Trikalinos TA, Trow P, Schmid CH. Simulation-Based Comparison of Methods for Meta-Analysis of Proportions and Rates. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013.
20 van Klaveren D, Steyerberg EW, Perel P, Vergouwe Y. Assessing discriminative ability of risk models in clustered data. *BMC Med Res Methodol* 2014; 14: 5.
21 IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol* 2014; 14: 1.
22 Snell KIE, Hua H, Debray TPA, Ensor J, Look MP, Moons KGM, Riley RD. Multivariate meta-analysis of individual participant data helped externally validate the performance and implementation of a prediction model. *J Clin Epidemiol* 2016; 69: 40–50.
23 Hobbs FD, Roalfe AK, Lip GYH, Fletcher K, Fitzmaurice DA, Mant J. Performance of stroke risk scores in older people with atrial fibrillation not taking warfarin: comparative cohort study from BAFTA trial. *BMJ* 2011; 342: d3653.
24 Olesen JB, Lip GYH, Hansen ML, Hansen PR, Tolstrup JS, Lindhardsen J, Selmer C, Ahlehoff O, Olsen A-MS, Gislason GH, Torp-Pedersen C. Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study. *BMJ* 2011; 342: d124.
25 Sandhu RK, Bakal JA, Ezekowitz JA, McAlister FA. Risk stratification schemes, anticoagulation use and outcomes: the risk—treatment paradox in patients with newly diagnosed non-valvular atrial fibrillation. *Heart* 2011; 97: 2046–50.
26 Abu-Assi E, Otero-Ravina F, Allut Vidal G, Coutado Mendez A, Vaamonde Mosquera L, Sanchez Loureiro M, Caneda Villar MC, Fernandez Villaverde JM, Maestro Saavedra FJ, Gonzalez-Juanatey JR, Grupo Barbanza researchers. Comparison of the reliability and validity of four contemporary risk stratification schemes to predict thromboembolism in non-anticoagulated patients with atrial fibrillation. *Int J Cardiol* 2013; 166: 205–9.
27 Guo Y, Apostolakis S, Blann AD, Wang H, Zhao X, Zhang Y, Zhang D, Ma J, Wang Y, Lip GYH. Validation of contemporary stroke and bleeding risk stratification scores in non-anticoagulated Chinese patients with atrial fibrillation. *Int J Cardiol* 2013; 168: 904–9.
28 Singer DE, Chang Y, Borowsky LH, Fang MC, Pomernacki NK, Udaltsova N, Reynolds K, Go AS. A new risk scheme to predict ischemic stroke and other thromboembolism in atrial fibrillation: the ATRIA study stroke risk score. *J Am Heart Assoc* 2013; 2: e000250–0.
29 1 Forslund T, Wettermark B, Wandell P, Euler von M, Hasselstrom J, Hjemdahl P. Risks for stroke and bleeding with warfarin or aspirin treatment in patients with atrial fibrillation at different CHA(2)DS(2)VASc scores: experience from the Stockholm region. *Eur J Clin Pharmacol* 2014; 70: 1477–85.
30 Komatsu T, Sato Y, Ozawa M, Kunugita F, Yoshizawa R, Morino Y, Nakamura M. Comparison between CHADS2 and CHA2DS2-VASc score for risk stratification of ischemic stroke in Japanese patients with non-valvular paroxysmal atrial fibrillation not receiving anticoagulant therapy. *Int Heart J* 2014; 55: 119–25.

31 Siu C-W, Lip GYH, Lam K-F, Tse H-F. Risk of stroke and intracranial hemorrhage in 9727 Chinese with atrial fibrillation in Hong Kong. *Heart Rhythm* 2014; 11: 1401–8.

32 Abumuaileq RRY, Abu-Assi E, López-López A, Raposeiras-Roubin S, Rodríguez-Mañero M, Martínez-Sande L, García-Seara J, Fernandez-López XA, Peña-Gil C, González-Juanatey JR. Comparison between CHA2DS2-VASc and the new R2CHADS2 and ATRIA scores at predicting thromboembolic event in non-anticoagulated and anticoagulated patients with non-valvular atrial fibrillation. *BMC Cardiovasc Disord* 2015; 15: 156.

33 Saliba W, Barnett-Griness O, Elias M, Rennert G. The Association Between Red Cell Distribution Width and Stroke in Patients with Atrial Fibrillation. *Am J Med* 2015; 128: 192.e11–8.

34 Suzuki S, Yamashita T, Okumura K, Atarashi H, Akao M, Ogawa H, Inoue H. Incidence of Ischemic Stroke in Japanese Patients With Atrial Fibrillation Not Receiving Anticoagulation Therapy. *Circ J* 2015; 79: 432–8.

35 van den Ham HA, Klungel OH, Singer DE, Leufkens HGM, van Staa TP. Comparative Performance of ATRIA, CHADS2, and CHA2DS2-VASc Risk Scores Predicting Stroke in Patients With Atrial Fibrillation: Results From a National Primary Care Database. *J Am Coll Cardiol* 2015; 66: 1851–9.

36 Aspberg S, Chang Y, Atterman A, Bottai M, Go AS, Singer DE. Comparison of the ATRIA, CHADS2, and CHA2DS2-VASc stroke risk scores in predicting ischaemic stroke in a large Swedish cohort of patients with atrial fibrillation. *Eur Heart J* 2016; 37: 3203–10.

37 Chao T-F, Liu C-J, Tuan T-C, Chen S-J, Wang K-L, Lin Y-J, Chang S-L, Lo L-W, Hu Y-F, Chen T-J, Chiang C-E, Chen S-A. Comparisons of CHADS2 and CHA2DS2-VASc scores for stroke risk stratification in atrial fibrillation: Which scoring system should be used for Asians? *Heart Rhythm* 2016; 13: 46–53.

38 Nielsen PB, Larsen TB, Skjøth F, Overvad TF, Lip GYH. Stroke and thromboembolic event rates in atrial fibrillation according to different guideline treatment thresholds: A nationwide cohort study. *Sci Rep* 2016; 6: 27410.

39 Xing Y, Ma Q, Ma X, Wang C, Zhang D, Sun Y. CHADS2 score has a better predictive value than CHA2DS2-VASc score in elderly patients with atrial fibrillation. *Clin Interv Aging* 2016; 11: 941–6.

40 Allan V, Banerjee A, Shah AD, Patel R, Denaxas S, Casas J-P, Hemingway H. Net clinical benefit of warfarin in individuals with atrial fibrillation across stroke risk and across primary and secondary care. *Heart* 2017; 103: 210–8.

41 McAlister FA, Wiebe N, Jun M, Sandhu R, James MT, McMurtry MS, Hemmelgarn BR, Tonelli M. 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: 243–52.

42 Vergouwe Y, Steyerberg EW, Eijkemans MJC, Habbema JDF. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol* 2005; 58: 475–83.

43 Gorelick MH. Bias arising from missing data in predictive models. *J Clin Epidemiol* 2006; 59: 1115–23.

44 Marshall A, Altman DG, Royston P, Holder RL. Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. *BMC Med Res Methodol* 2010; 10: 7.

45 Donders ART, van der Heijden GJMG, Stijnen T, Moons KGM. Review: A gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006; 59: 1087–91.

46 Chen J-Y, Zhang A-D, Lu H-Y, Guo J, Wang F-F, Li Z-C. CHADS2 versus CHA2DS2-VASc score in assessing the stroke and thromboembolism risk stratification in patients with atrial fibrillation: a systematic review and meta-analysis. *J Geriatr Cardiol* 2013; 10: 258–66.
47 Joundi RA, Cipriano LE, Sposato LA, Saposnik G, Group OBOTSORW. Ischemic Stroke Risk in Patients With Atrial Fibrillation and CHA2DS2-VASc Score of 1 Systematic Review and Meta-Analysis. *Stroke* 2016; 47: 1364–7.
48 Quinn GR, Severdija ON, Chang Y, Singer DE. Wide Variation in Reported Rates of Stroke Across Cohorts of Patients With Atrial Fibrillation. *Circulation* 2017; 135: 208–19.
49 NICE. Atrial fibrillation: the management of atrial fibrillation. (Clinical guideline 180). NICE; 2014.
50 Friberg L, Rosenqvist M, Lip GYH. Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: the Swedish Atrial Fibrillation cohort study. *Eur Heart J* 2012; 33: 1500–10.
51 Eckman MH, Singer DE, Rosand J, Greenberg SM. Moving the tipping point: the decision to anticoagulate patients with atrial fibrillation. *Circ Cardiovasc Qual Outcomes* 2011; 4: 14–21.
52 Singer DE, Chang Y, Fang MC, Borowsky LH, Pomernacki NK, Udaltsova N, Go AS. The net clinical benefit of warfarin anticoagulation in atrial fibrillation. *Ann Intern Med* 2009; 151: 297–305.
53 Cairns JA, Healey JS, Macle L, Mitchell LB, Verma A. The New Canadian Cardiovascular Society Algorithm for Antithrombotic Therapy of Atrial Fibrillation Is Appropriately Based on Current Epidemiologic Data. *Can J Cardiol* 2015; 31: 20–3.
54 Lip GYH, Nielsen PB, Skjøth F, Rasmussen LH, Larsen TB. Atrial Fibrillation Patients Categorized as "Not for Anticoagulation" According to the 2014 Canadian Cardiovascular Society Algorithm Are Not 'Low Risk'. *Can J Cardiol* 2015; 31: 24–8.
55 Nielsen PB, Larsen TB, Skjøth F, Gorst-Rasmussen A, Rasmussen LH, Lip GYH. Restarting Anticoagulant Treatment After Intracranial Hemorrhage in Patients With Atrial Fibrillation and the Impact on Recurrent Stroke, Mortality, and Bleeding A Nationwide Cohort Study. *Circulation* 2015; 132: 517–25.
56 Steinberg BA, Hellkamp AS, Lokhnygina Y, Patel MR, Breithardt G, Hankey GJ, Becker RC, Singer DE, Halperin JL, Hacke W, Nessel CC, Berkowitz SD, Mahaffey KW, Fox KAA, Califf RM, Piccini JP, ROCKET AF Steering Committee and Investigators. Higher risk of death and stroke in patients with persistent vs. paroxysmal atrial fibrillation: results from the ROCKET-AF Trial. *Eur Heart J* 2015; 36: 288–96.
57 Olesen JB, Lip GYH, Kamper A-L, Hommel K, Kober L, Lane DA, Lindhardsen J, Gislason GH, Torp-Pedersen C. Stroke and bleeding in atrial fibrillation with chronic kidney disease. *N Engl J Med* 2012; 367: 625–35.
58 Chao T-F, Liu C-J, Wang K-L, Lin Y-J, Chang S-L, Lo L-W, Hu Y-F, Tuan T-C, Chen T-J, Lip GYH, Chen S-A. Should Atrial Fibrillation Patients With 1 Additional Risk Factor of the CHA2DS2-VASc Score (Beyond Sex) Receive Oral Anticoagulation? *J Am Coll Cardiol* 2015; 65: 635–42.
59 Huang D, ANGUO L, YUE W-S, YIN L, Tse H-F, Siu C-W. Refinement of ischemic stroke risk in patients with atrial fibrillation and CHA2 DS2 -VASc score of 1. *Pacing Clin Electrophysiol* 2014; 37: 1442–7.
60 Chang K-C, Wang Y-C, Ko *P*-Y, Wu H-P, Chen Y-W, Muo C-H, Sung F-C, Li T-C, Hsu CY. Increased Risk of First-Ever Stroke in Younger Patients With Atrial Fibrillation Not Recom-

mended for Antithrombotic Therapy by Current Guidelines: A Population-Based Study in an East Asian Cohort of 22 Million People. *Mayo Clinic Proceedings* 2014; 89: 1487–97.

61 Hajat C, Tilling K, Stewart JA, Lemic-Stojcevic N, Wolfe CDA. Ethnic differences in risk factors for ischemic stroke: a European case-control study. *Stroke* 2004; 35: 1562–7.

62 Ayala C, Greenlund KJ, Croft JB, Keenan NL, Donehoo RS, Giles WH, Kittner SJ, Marks JS. Racial/ethnic disparities in mortality by stroke subtype in the United States, 1995-1998. *Am J Epidemiol* 2001; 154: 1057–63.

63 Vickers AJ. Decision analysis for the evaluation of diagnostic tests, prediction models, and molecular markers. *Am Stat* 2008; 62: 314-20.

64 Steyerberg EW, Pencina MJ, Lingsma HF, Kattan MW, Vickers AJ, Van Calster B. Assessing the incremental value of diagnostic and prognostic markers: a review and illustration. *Eur J Clin Invest* 2012; 42: 216–28.

65 Walker AM. Confounding by indication. *Epidemiology* 1996; 7: 335–6.

66 Hijazi Z, Oldgren J, Siegbahn A, Granger CB, Wallentin L. Biomarkers in atrial fibrillation: a clinical review. *Eur Heart J* 2013; 34: 1475–80.

67 Wu N, Chen X, Cai T, Wu L, Xiang Y, Zhang M, Li Y, Song Z, Zhong L. Association of Inflammatory and Hemostatic Markers With Stroke and Thromboembolic Events in Atrial Fibrillation: A Systematic Review and Meta-analysis. *Can J Cardiol* 2015; 31: 278–86.

68 Providencia R, Botelho A, Trigo J, Quintal N, Nascimento J, Mota P, Leitão-Marques A. Possible refinement of clinical thromboembolism assessment in patients with atrial fibrillation using echocardiographic parameters. EP Europace 2012; 14: 36–45.SUPPLEMENTAL MATERIAL

SUPPLEMENTAL MATERIAL 1.

The annual risks of thromboembolism (ischaemic stroke, peripheral embolism, or pulmonary embolism) for CHA2DS2-VASc, adjusted for aspirin use[1]

Score	CHA2DS2-VASc (events/persons)
0	0 (0/103)
1	0.7 (1/162)
2	1.9 (3/184)
3	4.7 (8/203)
4	2.3 (4/208)
5	3.9 (3/95)
6	4.5 (2/57)
7	10.1 (2/25)
8	14.2 (1/9)
9	100 (1/1)

The original study deriving the CHA2DS2-VASc consisting of 1084 AF patients with a follow-up of one year, considering ischaemic stroke, peripheral embolism, or pulmonary embolism as outcomes for thromboembolism.

SUPPLEMENTAL MATERIAL 2. SEARCH STRATEGY

Search string PubMed:

(((Validat* OR Predict*.ti. OR Rule*) OR (Predict* AND (Outcome* OR Risk* OR Model*)) OR ((History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*)) OR (Decision* AND (Model* OR Clinical* OR Logistic Models/)) OR (Prognostic AND (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor* OR Model*))) OR (stratification OR ("ROC Curve"[Mesh]) OR discrimination OR discriminate OR c-statistic OR (c AND statistic) OR (area under the curve) OR auc OR calibration OR indices OR algorithm OR multivariable)) AND (atrial fibrillation)

Search string Embase:

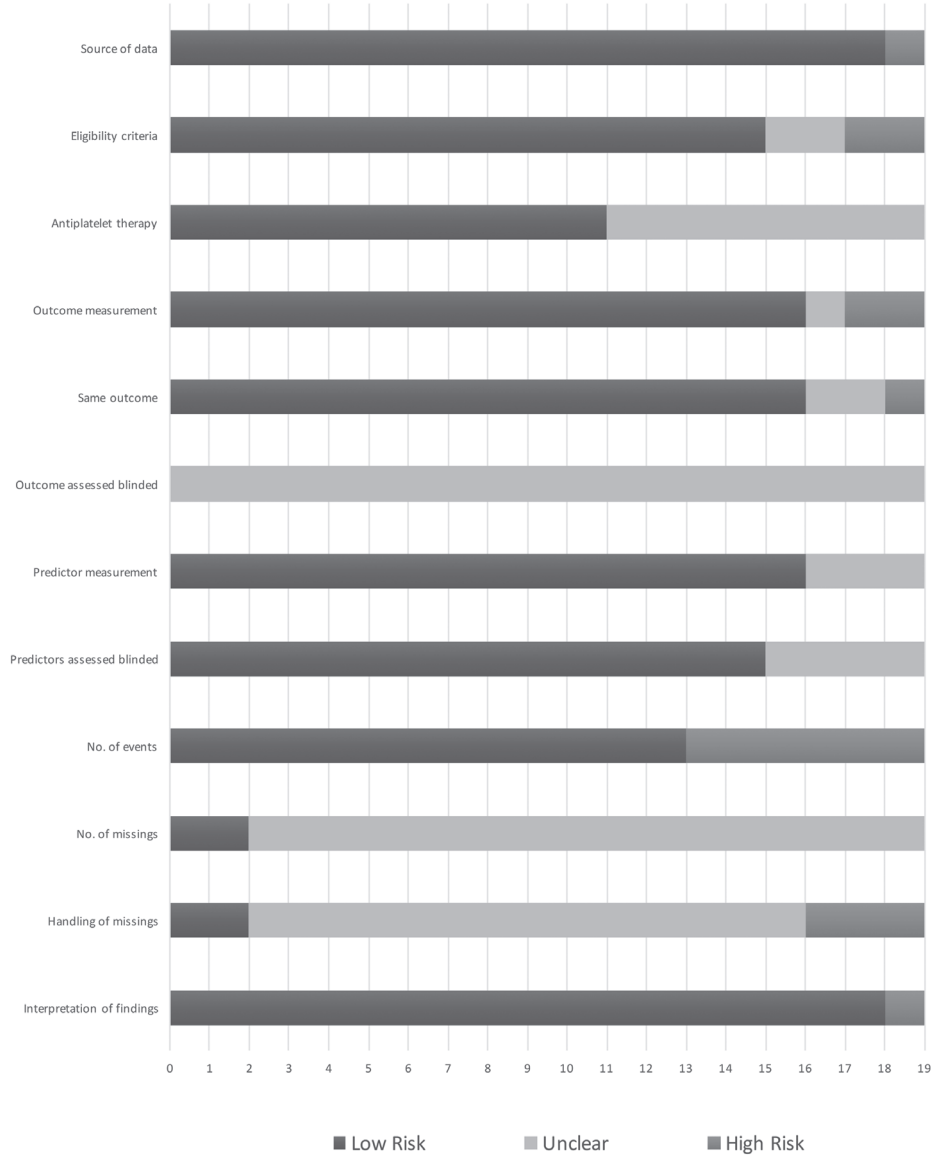
(((validat* OR predict*.ti. OR rule*) OR (predict* AND (outcome* OR risk* OR model*)) OR ((history/exp OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor*) AND (predict* OR model* OR decision* OR identif* OR prognos*)) OR (decision* AND (model* OR clinical* OR (logistic AND models))) OR (prognostic AND ('history'/exp OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor* OR model*))) OR ('stratification'/exp OR (roc

AND curve) OR discrimination OR discriminate OR 'c statistic' OR (c AND statistic) OR (area AND under AND the AND curve) OR 'auc'/exp OR 'calibration'/exp OR indices OR 'algorithm'/exp OR multivariable)) AND ('atrial fibrillation'/exp OR 'atrial fibrillation') AND 'article'/it AND (2001:py OR 2002:py OR 2003:py OR 2004:py OR 2005:py OR 2006:py OR 2007:py OR 2008:py OR 2009:py OR 2010:py OR 2011:py OR 2012:py) AND [embase]/lim

SUPPLEMENTAL MATERIAL 3. RISK OF BIAS ITEMS

Item
Source of data
Eligibility criteria
Details of antiplatelet therapy received
Definition and method for measurement of outcome
Same outcome definition in all patients
Outcome measurement blinded for predictors
Definition and method for measurement of predictors
Predictor measurement blinded for outcome
Number of events
Amount of missing data
Handling of missing data
Interpretation

SUPPLEMENTAL MATERIAL 4. OVERVIEW RISK OF BIAS



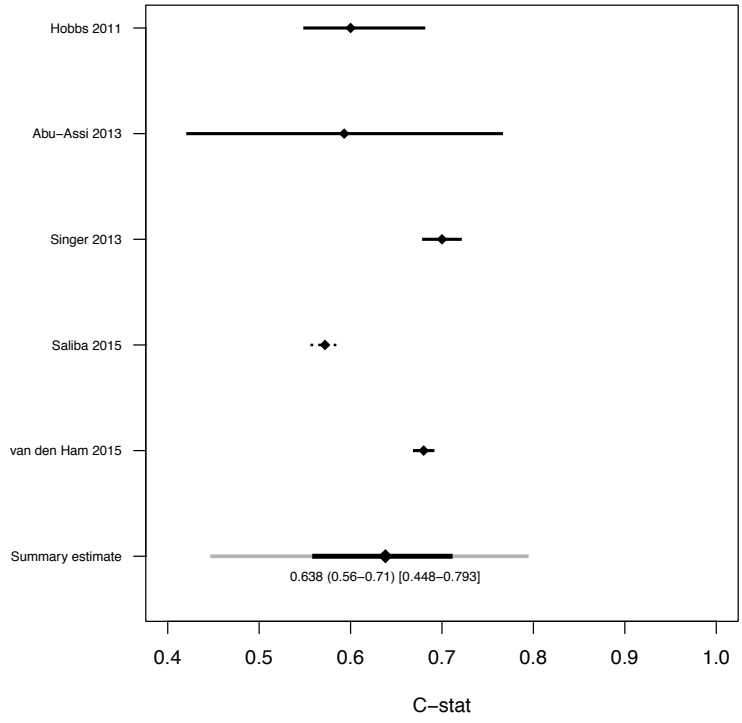
SUPPLEMENTAL MATERIAL 5. RISK OF BIAS OF INDIVIDUAL STUDIES

		Source of data	Eligibility criteria	Antiplatelet therapy	Outcome meas.	Same outcome	Outcome ass. blinded	Predictor meas.	Predictors ass. blinded	No. of events	No. of missings	Handling missings	Interpretation of findings
Hobbs [2]	2011	-	+	+	+	+	±	+	+	-	±	±	+
Olesen [3]	2011	+	+	+	+	+	±	+	+	+	±	+	+
Sandhu [4]	2011	+	+	±	-	+	±	+	+	+	±	±	+
Abu-Assi [5]	2013	+	+	+	+	+	±	±	+	-	±	±	+
Guo [6]	2013	+	±	+	+	+	±	+	+	-	+	+	+
Singer [7]	2013	+	+	±	+	+	±	+	+	+	±	±	+
Forslund [8]	2014	+	+	+	+	+	±	+	+	+	±	±	+
Komatsu [9]	2014	+	-	+	+	+	±	+	±	+	±	±	+
Siu [10]	2014	+	±	+	-	±	±	±	+	+	±	-	+
Abumuaileq [11]	2015	+	+	+	+	+	±	+	±	-	+	-	+
Saliba [12]	2015	+	+	±	±	±	±	+	+	+	±	±	+
Suzuki [13]	2015	+	+	+	+	-	±	+	±	-	±	±	+
vd Ham [14]	2015	+	+	±	+	+	±	+	+	+	±	±	+
Aspberg [15]	2016	+	+	±	+	+	±	+	+	+	±	±	+
Chao [16]	2016	+	+	+	+	+	±	+	+	+	±	±	+
Nielsen [17]	2016	+	+	±	+	+	±	+	+	+	±	±	+
Xing [18]	2016	+	-	+	+	+	±	±	±	-	±	-	-
Allan [19]	2017	+	+	±	+	+	±	+	+	+	±	±	+
McAlister [20]	2017	+	+	±	+	+	±	+	+	+	±	±	+

+ low risk of bias; ± unclear risk of bias; - high risk of bias

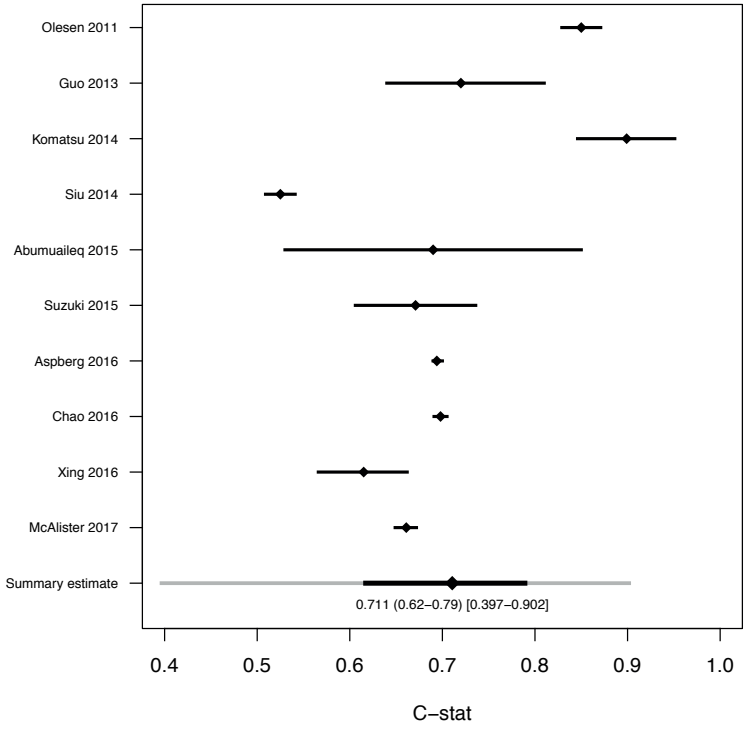
SUPPLEMENTAL MATERIAL 6.

Discriminative ability of CHA2DS2-VASc in studies recruiting from the general population



Solid bars represent 95% confidence intervals
Dashed bars represent 95% confidence intervals, estimated
Grey bars represent 95% prediction intervals
Summary estimate is c-statistic (95% CI) [95% PI]

Discriminative ability of CHA2DS2-VASc in studies recruiting from hospitals



Solid bars represent 95% confidence intervals
Dashed bars represent 95% confidence intervals, estimated
Grey bars represent 95% prediction intervals
Summary estimate is c-statistic (95% CI) [95% PI]

SUPPLEMENTAL MATERIAL 7. META-REGRESSION AND BEST AVAILABLE EVIDENCE

Meta regression

To investigate potential sources of heterogeneity, we included study characteristics and summarised patient characteristics as covariates in the random effects model.

For categorical covariates, we studied the 95% prediction interval. For instance, in studies enrolling patients from hospitals a meta-regression model with the outcome under study as a categorical covariate, yielded approximate 95% prediction intervals for risk of ischaemic stroke for score zero, one and two of 0.0 to 3.6%, 0.1 to 8.8% and 0.0 to 11.1%, respectively. For the outcome thromboembolism, this was 0.0 to 2.3%, 0.7 to 7.2% and 0.1 to 10.1%, respectively.

See table Model 1.

Model 1 - the type of outcome added as a categorical covariate, studies in hospital care patients

	Ischaemic stroke		**TE**		**All strokes**	
		95% PI		95% PI		
C-statistic	0.65	0.36 – 0.86	0.81	0.56 – 0.94	NA	NA
Score 0	0.8	0.0 – 3.6	0.2	0.0 – 2.3	NA	NA
Score 1	1.7	0.1 – 8.8	0.9	0.7 – 7.2	1.75	0.35 – 10.46
Score 2	2.9	0.0 – 11.1	2.1	0.1 – 10.1	NA	NA
Score 3	3.9	0.0 –16.5	3.5	0.0 – 13.7	NA	NA

NA = not available; 95% PI = approximate 95% prediction interval

Model 2 - Risk of bias added as a categorical covariate, studies in hospital care patients

	Low risk		**High risk**	
		95% PI	c-stat	95% PI
C-statistic	0.76	0.33 – 0.95	0.71	0.30 – 0.94
Score 0	0.7	0.1 – 3.9	0,4	0.3 – 3.4
Score 1	1.4	0.2 – 7.6	1.4	0.2 – 7.5
Score 2	2.7	0.0 – 11.2	2.5	0.0 – 10.8
Score 3	4.0	0.1 – 14.0	3.5	0.00 – 13.2

95% PI = approximate 95% prediction interval

For continuous covariates, we created regression plots and visually assessed the 95% prediction interval. As an example, for any observed annual incidence of the main outcome, future validation studies enrolling AF patients in hospitals may find a c-statistic between approximately 0.4 and 0.9. Likewise, the 95% PI of a model containing the mean (or median) age of the study population remained wide for both discrimination and the stroke risk per score, see Figures. Adding

summarised patient characteristics as continuous covariates to the model did not cause for a meaningful reduction in the 95% prediction interval. The same pattern was observed by adding the following covariates, both for studies recruiting form hospitals and studies recruiting from the general population (data not shown):

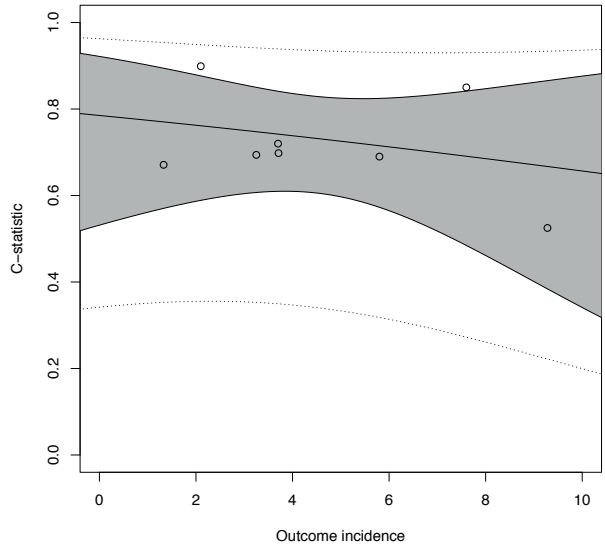
– The total number of person years of follow up in the study *;
– The (estimated) mean CHA2DS2-VASc score in the study population;
– Prevalence of heart failure in the study population;
– Prevalence of platelet inhibitor use in the study population;

* If not reported, the number of person years were calculated using the number of persons and the mean follow up time. If mean follow up was not reported, the median follow up time was used.

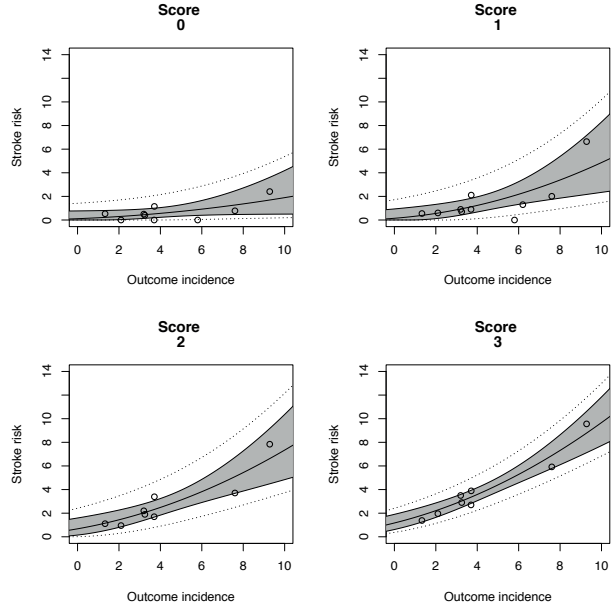
Best available evidence

We considered studies to provide best available evidence if they low/unclear risk of bias on less than five items, with clear reporting of the definition and the measurement of outcome and predictors, and with a sufficiently large number of outcomes. For studies enrolling non-selected patients, four studies[7,8,14,19] fulfilled these criteria. Among these, there was still substantial heterogeneity in the results. The reported stroke risks for CHA2DS2-VASc 0, 1 and 2 ranged from 0.04 to 0.4, from 0.2 to 0.8, and 0.8 to 1.9, respectively. For five studies enrolling patients in hospital care that were considered to provide the best available evidence,[3,15,16,20,21] corresponding stroke rates ranged from 0.4 to 1.2, 0.7 to 2.1 and 1.9 to 3.7 for CHA2DS2-VASc 0, 1 and 2, respectively.

Meta-regression with outcome incidence as a continuous covariate – discriminative ability

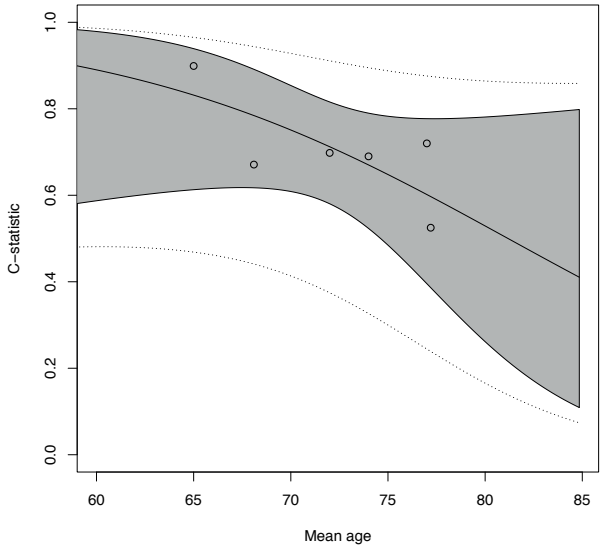


Meta-regression with outcome incidence as a continuous covariate – stroke risk per score

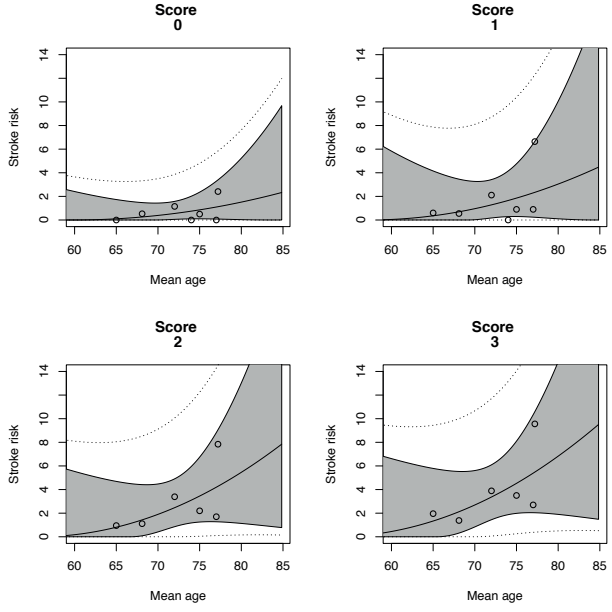


Dots represent the observed c-statistic and stroke risk per score, respectively. Shaded area represents 95% confidence interval. Dashed lines represent approximate 95% prediction interval.

Meta-regression with mean age of the study population as a continuous covariate – discriminative ability



Meta-regression with mean age of the study population as a continuous covariate – stroke risk per score



Dots represent the observed c-statistic and stroke risk per score, respectively. Shaded area represents 95% confidence interval. Dashed lines represent approximate 95% prediction interval.

SUPPLEMENTAL MATERIAL REFERENCES

1 Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. 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.

2 Hobbs FD, Roalfe AK, Lip GYH, Fletcher K, Fitzmaurice DA, Mant J. Performance of stroke risk scores in older people with atrial fibrillation not taking warfarin: comparative cohort study from BAFTA trial. *BMJ* 2011; **342**: d3653.

3 Olesen JB, Lip GYH, Hansen ML, Hansen PR, Tolstrup JS, Lindhardsen J, Selmer C, Ahlehoff O, Olsen A-MS, Gislason GH, Torp-Pedersen C. Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study. *BMJ* 2011; **342**: d124.

4 Sandhu RK, Bakal JA, Ezekowitz JA, McAlister FA. Risk stratification schemes, anticoagulation use and outcomes: the risk—treatment paradox in patients with newly diagnosed non-valvular atrial fibrillation. *Heart* 2011; **97**: 2046–50.

5 Abu-Assi E, Otero-Ravina F, Allut Vidal G, Coutado Mendez A, Vaamonde Mosquera L, Sanchez Loureiro M, Caneda Villar MC, Fernandez Villaverde JM, Maestro Saavedra FJ, Gonzalez-Juanatey JR, Grupo Barbanza researchers. Comparison of the reliability and validity of four contemporary risk stratification schemes to predict thromboembolism in non-anticoagulated patients with atrial fibrillation. *Int J Cardiol* 2013; **166**: 205–9.

6 Guo Y, Apostolakis S, Blann AD, Wang H, Zhao X, Zhang Y, Zhang D, Ma J, Wang Y, Lip GYH. Validation of contemporary stroke and bleeding risk stratification scores in non-anticoagulated Chinese patients with atrial fibrillation. *Int J Cardiol* 2013; **168**: 904–9.

7 Singer DE, Chang Y, Borowsky LH, Fang MC, Pomernacki NK, Udaltsova N, Reynolds K, Go AS. A new risk scheme to predict ischemic stroke and other thromboembolism in atrial fibrillation: the ATRIA study stroke risk score. *J Am Heart Assoc* 2013; **2**: e000250–0.

8 Forslund T, Wettermark B, Wandell P, Euler von M, Hasselstrom J, Hjemdahl P. Risks for stroke and bleeding with warfarin or aspirin treatment in patients with atrial fibrillation at different CHA2DS2VASc scores: experience from the Stockholm region. *Eur J Clin Pharmacol* 2014; **70**: 1–9.

9 Komatsu T, Sato Y, Ozawa M, Kunugita F, Yoshizawa R, Morino Y, Nakamura M. Comparison between CHADS2 and CHA2DS2-VASc score for risk stratification of ischemic stroke in Japanese patients with non-valvular paroxysmal atrial fibrillation not receiving anticoagulant therapy. *Int Heart J* 2014; **55**: 119–25.

10 Siu C-W, Lip GYH, Lam K-F, Tse H-F. Risk of stroke and intracranial hemorrhage in 9727 Chinese with atrial fibrillation in Hong Kong. *Heart Rhythm* 2014; **11**: 1401–8.

11 Abumuaileq RRY, Abu-Assi E, López-López A, Raposeiras-Roubin S, Rodríguez-Mañero M, Martínez-Sande L, García-Seara J, Fernandez-López XA, Peña-Gil C, González-Juanatey JR. Comparison between CHA2DS2-VASc and the new R2CHADS2 and ATRIA scores at predicting thromboembolic event in non-anticoagulated and anticoagulated patients with non-valvular atrial fibrillation. *BMC Cardiovasc Disord* 2015; **15**: 156.

12 Saliba W, Barnett-Griness O, Elias M, Rennert G. The Association Between Red Cell Distribution Width and Stroke in Patients with Atrial Fibrillation. *Am J Med* 2015; **128**: 192.e11–8.

13 Suzuki S, Yamashita T, Okumura K, Atarashi H, Akao M, Ogawa H, Inoue H. Incidence of Ischemic Stroke in Japanese Patients With Atrial Fibrillation Not Receiving Anticoagulation Therapy. *Circ J* 2015; **79**: 432–8.

14 van den Ham HA, Klungel OH, Singer DE, Leufkens HGM, van Staa TP. Comparative Performance of ATRIA, CHADS2, and CHA2DS2-VASc Risk Scores Predicting Stroke in Patients With Atrial Fibrillation: Results From a National Primary Care Database. *J Am Coll Cardiol* 2015; **66**: 1851–9.

15 Aspberg S, Chang Y, Atterman A, Bottai M, Go AS, Singer DE. Comparison of the ATRIA, CHADS2, and CHA2DS2-VASc stroke risk scores in predicting ischaemic stroke in a large Swedish cohort of patients with atrial fibrillation. *Eur Heart J* 2016; **37**: 3203–10.

16 Chao T-F, Liu C-J, Tuan T-C, Chen S-J, Wang K-L, Lin Y-J, Chang S-L, Lo L-W, Hu Y-F, Chen T-J, Chiang C-E, Chen S-A. Comparisons of CHADS2 and CHA2DS2-VASc scores for stroke risk stratification in atrial fibrillation: Which scoring system should be used for Asians? *Heart Rhythm* 2016; **13**: 46–53.

17 Nielsen PB, Larsen TB, Skjøth F, Overvad TF, Lip GYH. Stroke and thromboembolic event rates in atrial fibrillation according to different guideline treatment thresholds: A nationwide cohort study. *Sci Rep* 2016; **6**: 27410.

18 Xing Y, Ma Q, Ma X, Wang C, Zhang D, Sun Y. CHADS2 score has a better predictive value than CHA2DS2-VASc score in elderly patients with atrial fibrillation. *Clin Interv Aging* 2016; **11**: 941–6.

19 Allan V, Banerjee A, Shah AD, Patel R, Denaxas S, Casas J-P, Hemingway H. Net clinical benefit of warfarin in individuals with atrial fibrillation across stroke risk and across primary and secondary care. *Heart* 2017; **103**: 210–8.

20 McAlister FA, Wiebe N, Jun M, Sandhu R, James MT, McMurtry MS, Hemmelgarn BR, Tonelli M. 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**: 243–52.

21 Nielsen PB, Larsen TB, Skjøth F, Overvad TF, Lip GYH. Stroke and thromboembolic event rates in atrial fibrillation according to different guideline treatment thresholds: A nationwide cohort study. Sci Rep 2016; 6: 27410.

2

Reasons for non-adherence to stroke prevention guidelines in patients with atrial fibrillation

Sander van Doorn, Floor Hartman-Weide, Geert-Jan Geersing, Ruud Oudega, Arno W. Hoes, Frans H. Rutten

Based on:
van Doorn S, Hartman-Weide F, Geersing G-J, Oudega R, Hoes AW, Rutten FH.Reasons for non-adherence to practice guidelines on stroke prevention in patients with atrial fibrillation: A cross-sectional study in primary care. Int J Cardiol. 2015;187:525–526.

ABSTRACT

BACKGROUND: Non-adherence to practice guidelines on stroke prevention in atrial fibrillation (AF) is common, yet the reasons why are unclear.

AIM: To assess current non-adherence and underlying reasons of general practitioners (GPs) to practice guidelines on stroke prevention in AF.

DESIGN and settings: An observational cross-sectional study in Dutch general practice.

METHODS: the management of AF patients from 19 practices was analysed, and reasons for non-adherence per individual patient assessed by asking the GP.

RESULTS: The median age of 440 included patients was 76.0 (IQR 67.0 – 83.0) years, 55% were male, and 61.6% received cooperative care from the cardiologist. Undertreatment according to the CHADS2 and CHA2DS2-VASc decision rules occurred in 93 (21.1%) and 104 (23.6%) patients, respectively. Overtreatment occurred in 84 (19.1%) and 29 (3.4%) patients, respectively. The main reasons mentioned by GPs for non-adherence per individual undertreated case was i) sustained sinus rhythm after an episode of AF, and ii) the cardiologist was considered responsible for the anticoagulation. Adverse effects or contra-indications to drugs, or patient's preferences were seldomly mentioned as reasons for undertreatment.

CONCLUSIONS: Uncertainty about how to manage patients who have sinus rhythm after an episode of AF, and whom is responsible for anticoagulation seem to be more important reasons for GPs to non-adhere to guidelines on stroke prevention than fear of bleedings. These barriers may be overcome by clearer evidence-based recommendations on the management of patients with sinus rhythm after AF, and better agreement on responsibility of initiation and monitoring of anticoagulation therapy between GPs and cardiologists.

INTRODUCTION

Atrial fibrillation (AF) is associated with an increased risk of stroke and mortality if left untreated.[1] Anticoagulants – such as vitamin K antagonists (VKA) and non-VKA oral anticoagulants (NOACs) – are highly effective in preventing stroke [2–5] and superior to aspirin.[6] Importantly, anticoagulants inherently carry the risk of bleeding complications. Thus, prescription of anticoagulants should be based on the absolute risk of stroke and thereby restricted to those in whom the absolute stroke risk outweighs the risk of bleeding complications. Hereto, clinical decision rules have been developed, validated and subsequently recommended in current practice guidelines for daily use in clinical practice.

Previous studies showed that such decision rule-guided prescription of anticoagulants in everyday clinical practice was modest to low.[7–10] Up to 40% of the patients with AF in daily practice did not to receive adequate stroke prevention according to the often recommended CHADS2 rule.[11,12]

Knowledge of the exact reasons for non-adherence of physicians to available guidelines on AF is scarce and mainly based on vignette studies or qualitative studies in which physicians were asked to mention in general the reasons for non-adherence.[13–17] Although informative, such studies may be prone to bias by including highly selected and small numbers of patients. A quantitative study based on per individual assessment of the reasons for non-adherence in a large enough unselected patient population is lacking.

In the current study we included an unselected large group of primary care patients with AF, and we i) assessed the adherence to prevailing AF guidelines on stroke prevention and ii) identified reasons for non-adherence as mentioned by the general practitioners (GPs) per individual deviant managed case.

METHODS

Study population

Four large group practices participated with in total nineteen general practitioners and 26,400 enlisted persons. Practices covered urban, sub-urban, and rural areas in the Netherlands. All citizens in the Netherlands are registered with a GP, irrespective of cooperative care from a medical specialist, including patients living in a home for the elderly, but with the exception of those living in a nursing home or hospice. Our study population is therefore a representative sample of community-dwelling persons with AF.

Data collection

Between July 2012 and November 2012, patients with a diagnosis of atrial fibrillation (International Coding for Primary Care (ICPC) code K78) were identified from the electronic medical files of the participating general practices. Only electrocardiographically confirmed cases were included in the study. Patient characteristics, comorbidities, visits to a cardiologist in the last two years, and cardiovascular drug use were manually extracted from the files of the general practitioners. Next, for each patient with AF, anticoagulation status was recorded.

The medical ethical committee of the University Medical Centre Utrecht, the Netherlands approved the study protocol. Anonymised data was used for analysis.

The stroke prevention recommendations on AF available in the Netherlands when the study was executed

At the time of the study, 2012, two different guidelines on AF were available with different recommendations for stroke prevention, with GPs following the 2010 Dutch General Practice guidelines and cardiologists the 2010 ESC guidelines on AF.

The Dutch General Practice Guidelines on AF recommended the CHADS2 clinical decision rule for risk stratification of patients with AF for anticoagulation.[18] This rule assigns 1 point for Congestive heart failure, Hypertension, Age 75 years or older and Diabetes, and 2 points for prior Stroke/TIA.[19] A platelet inhibitor was recommended for those with a score 0 (low risk of stroke) and 1 (intermediate risk), and oral anticoagulation for scores ≥ 2 (high risk). In our study, patients with a CHADS2 score of 0 or 1 who did not receive a platelet inhibitor, or patients with a score ≥ 2 without anticoagulation were considered as undertreated. Those with a CHADS2 score of 0 or 1 who received anticoagulation were considered as overtreated. Patients receiving a combination of an anticoagulant and a platelet inhibitor were analysed as receiving anticoagulation.

Non-adherence to the stroke recommendations by the general practitioner

The general practitioner was asked to fill out a questionnaire on reasons for non-adherence per individual case not managed according to the CHADS2-based recommendations. We classified afterwards their arguments into categories; related to overtreatment or undertreatment.

In additional analyses, we also evaluated the adherence to the newer CHA2DS2-VASc, the decision rule that cardiologists used in the Netherlands during the study period.[20,21] Compared to the CHADS2, this score assigns one additional point for vascular disease and female sex, and one or two points for age 65 to 75, and 75 years and older, respectively. According to CHA2DS2-VASc, we considered

patients undertreated if no stroke prevention (either aspirin or VKA/NOAC) was prescribed for score 1, or no anticoagulation (VKA/NOAC) for a score ≥ 2. Patients on anticoagulation with a CHA2DS2-VASc score 0 were considered overtreated.

Data analysis

First, we calculated the proportion of patients not managed according to the 2010 Dutch General Practice guidelines on AF recommending the CHADS2 decision rule, and divided them in overtreated and undertreated. Next, we calculated the proportion of patients not managed according to the 2010 ESC guidelines on AF recommending the CHA2DS2-VASc rule and divided them in undertreated and overtreated. Reasons for non-adherence as mentioned by the GPs were assessed per individual case non-adherent to CHADS2. Data was analysed with SPSS version 20.0 for Mac OSX (SPSS inc.).

RESULTS

Patient characteristics

In total, 440 community-dwelling persons had electrocardiographically-confirmed AF, which corresponds to a prevalence of 1.7%. The median age was 76 (IQR 67.0 – 83.0) years, and 55% were male. Hypertension (58%), heart failure (29%), and vascular diseases (21%) were the most common co-morbidities. See Table 1.

TABLE 1. Patient characteristics of 440 patients with atrial fibrillation

Patient characteristics	Patients n = 440
Median age (IQR)	76.0 (67.0 – 83.0)
Mean CHADS2 score (SD)	2.0 (1.43)
CHADS2 score 0	16.8
CHADS2 score 1	23.2
CHADS2 score ≥ 2	60.0
Female sex	45.5
History of heart failure	29.3
Hypertension	57.7
Age ≥ 75 years	55.9
Diabetes mellitus	19.1
Prior stroke/TIA	18.4
Vascular disease*	21.4
Age between 65 and 75 years	23.2

Numbers are percentages unless otherwise specified.
IQR = inter quartile range
* Vascular disease constitutes of coronary artery disease and peripheral vascular disease.

In total, 177 (40.2%) patients were not managed according to CHADS2; 93 (21.1%) were undertreated, and 84 (19.1%) overtreated. 119 AF patients were not managed according to CHA2DS2-VASc: 104 (23.6%) were undertreated, and 15 (3.4%) overtreated.

The majority of patients (61.6%) received cooperative care from the cardiologist, and non-adherence to CHA2DS2-VASc in this specific subgroup was lower; 16.6% undertreated, and 3.7% overtreated, respectively.

Questions posed to general practitioners on AF guidelines adherence

Of the 19 GPs, 17 (89.5%) completed the questions on guideline adherence. The majority of them (69%) considered stroke prevention to be the primary responsibility of the cardiologist. Four GPs admitted not to be familiar with the guideline recommendations and the CHADS2 decision rule, and in total eight GPs did not know the CHA2DS2-VASc rule at the time of the study. The remaining 11 GPs familiar with either rule filled out that they 'always' (53%), 'sometimes' (24%), and 'never' (33%) used a rule for deciding on stroke prevention in AF.

Reasons mentioned by GPs for undertreatment according to CHADS2

In 37 patients (40% of the reasons for undertreatment) the GP mentioned suspected sinus rhythm for more than one year after a period of AF. In 25 cases (27%) the GP could not reproduce the reason, and he/she thought that the cardiologist monitored the stroke prevention management. Of these 25 patients, 17 (68.0%) indeed consulted a cardiologist in the two years before the study, but the remaining 8 (32.0%) did not. In 9 cases a contra-indication for anticoagulation, and in 4 cases patient's preferences were mentioned as the reasons for undertreatment according to CHADS2. In none of the cases the GP mentioned adverse effects of anticoagulation, see also Table 2.

Reasons mentioned by GPs for overtreatment according to CHADS2

In 62 patients (74%) the reasons for overtreatment as mentioned by the GPs was considering the cardiologist responsible for the stroke prevention management. Of these 62 patients, 48 (77.4%) indeed consulted a cardiologist in the two years before the study, but 14 (22.6%) had not. In 10 patients, the reason was preferring the CHA2DS2-VASc score over the CHADS2. Indeed, all 10 overtreated according to CHADS2 were adequately managed according to CHA2DS2-VASc. In 10 patients the GP could not provide a reason for the overtreatment according to CHADS2.

TABLE 2. Reasons mentioned by GPs for inadequate stroke prevention according to the CHADS2 decision rule score per individual non-adherent case with atrial fibrillation (N = 177)

Reasons	93 undertreated cases n (%) *	84 overtreated cases n (%) *
Cardiologist considered to be responsible for anticoagulation	25 (26.9)	62 (73.8)
A comorbid condition meriting another choice	3 (3.2)	15 (17.9)
Considered contra-indicated	9 (9.7)	0 (0.0)
Prior adverse effects on OACs	0 (0.0)	0 (0.0)
Did not monitor anticoagulation for more than a year	8 (8.6)	4 (4.8)
Patient's preference	4 (4.3)	0 (0.0)
Did not know I should apply the CHADS2 rule	2 (2.2)	0 (0.0)
Suspected sinus rhythm after an episode of AF	37 (39.8)	0 (0.0)
CHA2DS2-VASc applied instead of CHADS2	1 (1.1)	10 (11.9)
Other reasons or 'unknown'	7 (7.5)	13 (15.5)

* Cells do not add up to hundred percent because more than one answer was allowed.

DISCUSSION

In our cross-sectional practice study among AF patients in the primary care setting, stroke prevention was not according to the CHADS2 decision rule in 40.2% of the patients, and in 27.0% not according to CHA2DS2-VASc. Undertreatment was more common than overtreatment according to CHADS2 (21.1% vs. 19.1%), and CHA2DS2-VASc (23.6% vs. 3.4%). GPs mentioned suspected sinus rhythm for more than a year after a period of AF as the main reason for withholding adequate stroke prevention. The mean reason for overtreatment was that GPs considered the cardiologist responsible for the anticoagulation. Contra-indications for anticoagulation, interactions with other drugs, or patient preferences were seldomly mentioned as reasons for non-adherence.

Strengths and limitations

We could include a large representative sample of community-dwelling patients with AF from primary care in our study. The distribution over the CHADS2 score of our population was comparable to that of other population-based studies. [19,22–25] We included electrocardiographically-confirmed AF to prevent including misclassified cases. Our study design allowed us to receive case-specific barriers to guideline adherence.

That we did not evaluate the cardiologist's opinion could be considered as a limitation.

Comparison with existing literature

A previously published meta-analysis on underuse of anticoagulation in AF patients found a pooled prevalence of undertreatment of 40% in patients identified as being at high risk of stroke.[11] In our current study, we found a substantially lower prevalence of undertreatment of around 20% with the CHADS2 or CHA2DS2-VASc decision rule as the reference.

Previous survey-based studies exploring reasons for non-adherence to stroke prevention in atrial fibrillation found that older age, high BMI, and male gender were positively related to anticoagulation use, while previous stroke, dementia, and low stroke risk were negatively related.[26–28] Vignette studies exploring barriers to prescribing anticoagulants showed that physicians often mention bleeding risk as the main barrier.[16,17] A systematic review summarised increasing age, comorbidities, and inability to comply with treatment as important barriers.[15] In interviews, clinicians also mentioned patients' social situation and past medication-taking behaviour as challenges to anticoagulation therapy.[13] In a study published in 2004, GPs and cardiologists were asked for reasons 227 AF patients admitted for a stroke were not on anticoagulation on admission. 'Contra-indicated' and 'no indication' were mentioned as the most important reasons.[29] In a recent primary care study, GPs mentioned 'not indicated', 'contra-indicated', and 'non-compliance' as main reasons for not prescribing anticoagulants in a sample of 15 patients with AF while recommended in the prevailing guideline. This was, however, only in a small selection of patients.[30]

The aforementioned studies included small numbers of often selected patients, without considering reasons for non-adherent per individual case. This may have created bias, including 'preferred answering' and recall bias of physicians.

Our study clearly shows that co-morbidities, non-compliance, adverse effects and contra-indications do not play an essential role according to the GPs when considering each individual non-adherent to stroke prevention according to CHADS2. 'Suspected sinus rhythm' was an important reason for undertreatment in our study. Multiple Holter studies make this to a dubious argument as nearly all patients studied because of suspicion of sinus rhythm after an episode of AF showed to have paroxysms of AF, including asymptomatic episodes.[31,32] On the other hand, a recent large study suggested a lower stroke and mortality risk for those with paroxysmal AF than patients in permanent AF.[33] In a recently published small proof-of-principle study, it was shown that patients who underwent closing of the left atrial appendix had less ischaemic stroke and less bleeds during follow up while not receiving anticoagulation than those randomised to care as usual with continuation of anticoagulation.[34]

Guidelines on AF recommend the same stroke prevention for those with paroxysmal and permanent/persistent AF. They also recommend to continue anticoagulation after cardioversion or pulmonary vein isolation. They do, however, not specifically address what to do in cases with suspected sinus rhythm after an episode of AF.[35–37] Thus, there is a need for clear guidance how to manage patients with (long periods of) sinus rhythm after an episode of AF, because these patients may – possibly unjustified – not receive anticoagulation.

An easy to tackle barrier to adherence was uncertainty about who was responsible for anticoagulation; the cardiologist or the GP. Regional agreement about whom is responsible for initiation, and monitoring of the anticoagulation would improve adherence to stroke prevention in AF.

Implications for practice

Evidence-based recommendations are needed on how to manage patients suspected of sinus rhythm after a period of AF, and GPs and cardiologists should work more closely together in patients with AF.

CONCLUSIONS

Non-adherence to stroke prevention guidelines is common in community-dwelling patients with AF. Suspected sinus rhythm and uncertainty about whether the GP or cardiologist is responsible are critical reasons for non-adherence.

ACKNOWLEDGEMENTS

We would like to thank the participating general practitioners PJE Bindels, GJG Geersing, MA de Meij, BJ Noort, A. Wewerinke (Buitenhof), M Bosman, BE van der Snoek, M. Kocks, MW van den Broek, BD Frijling, FH Rutten (de Grebbe), HAM Asbreuk, JG Blommestein, I Brouwer, AE Denteneer, NC de Hartog, YCJ van der Helm (Mondriaanlaan), LJ Meijer, MJM Valk (Amersfoort).

COMPETING INTEREST

GJG and FHR received an unrestricted institutional grand from Boehringer Ingelheim.

REFERENCES

1 Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991;22:983–8.

2 Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Ann Intern Med* 2007;146:857–67.

3 Connolly SJ, Ezekowitz MD, Yusuf S, *et al.* Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2009;361:1139–51. doi:10.1056/NEJMoa0905561

4 Patel MR, Mahaffey KW, Garg J, *et al.* Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med* 2011;365:883–91. doi:10.1056/NEJMoa1009638

5 Granger CB, Alexander JH, McMurray JJ, *et al.* Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011;365:981–92. doi:10.1056/NEJMoa1107039

6 van Walraven C, Hart RG, Connolly S, *et al.* Effect of age on stroke prevention therapy in patients with atrial fibrillation: the atrial fibrillation investigators. *Stroke* 2009;40:1410–6. doi:10.1161/STROKEAHA.108.526988

7 Kalra L, Yu G, Perez I, *et al.* Prospective cohort study to determine if trial efficacy of anticoagulation for stroke prevention in atrial fibrillation translates into clinical effectiveness. *BMJ* 2000;320:1236–9.

8 Rutten FH, Hak E, Stalman WA, *et al.* Is treatment of atrial fibrillation in primary care based on thromboembolic risk assessment? *Fam Pract* 2003;20:16–21.

9 Mazzaglia G, Filippi A, Alacqua M, *et al.* A national survey of the management of atrial fibrillation with antithrombotic drugs in Italian primary care. *Thromb Haemost* 2010;103:968–75. doi:10.1160/th09-08-0525

10 DeWilde S, Carey IM, Emmas C, *et al.* Trends in the prevalence of diagnosed atrial fibrillation, its treatment with anticoagulation and predictors of such treatment in UK primary care. *Heart* 2006;92:1064–70. doi:10.1136/hrt.2005.069492

11 Ogilvie IM, Newton N, Welner SA, *et al.* Underuse of oral anticoagulants in atrial fibrillation: a systematic review. *Am J Med* 2010;123:638–645e4. doi:10.1016/j.amjmed.2009.11.025

12 Scowcroft ACE, Cowie MR. Atrial fibrillation: improvement in identification and stroke preventive therapy - data from the UK Clinical Practice Research Datalink, 2000-2012. *Int J Cardiol* 2014;171:169–73. doi:10.1016/j.ijcard.2013.11.086

13 Decker C, Garavalia, Garavalia B, *et al.* Exploring barriers to optimal anticoagulation for atrial fibrillation: interviews with clinicians. *JMDH* 2012;:129. doi:10.2147/JMDH.S33045

14 Pugh D, Pugh J, Mead GE. Attitudes of physicians regarding anticoagulation for atrial fibrillation: a systematic review. *Age and Ageing* 2011;40:675–83. doi:10.1093/ageing/afr097

15 Gattellari M, Worthington J, Zwar N, *et al.* Barriers to the use of anticoagulation for nonvalvular atrial fibrillation: a representative survey of Australian family physicians. *Stroke* 2008;39:227–30. doi:10.1161/strokeaha.107.495036

16 Peterson GM, Boom K, Jackson SL, *et al.* Doctors' beliefs on the use of antithrombotic therapy in atrial fibrillation: identifying barriers to stroke prevention. *Intern Med J* 2002;32:15–23.

17 Gross CP, Vogel EW, Dhond AJ, *et al.* Factors influencing physicians' reported use of anticoagulation therapy in nonvalvular atrial fibrillation: a cross-sectional survey. *Clin Ther* 2003;25:1750–64. doi:10.1016/S0149-2918(03)80167-4

18 Opstelten W, Boode BS, Heeringa J, *et al.* [Summary of the practice guideline "Atrial fibrillation" (first revision) from the Dutch College of General Practitioners]. *Ned Tijdschr Geneeskd* 2010;154:A1570.

19 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.

20 European Heart Rhythm Association, European Association for Cardio-Thoracic Surgery, Camm AJ, *et al.* Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). *Eur Heart J.* 2010;31:2369–429. doi:10.1093/eurheartj/ehq278

21 Lip GYH, 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. doi:10.1378/chest.09-1584

22 Friberg L, Rosenqvist M, Lip GY. Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: the Swedish Atrial Fibrillation cohort study. *Eur Heart J* 2012;33:1500–10. doi:10.1093/eurheartj/ehr488

23 Olesen JB, Torp-Pedersen C, Hansen ML, *et al.* The value of the CHA2DS2-VASc score for refining stroke risk stratification in patients with atrial fibrillation with a CHADS2 score 0-1: a nationwide cohort study. *Thromb Haemost* 2012;107:1172–9. doi:10.1160/th12-03-0175

24 Rietbrock S, Heeley E, Plumb J, *et al.* Chronic atrial fibrillation: Incidence, prevalence, and prediction of stroke using the Congestive heart failure, Hypertension, Age > 75, Diabetes mellitus, and prior Stroke or transient ischemic attack (CHADS2) risk stratification scheme. *Am Heart J* 2008;156:57–64. doi:10.1016/j.ahj.2008.03.010

25 Singer DE, Chang Y, Fang MC, *et al.* The net clinical benefit of warfarin anticoagulation in atrial fibrillation. *Ann Intern Med* 2009;151:297–305.

26 Ewen E, Zhang Z, Simon TA, *et al.* Patterns of warfarin use and subsequent outcomes in atrial fibrillation in primary care practices. *Vasc Health Risk Manag* 2012;8:587–98.

27 Arts DL, Visscher S, Opstelten W, *et al.* Frequency and risk factors for under- and over-treatment in stroke prevention for patients with non-valvular atrial fibrillation in general practice. *PLoS One* 2013;8:e67806.

28 Mohammed MA, Marshall T, Nirantharakumar K, *et al.* Patterns of warfarin use in subgroups of patients with atrial fibrillation: a cross-sectional analysis of 430 general practices in the United Kingdom. *PLoS One* 2013;8:e61979. doi:10.1371/journal.pone.0061979

29 Deplanque D, Leys D, Parnetti L, *et al.* Stroke prevention and atrial fibrillation: reasons leading to an inappropriate management. Main results of the SAFE II study. *Br J Clin Pharmacol* 2004;57:798–806.

30 Furthauer J, Flamm M, Sonnichsen A. Patient and physician related factors of adherence to evidence based guidelines in diabetes mellitus type 2, cardiovascular disease and prevention: a cross sectional study. *BMC Fam Pract* 2013;14:47. doi:10.1186/1471-2296-14-47

31 Israel CW, Grönefeld G, Ehrlich JR, *et al.* Long-term risk of recurrent atrial fibrillation as documented by an implantable monitoring device: implications for optimal patient care. *J Am Coll Cardiol* 2004;43:47–52.

32 Disertori M, Lombardi F, Barlera S, *et al.* Clinical characteristics of patients with asymptomatic recurrences of atrial fibrillation in the Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico-Atrial Fibrillation (GISSI-AF) trial. *Am Heart J* 2011;162:382–9. doi:10.1016/j.ahj.2011.05.008

33 Steinberg BA, Hellkamp AS, Lokhnygina Y, *et al.* Higher risk of death and stroke in patients with persistent vs. paroxysmal atrial fibrillation: results from the ROCKET-AF Trial. *Eur Heart J* 2015;36:288–96. doi:10.1093/eurheartj/ehu359

34 Reddy VY, Sievert H, Halperin J, *et al.* Percutaneous Left Atrial Appendage Closure vs Warfarin for Atrial Fibrillation: A Randomized Clinical Trial. *JAMA* 2014;312:1988–98. doi:10.1001/jama.2014.15192

35 January CT, Wann LS, Alpert JS, *et al.* 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol* 2014;64:e1–76. doi:10.1016/j.jacc.2014.03.022

36 You JJ, Singer DE, Howard PA, *et al.* Antithrombotic Therapy for Atrial Fibrillation: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 2012;141:e531S–e575S. doi:10.1378/chest.11-2304

37 Camm AJ, Lip GYH, De Caterina R, *et al.* 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation. Developed with the special contribution of the European Heart Rhythm Association. *Eur Heart J* 2012;33:2719–47. doi:10.1093/eurheartj/ehs253

3

Effectiveness of CHA2DS2-VASc based decision support on stroke prevention in atrial fibrillation: A cluster randomised trial in general practice

Sander van Doorn, Frans H. Rutten, Caitriona M. O'Flynn, Ruud Oudega, Arno W. Hoes, Karel G.M. Moons, Geert-Jan Geersing

Submitted

ABSTRACT

BACKGROUND: Guidelines on atrial fibrillation (AF) recommend using the CHA2DS2-VASc clinical decision rule for estimating stroke risk and to subsequently guide decision-making on oral anticoagulant therapy. Nevertheless, underuse of anticoagulants exists, especially in primary care. Providing general practitioners (GPs) with an automated decision support based on CHA2DS2-VASc may improve anticoagulant use and eventually prevent more strokes.

AIM: to study the impact of an automated decision support on stroke prevention in patients with AF.

DESIGN: A cluster randomised trial in general practice.

METHODS: A total of 38 practices were randomised. In the index practices, GPs were provided with an automatically generated CHA2DS2-VASc based anticoagulant treatment recommendation. The GPs in the reference practices provided care as usual. The primary outcome was the incidence of ischaemic stroke, TIA and/or thromboembolism (TE). Secondary outcomes were bleeding and the proportion of patients on guideline recommended anticoagulant treatment.

RESULTS: In total, 1129 AF patients were included in the 19 index practices and 1226 AF patients in the 19 reference practices. The median age was 77 (interquartile range (IQR) 68 – 75) years, and 48% were male. The median CHA2DS2-VASc score was 3.0 (IQR 2.0 – 5.0). Underuse of anticoagulants as recommended for patients with CHA2DS2-VASc score ≥ 2 was 6.6%. After a median follow-up of 2.7 years (IQR 2.3 – 3.0), the incidence rate per 100 person-years of ischaemic stroke/TIA/TE was 1.96 in the index group and 1.42 in the reference group (hazard ratio (HR) 1.3, 95% C 0.8 – 2.1). No difference was observed in the rate of bleeding (0.79 versus 0.82), or in the underuse (7.2% versus 8.2%) or overuse (8.0% versus 7.9%) of anticoagulation.

CONCLUSION: In this study in patients with AF in general practice, underuse of anticoagulants was relatively low. Nevertheless, providing general practitioners with CHA2DS2-VASc based decision support did not result in a reduction in stroke incidence, nor did it affect bleeding risk or anticoagulant over- or underuse.

INTRODUCTION

Atrial fibrillation (AF) is the most common cardiac arrhythmia[1] and a major risk factor for ischaemic stroke. Left untreated, stroke risk may increase fivefold and one in every five strokes is believed to be a direct consequence of atrial fibrillation.[2] The effectiveness of oral anticoagulants – both vitamin K antagonists (VKA) and direct oral anticoagulants (DOAC) – in preventing stroke in patients with AF is well established.[3–7] Practice guidelines on AF[8–10] provide clear anticoagulant treatment recommendations based on stroke risk prediction using the CHA2DS2-VASc clinical decision rule.[11]

However, underuse of anticoagulants in patients with AF is common, notably in primary care.[12–14] A meta-analysis published in 2010[15] showed that underuse of anticoagulants often exceeded 30% in AF patients at high risk of stroke as based on risk factors or the CHADS2 decision rule.[16] Many studies aimed to improve stroke prevention, for instance by evaluating automated decision support.[17–20] However, most used surrogate primary endpoints such as the relative increase in anticoagulant use, and were not powered to demonstrate an effect on a more clinically relevant outcome such as ischaemic stroke incidence. Additionally, such decision support often activates alerts with treatment recommendations when a patient's record is activated, i.e. *only* at the time of consultation. In busy daily practice, such unrequested alerts are easily ignored, thereby reducing the potential effect of such automated interventions. A strategy to improve anticoagulant treatment in *all* patients with AF (i.e. not only in those attending the GP) with ischaemic stroke and/or thromboembolism (TE) as the primary outcome has not yet been evaluated. In a cluster randomised trial, we assessed the effectiveness of an automated CHA2DS2-VASc decision support providing GPs with a treatment recommendation for all community-dwelling AF patients on the occurrence of the composite endpoint ischaemic stroke, TIA and/or thromboembolism; and on bleeding and the proportion of patients on guideline recommended anticoagulant treatment.

METHODS

Design

This study entitled CAFe *(Cost-effectiveness of balancing stroke and bleeding risk using CHA2DS2-VASc in primary care patients with Atrial Fibrillation)* is a pragmatic, parallel cluster randomised trial, trial registration number NTR3741 (www.trialregister.nl). Clusters are general care practices of varying size, and include both solo and

group practices. General practices in the region of Utrecht, the Netherlands were invited to participate. The initial invitation letter outlined an evaluation of the treatment of atrial fibrillation as the general purpose of the study but did not mention the specific aim of improving anticoagulant treatment in atrial fibrillation.

From February 2013 until September 2014 a planned total of 38 practices were enrolled in the study, covering urban, suburban and rural areas. After inclusion, practices were allocated by computer-generated randomisation to one of two groups. The clusters were stratified according to the number of electronic files screened for AF diagnosis (< 110, 110–180 and > 180 records).

After randomisation, only general practitioners in de index cluster received information on the specific study aim.

Patients

At baseline, an automated search in the 38 practices was used to identify all patients aged 18 years and older, diagnosed with AF as based on the International Classification of Primary Care (ICPC)[21] code K78 'atrial fibrillation/flutter'. In addition, files of patients with other ICPC codes for cardiac arrhythmias (K97 'Paroxysmal tachycardia' and K80 'Ectopic beats/Extra-systoles') as well as patients with a prescription of antiarrhythmic drugs (amiodarone, sotalol, digoxine and flecainide) or oral anticoagulants (acenocoumarol, phenprocoumon) were selected. Subsequently, all patient files thus identified were manually scrutinized by a clinical researcher (SD) for an electrocardiographically-confirmed AF diagnosis, following recommendations from the European Society of Cardiology.[22]

The CHA2DS2-VASc clinical decision rule and anticoagulant use

The CHA2DS2-VASc clinical decision rule was developed in 2010 as an update to the original CHADS2 rule[16] by including additional predictors for stroke. Its aim was to help detect patients with AF at low risk of stroke, TIA and/or TE. Patients are assigned points for presence of congestive heart failure (1 point), hypertension (1 point), age above 75 years (2 points), diabetes (1 point) and prior stroke (2 points), age above 65 (1 point), vascular disease (1 point) and female sex (1 point). For the exact definition of the predictors in CHA2DS2-VASc, see Supplemental Material 1.

For every patient included in the study, the correctness of each CHA2DS2-VASc variable was manually verified and where necessary corrected by the researcher (SD) using all available information in the electronic patient file including diagnostic test results, out-of-hours office reports and specialists' letters.

These verified CHA2DS-VASc variables were automatically extracted from the electronic patient file and for every patient the CHA2DS2-VASc score was calculated. The ATC codes of anticoagulant and antiplatelet drug prescriptions in the

year previous to study inclusion were automatically extracted after verification, and where necessary correction, by manually studying the electronic patient files.

Intervention and control group

Index group – CHA2DS2-VASc based strategy

Based on the 2013 Guideline Atrial Fibrillation (2nd revision) of the Dutch College of General Practitioners,[23] patients with a CHA2DS2-VASc score of either 0 or 1 were considered as best treated without anticoagulant treatment, whereas those with a score ≥ 2 should receive anticoagulation. This differs from the prevailing guideline by the European Society of Cardiology at the time of our study, mentioning that anticoagulation should be considered for patients with CHA2DS2-VASc score of 1.[22] Platelet inhibitors are not recommended for the prevention of stroke in either guideline.

An off-site computer calculated the CHA2DS2-VASc score and the recommended treatment as defined above for each patient in the index group practices. Based on this calculation, general practitioners in the index practices received a list of all identified AF patients in their practice whose treatment was *not* according to these guideline recommendations. This list was personally handed to practitioners by the researcher (SD). It contained the patients' CHA2DS2-VASc score, their current treatment and the recommended treatment. Practitioners in the intervention practices were advised to follow the recommendation to optimise anticoagulant treatment in shared decision with the patient. They were allowed to deviate from the recommended treatment, without stating the reasons why. One month later, practitioners received an e-mail reminding of the list with treatment recommendations, along with additional general information on atrial fibrillation, stroke risk prediction and stroke prevention, all based upon the Dutch College of General Practitioners' guidelines.

Control group – care as usual

Practitioners in the reference group neither received information on the true aim of the study, nor any list of AF patients in their practice, nor any treatment recommendations, nor extra information on atrial fibrillation and stroke risk. They were asked to provide care as usual.

Outcomes

We evaluated the effect of the automated decision support, based on the CHA2DS2-VASc score and its associated anticoagulant treatment recommendations, at individual patient level. The primary outcome was a composite of stroke,

TIA and/or thromboembolism. Stroke was defined as a focal neurological deficit of sudden onset lasting > 24 hours not attributable to other identifiably causes. TIA was defined as a focal neurological deficit of sudden onset lasting < 24 hours. Thromboembolism was defined as peripheral embolism or pulmonary embolism. Peripheral embolism was defined as a sudden occlusion of an artery to an extremity or a visceral organ outside the brain, heart, eyes, and lungs, not attributable to concomitant atherosclerosis or other aetiology. Pulmonary embolism was defined as radiographic confirmation of a pulmonary arterial occlusion.

The secondary endpoints were i) bleeding and ii) the proportion of patients on guideline recommended anticoagulant treatment. Bleeding was further categorised as major bleeding or clinically relevant non-major bleeding as defined by the 2015 ISTH definitions.[24] Major bleeding was defined as i) fatal bleeding; and/or ii) bleeding in a critical area or organ, such as intracranial, intraspinal, intraocular, retroperitoneal, intra-articular or pericardial, or intramuscular with compartment syndrome; and/or iii) bleeding causing a fall in haemoglobin level of 20 g/L (1.24 mmol/L) or more, or leading to transfusion of two or more units of whole blood or red cells. Clinically relevant non-major bleeding was defined as bleeding i) requiring medical intervention by a healthcare professional; and/or ii) leading to hospitalisation or increased level of care; and/or iii) prompting a face to face (i.e., not just a telephone or electronic communication) evaluation. The diagnosis of the composite primary outcome stroke/TIA/TE and the secondary outcome bleeding was made at the discretion of the treating physician, i.e. either the GP or a specialist. All outcomes were manually extracted from the electronic patient files, after verification using all available information from diagnostic tests, out-of-hours office reports and specialists' letters.

Sample size

We anticipated 80% of the study population to have a CHA2DS2-VASc score ≥ 2, of which 50% would not receive anticoagulants. An incidence rate of stroke/TIA/TE in the total population of 5.3 per 100 person-years was expected. To show a decrease in incidence rate to 3.2 per 100 person-years if anticoagulant treatment of patients with CHA2DS2-VASc ≥ 2 is 100%, with a power of 80%, 1380 patients in total are needed with two years follow-up. To account for clustering in our study design,[25] we used an intra-cluster coefficient of 0.01, based on previous research in general practice.[26] Accounting for 10% loss-to-follow-up, we planned to include 38 practices with an expected average of 75 AF patients per practice (total 2,850 patients known with AF).

Ethics

All data extracted from the electronic patient files was de-identified by a "trusted third party". The decision support used a linking code and only within the practice could it be linked to the individual patient. This study complied with the Data protection law in the Netherlands. The medical ethics committee of the University Medical Centre Utrecht, the Netherlands, judged the CAFe study protocol as exempt from review as it was conducted outside the criteria for the Medical Research Involving Human Subjects Act (WMO). Participating general practitioners provided written informed consent.

Statistical analysis

There was no missing data in any of the baseline variables. Continuous variables were expressed as mean with standard deviation, or median with interquartile range (IQR). Categorical variables were expressed as percentages.

The incidence rate of the composite primary endpoint ischaemic stroke, TIA and/or thromboembolism, in both groups, was expressed as the number of events per 100 person-years follow-up. Patients were censored at time of first stroke/TIA/TE, time of death, loss to follow-up or end of study, whichever came first. To calculate the effect of the intervention accounting for clustering, we used Cox proportional hazard analysis with a random effect for clusters. Proportional hazards assumptions were checked by graphically evaluating the cumulative hazard, and tested for independence of scaled Schoenfeld residuals and time.

Cox proportional hazard was also used to calculate the effect of the automated treatment recommendation on the occurrence of the secondary endpoint, bleeding. The secondary endpoint, treatment according to guideline recommendations, was calculated as the proportion of patients with a CHA2DS2-VASc score 0 or 1 not on anticoagulant therapy, or patients with a CHA2DS2-VASc score ≥ 2 on anticoagulant therapy. Overuse was calculated as the proportion of patient with a score 0 or 1 on anticoagulants, and underuse as the percentage of patients with a score ≥ 2 not using anticoagulants. We used generalised estimating equation linear modelling to test for differences between the proportion of patients treated according to guideline recommendations at baseline and at follow-up, and calculated the effect of the intervention by adding this as an interaction term to the model.

As an additional analysis we assessed underuse according to the prevailing guideline of the European Society of Cardiology, where patients with CHA2DS2-VASc score ≥ 1 were recommended anticoagulation.

All analyses were performed in *R3.3.2* with the package *survival 2.40–1* and *geepack 1.2–1.*

Results are reported according to the CONSORT 2012 statement and extension for cluster randomised trials.[27]

RESULTS

Baseline characteristics

Figure 1 shows the study flow diagram. In total 38 practices were randomised, with in total 2,355 AF patients; 1,129 in the 19 index practices and 1,226 in the 19 reference practices (median of 49 (IQR 37 – 66) and 44 (IQR 28 – 63) AF patients per practice, respectively).

TABLE 1. Baseline characteristics of the individual patients in the index group and the reference group

		Index group n = 1,129 (%)	**Reference group n = 1,226 (%)**
Median Age (IQR)		77 (68 – 84)	77 (67 – 84)
Age < 65 years		194 (17.2)	214 (17.5)
Age 65 – 75 years		277 (24.5)	321 (26.2)
Age ≥ 75 years		658 (58.3)	691 (56.4)
Female sex		558 (49.4)	566 (46.2)
Heart failure		215 (19)	217 (17.7)
Hypertension		666 (59)	745 (60.8)
Diabetes		225 (19.9)	304 (24.8)
Prior Stroke/TIA		178 (15.8)	208 (17)
Vascular disease §		290 (25.7)	322 (26.3)
Anticoagulant use *		992 (87.9)	1076 (87.8)
Vitamin K antagonist		978 (98.6)	1068 (99.3)
Direct oral anticoagulant		37 (3.7)	32 (3.0)
CHA2DS2-VASc score	*0*	59 (5.2)	68 (5.5)
	1	104 (9.2)	103 (8.4)
	2	168 (14.9)	213 (17.4)
	3	242 (21.4)	229 (18.7)
	4	263 (23.3)	251 (20.5)
	5	147 (13)	196 (16)
	6	89 (7.9)	110 (9.0)
	7	44 (3.9)	37 (3.0)
	8	9 (0.8)	17 (1.4)
	9	4 (0.4)	2 (0.2)

IQR = Interquartile range

* Percentage exceeds100% due to prescriptions of multiple drugs in the year before baseline data collection

§ Vascular disease is defined as coronary heart disease, peripheral vascular disease or previous thromboembolism. See Supplemental Material 1.

The baseline characteristics of the individual patients are presented in Table 1. The median age was 77 in both study arms. The median CHA2DS2-VASc scores were 3 (IQR 2.0 – 5) and 3.5 (IQR 2.0 – 5) in the index and control group respectively. Underuse was 8.8% in the total index group (73 (7.6%) of the 966 (85.6%) patients with a CHA2DS2-VASc score ≥ 2) and 8.5% in the total control group (83 (7.9%) of 1055 (86.1%) patients with a score ≥ 2). According to the ESC guideline recommending anticoagulation for patients with CHA2DS2-VASc score ≥ 1, underuse was 9.2% in the total index group (104 (9.7%) of the 1070 (94.8%) patients with a CHA2DS2-VASc score ≥ 1) and 9.1% in the total control group (111 (9.6%) of 1058 (94.5%) patients with a score ≥ 1).

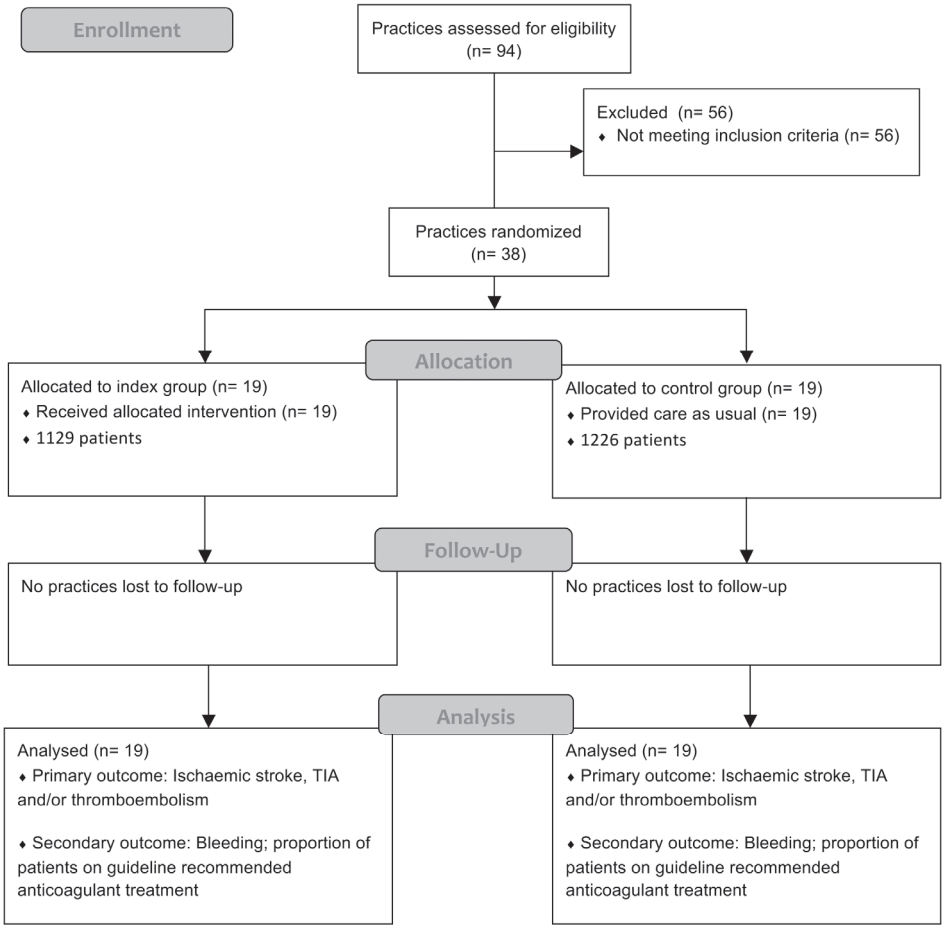


FIGURE 1. CONSORT study flow diagram

Ischaemic stroke, TIA and thromboembolism

During a median follow-up of 2.7 (IQR 2.3 – 3.0) years, the composite of ischaemic stroke, TIA and/or thromboembolism occurred in 97 (4.1%) patients in total. The incidence rate in the index group was 2.0 per 100 person-years, and in the reference group 1.4 per 100 person-years (HR 1.3, 95% CI 0.8 – 2.1). Of the 54 strokes/TIA/TE in the index group, five (9.3%) occurred in patients not on anticoagulation defined as high risk (underuse), and three (5.6%) in patients on anticoagulation defined as low risk (overuse). Forty-three strokes/TIA/TE occurred in the reference group; one (2.3%) in a patient not on anticoagulants while indicated (underuse) and two (4.7%) on anticoagulants while not indicated (overuse).

Bleeding

The incidence rate of major bleeding was 0.79 per 100 person-years in the index group, and 0.82 per 100 person-years in the reference group (HR 0.9, 95% CI 0.5 – 1.9). For clinically-relevant non-major bleeding, these numbers were 2.8 and 2.5 per 100 person-years, respectively (HR 1.1, 95% CI 0.9 – 1.4). For any bleeding, this was 2.9 in the index group, and 2.5 in the reference group (HR 1.2, 95% CI 0.9 – 1.5). Of the twenty-two bleeds in the index group, one (4.5%) occurred in a high risk patient not using anticoagulation, another occurred in a low risk patient using anticoagulation. In the reference group, two (8.0%) of the 25 bleeds occurred in patients at high risk not on anticoagulation.

Change in anticoagulant use and treatment according to guideline recommendations

At any time during follow-up, in 317 patients (13.5%) anticoagulant treatment was (temporarily) initiated or stopped; or changed from VKA to DOAC or vice versa. This was 12.6% in the index group and 14.3% in the reference group. In total 78 (3.3%) patients started a DOAC.

Ten of the 138 new users initiated anticoagulants after experiencing a stroke/TIA/TE. For the 71 patients that discontinued anticoagulants during the study, 18 did so after experiencing a bleeding.

Table 2 shows the proportion of patients treated according to the CHA2DS2-VASc, and the proportion overuse and underuse of anticoagulants in both groups. No statistically significant change was observed for the proportion of patients treated according to the guideline recommendations ($p = 0.08$), or the effect of the intervention (p-value for interaction = 0.52)

TABLE 2. Correct use, overuse and underuse of anticoagulants at baseline and after a median follow-up of 2.2 years based on CHA2DS2-VASc *

		Index n = 1129 (%)	Reference n = 1226 (%)
Baseline	*Correct use*	957 (84.8)	1039 (84.7)
	Underuse	73 (6.5)	83 (6.8)
	Overuse	99 (8.8)	104 (8.5)
After 2.2 years of follow-up	*Correct use*	958 (84.9)	1029 (83.9)
	Underuse	81 (7.2)	100 (8.2)
	Overuse	90 (8)	97 (7.9)

* Correct use is defined as a CHA2DS2-VASc score 0 or 1 not on anticoagulant therapy, or patients with a CHA2DS2-VASc score ≥ 2 on anticoagulant therapy. Overuse is defined as a score 0 or 1 on anticoagulants. Underuse is defined as a score ≥ 2 not using anticoagulants.

DISCUSSION

In this cluster-randomised trial among 2,355 patients with atrial fibrillation, a CHA2DS2-VASc based decision support did not seem to affect the hazard of ischaemic stroke, TIA and/or thromboembolism. The incidence rate of ischaemic stroke/TIA/TE was 1.96 in the index group and 1.42 in the reference group (HR 1.3, 95% CI 0.8 – 2.1). In addition, no difference was observed in the rate of bleeding, or in treatment according to guideline recommendations.

Several previous studies investigated the effect of automated decision support, both in the secondary and primary care setting (Table 3). In secondary care patients with AF, two recent studies found a positive effect. The electronic alert system evaluated by Silbernagel et al.,[28] recommending oral anticoagulation for those with a CHA2DS2-VASc score ≥ 1 in men or ≥ 2 in women (entailing 96% of the study population), resulted in an improvement in anticoagulant drug prescription (22% in the index group versus 15.9% in the control group, relative risk (RR) 1.4, $p = 0.02$). Wang et al.[29] found an increase in antithrombotic use in patients eligible for anticoagulants (89%) based on a CHA2DS2-VASc score ≥ 1 from 77% to 89% ($p < 0.001$).

In general practice, however, results are less positive. In a Dutch study from 2017, Arts et al.[17] found in an AF population with a baseline underuse of anticoagulants of 47% that a decision support tool did not improve adherence to the guideline recommending anticoagulants for those with CHA2DS2-VASc score ≥ 2. Similarly Eckman et al.[19] showed that underuse based on a decision analytic model recommending most patients with a CHA2DS2-VASc score ≥ 2 anticoagulation decreased from 44.7% to 44.5% ($p = 0.59$) after one year in the intervention practices. In a study by Holt et al.[20] a 37% underuse based on anticoagulants for a CHADS2 ≥ 2 reduced to 34% after 6 months of using automated screen reminders

TABLE 3. Overview of studies evaluating decision support systems in the anticoagulant management of patients with atrial fibrillation

Author and year of publication	Design	Setting	Sample size	Treatment recommendation	High risk patients	High risk patients and OAC use	Intervention	Primary Outcome	Effect
Arts 2017	cRCT	GP	781 patients in 19 clusters	OAC for CHA2DS2-VASc ≥ 2	80% CHA2DS2-VASc ≥ 2	47% non-adherence in the total study population	On-screen notifications at ToV	Guideline adherence	No improvement in guideline adherence after 11 months
Bajorek 2016	cRCT	GP	393 patients in 25 clusters	OAC for CHA2DS2-VASc ≥ 1 if at low bleeding risk	93.1% CHA2DS2-VASc ≥ 1	92% OAC use in the total study population	Computerised treatment recommendation	Change in antithrombotic therapy	Increase in OAC use, decrease in antiplatelet use during 1 year
Cook 2015	Cohort with historic controls	Hospital	268 newly-diagnosed AF patients, 226 controls	OAC for CHADS2 ≥ 2	51% CHADS2 ≥ 2	n.a.	Notification system at ToV	OAC prescription to eligible patients	No difference in rate of OAC 30 days after diagnosis
Eckman 2016	cRCT	GP	1493 patients in 15 clusters	Using decision analytics based on stroke and bleeding risk, with OAC recommended for 28% patients with CHA2DS2-VASc ≥ 1 and 93% with ≥ 2	~82% OAC indicated based on decision analytics	42% discordant antithrombotic therapy	Summary report and decision support	Discordant antithrombotic therapy	No improvement in discordant therapy after 1 year
Holt 2017	cRCT	GP	6429 patients in 46 clusters	OAC for CHADS2 ≥ 2	83% CHADS2 ≥ 2	37% underuse	Screen reminder at ToV	Proportion OAC use in eligible patients	No difference in OAC prescription after 1 year

TABLE 3. Overview of studies evaluating decision support systems in the anticoagulant management of patients with atrial fibrillation (*continued*)

Author and year of publication	Design	Setting	Sample size	Treatment recommendation	High risk patients	High risk patients and OAC use	Intervention	Primary Outcome	Effect
Robson 2014	Observational cohort	GP	~4000 patients in 139 practices	OAC for CHA2DS2-VASc ≥ 1	n.r.	~48% underuse	Education, feedback and decision support	Proportion OAC use in high risk patients	7.2% absolute increase in OAC use during 2 years
Silbernagel 2016	RCT	Hospital	889 OAC-naïve patients	OAC for CHA2DS2-VASc ≥ 1 in men and ≥ 2 in women	96.3% CHA2DS2-VASc ≥ 1 in man and ≥ 2 in women	n.a.	Electronic alert system	Adequate OAC prescription	More adequate OAC prescription (RR 1.4) during 1 year
Wang 2017	Observational cohort	Hospital	251 patients	OAC for CHA2DS2-VASc ≥ 1 if at low bleeding risk	89% eligible for OAC	11% underuse	Computerised treatment recommendation	Antithrombotic therapy at discharge	Significant increase in antithrombotic use

OAC= oral anticoagulant; cRCT = cluster randomised clinical trial; RCT = (parallel) randomised clinical trial; GP = general practice; n.a. = not applicable; ToV = time-of-visit; n.r. = not reported; RR = relative risk

for anticoagulant use (adjusted difference with control group 1.21%, $p = 0.21$). Another study applied a decision tool to AF patients in general practice in Australia, recommending anticoagulants for patients with a CHA2DS2-VASc ≥ 1 and low bleeding risk. They found a slight increase in anticoagulant use from 89.3% to 92.2% ($p = 0.02$).[18]

In contrast to previous research, our study evaluated a decision support specifically presenting GPs with treatment recommendations for *all* their AF patients, i.e. irrespective of any consultation. We used the clinically relevant endpoint ischaemic stroke, TIA and/or thromboembolism, as opposed to underuse of anticoagulants alone. Furthermore, automated decision support often relied on diagnosis codes in electronic patient files and concerns over inaccuracies have been raised.[30] To avoid this, data in our study were manually checked for correctness of AF status, CHA2DS2-VASc score and anticoagulation status. Despite these strengths, however, our study adds to the fairly consistent evidence base that improving stroke prevention in AF using automated decision support may not be effective. Several points need further attention in this respect.

First, one promising exception to the ineffectiveness of decision support in AF is the study by Robson et al.[31] It describes positive trends in anticoagulant use in general practice after the introduction of an extensive program consisting of guidance and education, evaluative feedback, and computer prompted reminders. The proportion of AF patients with CHA2DS2-VASc ≥ 1 (denominator of the total population unknown) using anticoagulants increased from 53% to 60% after introduction, though the observational nature of that study hampers drawing any definitive conclusions. It could be hypothesised that decision support alone (either computer prompts at the time of consultation or a list of treatment recommendations for all AF patients) is not sufficient for improving anticoagulant use in AF and that more extensive interventions are needed.

Second, importantly, our study showed at the outset a low proportion of underuse of anticoagulants of 6.6% at baseline. Previous practice studies found underuse as high as 30%[15], a substantial difference even when taking into consideration the change in definition of high risk patients from CHADS2 to the currently used CHA2DS2-VASc. More contemporary studies, however, showed low rates of underuse similar to our study.[32,33] With such initial high rates of stroke prevention according to CHA2DS2-VASc at the outset, it is difficult to proof a benefit of any intervention aiming to further increase anticoagulant use. Additionally, despite careful consideration of an automatically provided treatment recommendation, GPs in shared decision-making with their patient may still deviate from guideline recommendations for valid reasons. Many factors such as comorbid disorders, psychosocial circumstances, or patient preferences may lead to taking the 'cal-

culated risk' of stroke by not taking anticoagulants although indicated. On the other hand, we previously demonstrated that underuse of anticoagulants in AF in general practice seems less related to such patient characteristics and more related to problems in organising care (in particular defining the physician responsible for making anticoagulant decisions).[34] Our intervention aimed to overcome such reasons for underuse of anticoagulants but did not show an effect on guideline recommended treatment.

A third possible explanation for the ineffectiveness of our intervention is that the GPs in the index group did not pay sufficient attention to the treatment recommendations altogether. This has also been shown in previous studies that reported low uptake (< 5%) of decision support tools.[17,35] While our study benefitted from by-passing a common pitfall of 'too many notifications'[36] by providing a one-time list with individual treatment recommendations, this list could still have easily been neglected by the GPs.

Strengths and limitations

Strengths of our study include the large and representative sample of community-dwelling AF patients, with on average more than 2 years follow-up and the clinically relevant primary outcome, ischaemic stroke. We furthermore manually verified correctness of the AF diagnosis, the CHA2DS2-VASc score and the anticoagulant treatment status at baseline by scrutinising the complete electronic patient file, ascertaining the proportion of anticoagulant underuse and thereby ensuring that the automatically generated treatment recommendations were based on valid data.

A limitation is the lower statistical power than calculated a priori as the proportion of underuse in our study at 6% was far lower than expected. As a result the incidence of stroke and hence statistical power was low, though we are confident that an even larger study population would not result in different inferences. Additionally, due to our pragmatic approach, the treatment recommendation could easily be neglected, although all GPs were highly motivated to participate in the study and we aimed not to interfere with daily clinical practice. It cannot be excluded, however, that a more stringent, labour intensive intervention such as described by Robson et al.[31] may have an effect on stroke incidence.

Clinical implication and future considerations

In our study we found high rates of appropriate anticoagulant use in patients with AF from the community, and no effect of implementing an automated decision support on stroke risk or in the correctness of prescribed anticoagulants.

Making additional improvement to an already highly accurate use of anticoagulants according to guidelines is a challenge for future research. Insight in the amount of underuse in a particular care setting may be needed before any proposed intervention is evaluated. Perhaps (repeated) education or quality auditing, preferably in small peer-groups,[37] or the implementation of multidisciplinary care programs[38] for (anticoagulant) management could have a beneficial effect. Several studies evaluated, for instance, a consultant or pharmacist-led intervention with positive results on correct anticoagulant treatment.[39–41] Finally, the addition of patient involvement in a study protocol and decision-making on optimal anticoagulant use is currently lacking, and this approach may perhaps yield unexpected positive effects.

CONCLUSION

In this study in community-dwelling patients with atrial fibrillation, the use of anticoagulants was high. Providing general practitioners with automated decision support aimed to optimise a CHA2DS2-VASc based treatment recommendation neither resulted in a reduction in incidence of stroke, TIA and/or thromboembolism, nor did it affect bleeding risk or guideline recommended anticoagulant treatment.

FUNDING

This study was funded by an unrestricted institutional grant from Boehringer Ingelheim. The funding source had no role in the design, conduct, analyses, or reporting of the study or in the decision to submit the manuscript for publication.

CONFLICT OF INTEREST

GJG is supported by a VENI grant from The Netherlands Organization for Scientific Research (ZONMW 016.166.030). All other authors declared no conflict of interest.

REFERENCES

1 Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, Singer DE. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* 2001; 285: 2370–5.

2 Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991; 22: 983–8.

3 Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Ann Intern Med* 2007; 146: 857–67.

4 Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, Pogue J, Reilly PA, Themeles E, Varrone J, Wang S, Alings M, Xavier D, Zhu J, Diaz R, Lewis BS, Darius H, Diener H-C, Joyner CD, Wallentin L, et al. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2009; 361: 1139–51.

5 Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M, Al-Khalidi HR, Ansell J, Atar D, Avezum A, Bahit MC, Diaz R, Easton JD, Ezekowitz JA, Flaker G, Garcia D, Geraldes M, Gersh BJ, Golitsyn S, Goto S, et al. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011; 365: 981–92.

6 Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, Breithardt G, Halperin JL, Hankey GJ, Piccini JP, Becker RC, Nessel CC, Paolini JF, Berkowitz SD, Fox KAA, Califf RM, ROCKET AF Investigators. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med* 2011; 365: 883–91.

7 Giugliano RP, Ruff CT, Braunwald E, Murphy SA, Wiviott SD, Halperin JL, Waldo AL, Ezekowitz MD, Weitz JI, Špinar J, Ruzyllo W, Ruda M, Koretsune Y, Betcher J, Shi M, Grip LT, Patel SP, Patel I, Hanyok JJ, Mercuri M, et al. Edoxaban versus Warfarin in Patients with Atrial Fibrillation. *N Engl J Med* 2013; 369: 2093–104.

8 Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, Castella M, Diener H-C, Heidbuchel H, Hendriks J, Hindricks G, Manolis AS, Oldgren J, Popescu BA, Schotten U, Van Putte B, Vardas P, Agewall S, Camm J, Barón-Esquivias G, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J* 2016; 37: 2893–962.

9 January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, Conti JB, Ellinor PT, Ezekowitz MD, Field ME, Murray KT, Sacco RL, Stevenson WG, Tchou PJ, Tracy CM, Yancy CW, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol*. 2014. pp. e1–76.

10 NICE. Atrial fibrillation: the management of atrial fibrillation. (Clinical guideline 180). NICE; 2014.

11 Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. 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.

12 Turakhia MP, Hoang DD, Xu X, Frayne S, Schmitt S, Yang F, Phibbs CS, Than CT, Wang PJ, Heidenreich PA. Differences and trends in stroke prevention anticoagulation in primary care vs cardiology specialty management of new atrial fibrillation: The Retrospective Evaluation and Assessment of Therapies in AF (TREAT-AF) study. *Am Heart J* 2013; 165: 93–101.e1.

13 Haeusler KG, Gerth A, Limbourg T, Tebbe U, Oeff M, Wegscheider K, Treszl A, Ravens U, Meinertz T, Kirchhof P, Breithardt G, Steinbeck G, Nabauer M, AFNET registry investigators. Use of vitamin K antagonists for secondary stroke prevention depends on the treating healthcare provider in Germany - results from the German AFNET registry. *BMC Neurol* 2015; 15: 129.

14 Kirchhof P, Nabauer M, Gerth A, Limbourg T, Lewalter T, Goette A, Wegscheider K, Treszl A, Meinertz T, Oeff M, Ravens U, Breithardt G, Steinbeck G. Impact of the type of centre on management of AF patients: surprising evidence for differences in antithrombotic therapy decisions. *Thromb Haemost* 2011; 105: 1010–23.

15 Ogilvie IM, Newton N, Welner SA, Cowell W, Lip GYH. Underuse of oral anticoagulants in atrial fibrillation: a systematic review. *Am J Med* 2010; 123: 638–645e4.

16 Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of Clinical Classification Schemes for Predicting Stroke: Results From the National Registry of Atrial Fibrillation. *JAMA* 2001; 285: 2864–70.

17 Arts DL, Abu-Hanna A, Medlock SK, van Weert HCPM. Effectiveness and usage of a decision support system to improve stroke prevention in general practice: A cluster randomized controlled trial. Quinn TJ, editor. *PLoS One* Public Library of Science; 2017; 12: e0170974.

18 Bajorek BV, Magin PJ, Hilmer SN, Krass I. Optimizing Stroke Prevention in Patients With Atrial Fibrillation: A Cluster-Randomized Controlled Trial of a Computerized Antithrombotic Risk Assessment Tool in Australian General Practice, 2012-2013. *Prev Chronic Dis* 2016; 13: E90.

19 Eckman MH, Lip GYH, Wise RE, Speer B, Sullivan M, Walker N, Kissela B, Flaherty ML, Kleindorfer D, Baker P, Ireton R, Hoskins D, Harnett BM, Aguilar C, Leonard AC, Arduser L, Steen D, Costea A, Kues J. Impact of an Atrial Fibrillation Decision Support Tool on thromboprophylaxis for atrial fibrillation. *Am Heart J* 2016; 176: 17–27.

20 Holt TA, Dalton A, Marshall T, Fay M, Qureshi N. Automated Software System to Promote Anticoagulation and Reduce Stroke Risk. *Stroke* 2017.

21 Lamberts H, Wood M, World Organization of National Colleges, Academies, and Academic Associations of General Practitioners/Family Physicians, Party IW. ICPC, international classification of primary care. 1987.

22 Camm AJ, Lip GYH, De Caterina R, Savelieva I, Atar D, Hohnloser SH, Hindricks G, Kirchhof P, ESC Committee for Practice Guidelines (CPG), Kirchhof P, Kolh P. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation. Developed with the special contribution of the European Heart Rhythm Association. *Eur Heart J*. 2012. pp. 2719–47.

23 Dutch College of General Practitioners Guideline Development Group for Atrial fibrillation. Guideline Atrial fibrillation (second partial revision). *Huisarts Wet* 2013: 392–401.

24 Kaatz S, Ahmad D, Spyropoulos AC, Schulman S. Definition of clinically relevant nonmajor bleeding in studies of anticoagulants in atrial fibrillation and venous thromboembolic disease in non-surgical patients: communication from the SSC of the ISTH. *J Thromb Haemost* 2015; 13: 2119–26.

25 Cosby RH, Howard M, Kaczorowski J, Willan AR, Sellors JW. Randomizing patients by family practice: sample size estimation, intracluster correlation and data analysis. *Fam Pract* 2003; 20: 77–82.

26 Knox SA, Chondros P. Observed intra-cluster correlation coefficients in a cluster survey sample of patient encounters in general practice in Australia. *BMC Med Res Methodol* 2004; 4: 30.

27 Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. *BMJ*. 2012. p. e5661.

28 Silbernagel G, Spirk D, Hager A, Baumgartner I, Kucher N. Electronic Alert System for Improving Stroke Prevention Among Hospitalized Oral-Anticoagulation-Naïve Patients With Atrial Fibrillation: A Randomized Trial. *J Am Heart Assoc* 2016; 5: e003776.

29 Wang Y, Bajorek B. Pilot of a Computerised Antithrombotic Risk Assessment Tool Version 2 (CARATV2.0) for stroke prevention in atrial fibrillation. *Cardiol J* 2017; 24: 176–87.

30 Rice R, Roberts L, Fitzmaurice D. Can we trust studies using audit software? A case study of atrial fibrillation audit. *br j gen pract* 2015; 65: e784–5.

31 Robson J, Dostal I, Mathur R, Sohanpal R, Hull S, Antoniou S, Maccallum P, Schilling R, Ayerbe L, Boomla K. Improving anticoagulation in atrial fibrillation: observational study in three primary care trusts. *br j gen pract* 2014; 64: e275–81.

32 Huisman MV, Rothman KJ, Paquette M, Teutsch C, Diener H-C, Dubner SJ, Halperin JL, Ma C, Zint K, Elsaesser A, Bartels DB, Lip GYH, GLORIA-AF Investigators. Antithrombotic Treatment Patterns in Patients with Newly Diagnosed Nonvalvular Atrial Fibrillation: The GLORIA-AF Registry, Phase II. *Am J Med* 2015; 128: 1306–13.e1.

33 Gadsbøll K, Staerk L, Fosbøl EL, Sindet-Pedersen C, Gundlund A, Lip GYH, Gislason GH, Olesen JB. Increased use of oral anticoagulants in patients with atrial fibrillation: temporal trends from 2005 to 2015 in Denmark. *Eur Heart J* 2017; 38: 899–906.

34 van Doorn S, Hartman-Weide F, Geersing G-J, Oudega R, Hoes AW, Rutten FH. Reasons for non-adherence to practice guidelines on stroke prevention in patients with atrial fibrillation: A cross-sectional study in primary care. *Int J Cardiol* 2015; 187: 525–6.

35 Eccles M, McColl E, Steen N, Rousseau N, Grimshaw J. Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. *BMJ* 2002; 325: 941

36 Lugtenberg M, Weenink J-W, van der Weijden T, Westert GP, Kool RB. Implementation of multiple-domain covering computerized decision support systems in primary care: a focus group study on perceived barriers. *BMC Med Inform Decis Mak* 2015; 15: 82.

37 Camm AJ, Pinto FJ, Hankey GJ, Andreotti F, Hobbs FDR, alliance OBOTWCOTAFSP, Pinto FJ, Andreotti F, Hobbs FDR, Csiba L, de Freitas GR, Goto S, Cantú C, Gonzalez-Zuelgaray J, Hacke W, Hu HH, Mantovani L, Yoon B-W, Hu D, Sim K-H. Non-vitamin K antagonist oral anticoagulants and atrial fibrillation guidelines in practice: barriers to and strategies for optimal implementation. *Europace* 2015; 17: 1007–17.

38 Gallagher C, Elliott AD, Wong CX, Rangnekar G, Middeldorp ME, Mahajan R, Lau DH, Sanders P, Hendriks JML. Integrated care in atrial fibrillation: a systematic review and meta-analysis. *Heart* 2017; : heartjnl–2016–310952.

39 Bajorek BV, Krass I, Ogle SJ, Duguid MJ, Shenfield GM. Optimizing the use of antithrombotic therapy for atrial fibrillation in older people: a pharmacist-led multidisciplinary intervention. *J Am Geriatr Soc* 2005; 53: 1912–20.

40 Das M, Panter L, Wynn GJ, Taylor RM, Connor N, Mills JD, Kirchhof P, Gupta D. Primary Care Atrial Fibrillation Service: outcomes from consultant-led anticoagulation assessment clinics in the primary care setting in the UK. *BMJ Open* 2015; 5: e009267.

41 Virdee MS, Stewart D. Optimizing the use of oral anticoagulant therapy for atrial fibrilation in primary care: a pharmacist-led intervention. Int J Clin Pharm 2017; 39: 173–80.

SUPPLEMENTAL MATERIAL 1. DEFINITIONS OF THE CHA2DS2-VASC VARIABLES

Congestive heart failure was defined as the presence of signs and symptoms suggestive of heart failure, with confirmatory findings of cardiac dysfunction on echocardiography, either with preserved or reduced ejection fraction.

Hypertension was defined as a repeated systolic blood pressure measurement of 140 mmHg or higher.

Diabetes was defined as a repeated fasting blood glucose measurement of ≥ 7.0 mmol/L (126 mg/dL) or a non-fasting glucose measurement of ≥ 11.1 mmol/L (200 mg/dL). Previous stroke or TIA was defined as a focal neurological deficit of sudden onset lasting > 24 hours or < 24 hours, respectively.

Vascular disease was defined as either coronary heart disease, peripheral artery disease or previous thromboembolism. Coronary heart disease was defined as prior myocardial infarction (both ST-elevated myocardial infarction or non-ST-elevated myocardial infarction), angina pectoris or prior percutaneous coronary intervention (PCI) or coronary artery bypass graft surgery (CABG). Peripheral artery disease was defined as symptoms of intermittent claudication with ankle-branchial index ≤ 0.9 or prior surgery or percutaneous intervention on the abdominal or thoracic aorta or lower extremity vessels.

4

Prediction models for development of cardiac and non-cardiac adverse events in community-dwelling older patients with atrial fibrillation: a systematic review and independent validation

Sander van Doorn, Annerien Tavenier, Frans H. Rutten, Arno W. Hoes, Karel G.M. Moons, Geert-Jan Geersing

Submitted

ABSTRACT

Purpose: Patients with atrial fibrillation (AF) are at increased risk of many adverse events. Besides prevention of stroke, integrated care aims to prevent all clinically relevant outcomes. Whether models typically used to predict stroke can also identify anticoagulated patients at high risk of these outcomes is unknown.

We aimed to i) describe the risks of cardiac and non-cardiac hospitalisation and mortality in community-dwelling anticoagulated AF patients, ii) search for existing prediction models developed for stroke prediction in AF, and iii) validate these models for predicting mortality and hospitalisation.

Methods: In a cohort of 2,068 AF patients, we calculated incidence rates (IRs) of ischaemic stroke, hospitalisations and mortality. We systematically searched for existing stroke prediction models in AF. We calculated for every model the observed risk per score and c-statistic.

Results: During a median follow-up of 2.7 (IQR 2.2 – 3.0) years, the IR per 100 person-years was 22.1 for hospitalisations and 6.7 for all-cause mortality. Non-cardiac events outnumbered cardiac events (IRs 15.7 versus 7.6 per 100 person-years for hospitalisation, $p < 0.001$ and 5.0 versus 1.7, $p < 0.001$ for mortality). Four stroke prediction models were validated. The proportion of patients considered at 'low risk' by each model ranged from 3% (CHA2DS2-VASc) to 35% (ATRIA). The median c-statistic of all models was 0.65 for mortality outcomes, and 0.57 for hospitalisation outcomes.

Conclusions: In well-anticoagulated community-dwelling AF patients stroke risk is exceeded by risks of hospitalisation and mortality, importantly mainly for non-cardiac causes. The ATRIA model may be considered most suitable for predicting adverse events other than stroke.

INTRODUCTION

Atrial fibrillation (AF) is the most common cardiac arrhythmia with a prevalence of 1–2% in the general population, increasing to over 15% in those aged 85 years and older.[1,2] It is a major risk factor for stroke,[3] and prevention of stroke using oral anticoagulants – i.e. vitamin K antagonists (VKA) or direct oral antagonists (DOAC) – forms the mainstay of (chronic) treatment of AF. To estimate stroke risk and indicate anticoagulant treatment, multiple clinical prediction models have been developed and validated.[4–7] With adequate anticoagulation, patient stroke rates fell below 2% per year in the recent DOAC trials.[8–10]

However, it is increasingly recognised that after initiation of anticoagulation AF patients remain at increased risk of many other adverse events: hospital admission,[11–13] renal failure and death remain higher in patients with atrial fibrillation compared to non-AF patients of similar age.[14] Such adverse events in addition to many other co-morbidities may, through shared pathophysiological processes like inflammation, fibrosis, hypercoagulability and endothelial dysfunction, further accelerate the progression and burden of AF but also of related diseases like for instance heart failure.[15–17]

Thus, not surprisingly, targeting adverse events beyond stroke alone by means of *integrated disease managements* is recommended by the ESC guideline on AF.[6] Hereto, the prediction of risks of both cardiac and non-cardiac hospitalisation and mortality is an important first step. Because in patients with AF a prediction model is already used to predict stroke, it seems ideal to assess whether the same prediction model to predict stroke can also be applied to predict all (cardiac and non-cardiac) adverse events as it is preferable to use a single prediction model. Whether existing prediction models used for estimating stroke risk and initiating anticoagulation can be used to predict other adverse outcomes in anticoagulated AF patients is, however, yet unknown.

The aim of our study was three-fold. First, to describe the risks for cardiac and non-cardiac hospitalisations and mortality in a cohort of older community-dwelling AF patients already treated with oral anticoagulants; second, to perform a systematic literature review to identify and appraise existing prediction models for stroke risk in AF patients, and finally, to subsequently validate these models for predicting clinically relevant adverse (cardiac and non-cardiac) events in the cohort of anticoagulated AF patients.

METHODS

Study population

The cohort of anticoagulated community-dwelling AF patients was part of the CAFe trial, a large prospective cluster-randomised trial evaluating automated decision-support on the treatment and outcome of patients with atrial fibrillation in general practice in the Netherlands (Trial registration number NTR3741, www.trialregister.nl).

General practices in the region Utrecht, the Netherlands were invited to participate in the CAFe study. From February 2013 until September 2014 a planned total of 38 practices enrolled in the study. At baseline, an automated search in the registers of these 38 practices identified all patients diagnosed with AF based on the International Classification of Primary Care (ICPC)[18] code K78 'atrial fibrillation/flutter'. In addition, files were scrutinised containing ICPC codes for cardiac arrhythmias (K97 'Paroxysmal tachycardia' and K80 'Ectopic beats/Extrasystoles') as well as files of patients receiving a prescription of antiarrhythmic drugs (i.e. amiodarone, sotalol, digoxine and flecainide) and oral anticoagulants (i.e. acenocoumarol, phenprocoumon) . Subsequently, for every identified patient, the electronic patient file was manually screened by the researchers for correctness of AF diagnosis, i.e. whether it was indeed confirmed by electrocardiography. For the current study, all patients (already) using oral anticoagulants at baseline were included.

Sex and age (in years) were extracted from the electronic patient file. Correctness of ICPC codes of relevant cardiovascular and non-cardiovascular comorbidities was checked by screening the electronic patient file, and subsequently extracted.

Finally, ATC codes of all drug prescriptions, including oral anticoagulants and platelet inhibitors, in the year previous to study inclusion were collected. Drugs were categorised into pharmacological subgroups based on second or third level ATC code, see Supplemental Material 1.

Outcomes

Follow-up lasted a minimum of two years. At follow-up all electronic patient files of the included patients with AF were manually scrutinised. Using all available information in the electronic patient file, the following outcome variables were recorded in the two years after baseline data collection:

- admittance to a home for the elderly (yes/no);
- hospitalisation (yes/no), further classified as:
  - cardiac hospitalisation
  - non-cardiac hospitalisation including diagnosis at hospital admission

- mortality (yes/no), further classified as:
  - cardiac mortality
  - non-cardiac mortality including cause of death

Patients were censored at the time of the adverse event, at the time of death, at the time of loss to follow up or at the end of the follow-up, whichever came first.

Systematic review of prediction models

We formulated our review aim using the Checklist for critical Appraisal and Data Extraction for Systematic Reviews (CHARMS),[19] see Supplemental Material 2. To search for all relevant prediction models in AF, we used a previously published systematic search[20] on prediction studies and atrial fibrillation from January 1st 2001 until January 1st 2016, see Supplemental Material 3. For this study, we included all studies developing a stroke prediction model, specifically in patients with non-valvular AF that used clinical patient characteristics for individual risk predictions. Models including predictors that cannot easily be obtained in primary or community care – such as echocardiographic measurements or novel laboratory biomarkers – were excluded.

Two independent reviewers (SD, AT) used the Checklist for critical Appraisal and Data Extraction for Systematic Reviews (CHARMS)[19] for data extraction and critical appraisal of the included studies. Any disagreements were resolved by discussion.

Ethics Statement

As only de-identified data obtained from routine care were used, the medical ethics committee of the University Medical Center Utrecht, the Netherlands, judged the CAFe study protocol as exempt from review as it was conducted outside the criteria for the Medical Research Involving Human Subjects Act (WMO). Participating general practitioners provided written informed consent.

Data analyses

In every patient, a predictor was considered present if the electronic patient file contained a respective diagnosis code(s) (ICPC code), whereas a predictor was considered absent when the electronic patient file did not contain a respective diagnosis code. As such, there was (strictly speaking) no missing data for the available ICPC codes.

Incidence rates for each outcome of interest (admittance to a home for the elderly; cardiac, non-cardiac and all-cause hospitalisation; and cardiac, non-cardiac and all-cause mortality) were expressed as the number of events per 100 person-years of follow-up. We tested for differences between the incidence of outcomes from

cardiac causes versus non-cardiac causes using a Poisson test with a *p*-value of 0.05 as indicative for statistical significance.

We validated each prediction model for each outcome separately. Though the reporting of calibration – that is, the agreement between predictions and the observed outcome – is strongly advocated,[21] we a priori anticipated models to predominantly assign patients with a point-based score, as is most conventional for stroke risk prediction models in atrial fibrillation. Consequently, we described the observed incidence of each outcome per score. If available, we categorised scores into risk categories as defined in the original development studies. Discrimination of each model for the separate outcomes was expressed as the C-statistic, indicating the proportion in which a prediction model in a random pair of patients allocates the higher score to the patient that experiences the adverse event. For all prediction models we used Cox proportional hazards models. All analyses were performed in *R 3.3.2* with the package *rms 5.1–0.*

Results are reported in accordance with the TRIPOD statement.[21]

RESULTS

Patient characteristics

In the total number of 38 participating general practices, we identified 2,355 patients with a confirmed diagnosis of AF, of which a total 2,068 (88%) already used anticoagulation at baseline and they were included in the current study. The median age was 78 (interquartile range (IQR) 69 to 84) years and 1,255 (61%) was 75 years or older, and 51% of the patients were male. Of the common cardiovascular

TABLE 1. Characteristics of 2,068 community-dwelling AF patients using anticoagulants

	n = 2,068 (%)
Median age (IQR)	78 (69 – 84)
Age < 65 years	283 (13.7)
Age 65 – 74 years	530 (25.6)
Age ≥ 75 years	1255 (60.7)
Female Sex	1017 (49.2)
Heart failure	419 (20.3)
Hypertension	1296 (62.7)
Diabetes	503 (24.3)
Stroke	368 (17.8)
Renal Disorder	331 (16)
Vascular disease	571 (27.6)

IQR = interquartile range

risk factors, hypertension was most prevalent (63%). See Table 1 for the patient characteristics. At baseline 97.1% of patients used a VKA, and 2.9% used a DOAC (all dabigatran). Supplemental Material I shows the subgroups of most frequently prescribed drugs in the year before the start of the study.

Outcomes

Total follow-up was 5,133 person-years (median follow up 2.7 years, IQR 2.2 to 3.0 years). Despite anticoagulant treatment, stroke occurred in 87 patients with an incidence rate of 1.73 per 100 person-years.

Admissions to home for the elderly and hospital

In total 61 patients were institutionalised to a home for the elderly (incidence rate 1.19 per 100 person-years) and 879 patients were admitted at least once to the hospital (22.1 per 100 person-years). Hospitalisation occurred most frequently for non-cardiac causes as compared to cardiac causes (incidence rates 15.7 and 7.6 per 100 person-years, respectively). Table 2 shows the reasons for non-cardiac hospitalisation. Infectious diseases, internal diseases, cancer and pulmonology conditions accounted for half of all admission. About one third was admitted to surgery or orthopaedics, 116 because of falls of which 75 (65%) had traumatic fractures.

TABLE 2. Non-cardiac causes for hospitalisation

	n (%)		**n (%)**
Cancer	72 (10.5)		
		Breast	8 (11.1)
		Digestive tract	24 (33.3)
		Respiratory tract	8 (11.1)
		Urogenital tract	21 (29.2)
		Neurologic	3 (4.2)
		Hematologic	3 (4.2)
		Other	5 (6.9)
Infectious/Internal disease	232 (34.0)		
		Urogenital tract	28 (12.1)
		Respiratory tract	63 (27.2)
		Digestive tract	16 (6.9)
		Skin	16 (6.9)
		Other infectious disease	12 (5.2)
		Other internal disease	97 (41.8)
Orthopaedics	120 (17.6)		
		Elective	45 (37.5)
		Fractures	75 (62.5)
Surgery	119 (17.4)		

TABLE 2. Non-cardiac causes for hospitalisation (*continued*)

	n (%)		n (%)
		Vascular	27 (22.7)
		Fall, no fractures	41 (34.5)
		Other	51 (42.9)
Pulmonology	36 (5.3)		
		COPD	21 (58.3)
		Other	15 (41.7)
Neurology	40 (5.9)		40 (100)
Urology/Gynaecology	27 (4.0)		27 (100)
Other	37 (5.4)		37 (100)

Mortality

In total 343 patients died (crude incidence rate 6.7/100 person-years), of whom 87 patients (incidence rate 1.7/100 person-years) died from a cardiac cause. In the 258 non-cardiac deaths (incidence rate 5.0/100 person-years) cancer and infectious disease again were the most frequent causes.

Systematic review and included prediction models

The systematic literature search identified 10 studies that were subjected to full text evaluation. Of these, inclusion criteria were not met in four studies, resulting in a final selection of six model development studies. See Figure 1 for a flowchart of the systematic literature review. For the QStroke model[22] data on ethnicity, social deprivation score, smoking status and multiple clinical measurements were not available, as was 'time in therapeutic range' (TTR) in the AMADEUS model.[23] We were unable to validate these models, leaving a total of four models.[4,24–26]

Results of the critical appraisal of the included studies are shown in Supplemental Material 4. The data used to develop the models originated from a prospective cohort (two studies) and from routinely collected insurance claim healthcare data (two studies). All models assigned patients a point-based score, and categorised scores into three (n = 3) or four (n = 1) risk categories. Two models predicted stroke, two models predicted ischaemic stroke and/or thromboembolism, defined as "sudden occlusion of an artery to a visceral organ or extremity"[25] and "peripheral embolism or pulmonary embolism",[4] respectively.

All models included age stratified in two (n = 1), three (n = 1) or more categories (n = 2), previous stroke/TIA, diabetes, heart failure and hypertension. Three considered female sex a risk factor. One model included renal disease.

Two models were internally validated using bootstrapping, one used a random 1/3 sample of the total dataset. Two studies reported on an external validation in an independent population.

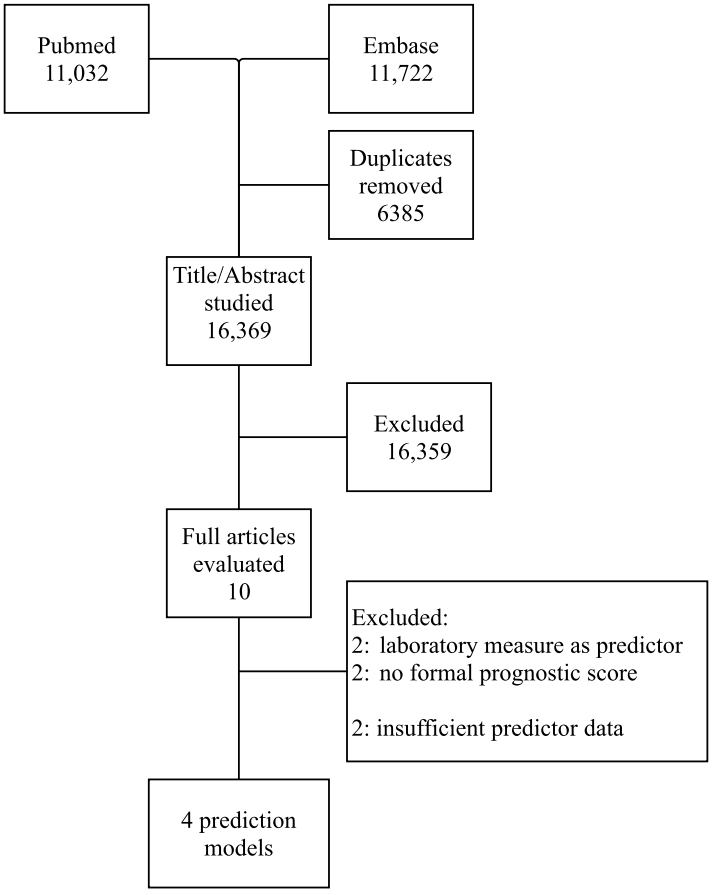


FIGURE 1. Flowchart of the systematic literature review

Validation of the prediction models

Table 3 shows the distribution of patients across scores on CHADS2, Framingham Stroke Risk Score, CHA2DS2-VASc and ATRIA. Using the risk categories as presented in the original development studies, there were substantial differences in the proportion of patients considered at low risk by each model. As a consequence, the observed incidence rates for patients in the low risk category varied widely for all outcomes, see Table 4. This is further illustrated by Figure 2, showing the number of patients and all-cause hospitalisations in the 'low risk' and the 'intermediate/high risk' categories for each model. For instance, the Framingham and the CHA2DS2-VASc models classified only 2.9% and 2.7% patients as low risk, respectively with as a result very few all-cause hospitalisations in this category (6.1 and 4.7 per 100 person-years). Conversely, the ATRIA model considered 34.7% of patients as low risk, with a resulting IR of 15.7%.

TABLE 3. The distribution of 2,068 community-dwelling AF patients using anticoagulants across scores on the CHADS2, Framingham Stroke Risk Score, CHA2DS2-VASc and ATRIA models.

		CHADS2	Framingham	CHA2DS2-VASc	ATRIA
Score	0	224 (10.8)	44 (2.1)	55 (2.7)	58 (2.8)
	1	528 (25.)	17 (0.8)	148 (7.2)	104 (5.0)
	2	669 (32.4)	63 (3.0)	338 (16.3)	69 (3.3)
	3	368 (17.8)	55 (2.7)	424 (20.5)	95 (4.6)
	4	175 (8.5)	64 (3.1)	476 (23.0)	179 (8.7)
	5	85 (4.1)	99 (4.8)	329 (15.9)	213 (10.3)
	6	19 (0.9)	97 (4.7)	187 (9.0)	256 (12.4)
	7		103 (5.0)	79 (3.8)	329 (15.9)
	8		144 (7.0)	26 (1.3)	312 (15.1)
	8+		1381 (66.5)	6 (0.3)	452 (21.9)
Category *	Low	224 (10.8)	61 (2.9)	55 (2.7)	718 (34.7)
	Int.	528 (25.5)	182 (8.8)	148 (7.2)	256 (12.4)
	High	1316 (63.6)	1825 (88.3) §	1865 (90.2)	1094 (52.9)

Numbers are counts (%); Int = intermediate
* Low risk category as defined in the original development studies:
CHADS2 = 0 points; Framingham = 0–1 points; CHA2DS2-VASc = 0 points; ATRIA 0–5 points
§ Two highest risk categories combined

TABLE 4. Incidence rates for each score on the CHADS2, Framingham Stroke Risk Score, CHA2DS2-VASc and ATRIA models.

Table 4.1 *All-cause hospitalisation:*

		CHADS2			Framingham			CHA2DS2-VASc			ATRIA		
		n	py	IR	n	py	IR	n	py	IR	n	py	IR
Score	0	59	609	9.7	7	117	6.0	7	150	4.7	7	150	4.7
	1	200	1353	14.8	3	48	6.2	40	400	10.0	30	225	13.3
	2	306	1662	18.4	15	168	8.9	125	883	14.2	28	146	19.2
	3	175	873	20.1	13	145	8.9	187	1084	17.3	29	214	13.6
	4	87	403	21.6	28	165	16.9	210	1183	17.8	67	379	17.7
	5	37	187	19.8	37	262	14.1	152	761	20.0	84	445	18.9
	6	15	47	32.2	42	254	16.6	104	416	25.0	114	504	22.6
	7				46	262	17.5	35	184	19.0	161	596	27.0
	8				61	364	16.8	14	60	23.3	145	551	26.3
	8+				627	3347	18.7	5	14	36.7	214	762	28.1
Category *	Low	59	609	9.7	10	165	6.1	7	150	4.7	245	1558	15.7
	Int.	200	1353	14.8	56	479	11.7	40	400	10.0	114	504	22.6
	High	620	3171	19.6	813 §	4489	18.1	832	4583	18.2	520	1909	27.2

Table 4.2 *All-cause mortality*

		CHADS2			Framingham			CHA2DS2-VASc			ATRIA		
		n	py	IR	n	py	IR	n	py	IR	n	py	IR
Score	0	10	609	1.6	1	117	0.9	1	142	0.7	1	157	0.6
	1	60	1353	4.4	0	48	0.0	3	336	0.9	2	242	0.8
	2	102	1662	6.1	0	168	0.0	35	700	5.0	5	175	2.9
	3	85	873	9.7	4	145	2.8	51	845	6.0	7	234	3.0
	4	49	403	12.2	6	165	3.6	75	897	8.4	13	417	3.1
	5	32	187	17.2	10	262	3.8	84	579	14.5	18	509	3.5
	6	6	47	12.9	12	254	4.7	57	288	19.8	33	595	5.6
	7				9	262	3.4	26	141	18.4	58	735	7.9
	8				21	364	5.8	10	37	26.7	72	660	10.9
	8+				281	3347	8.4	2	7	30.8	135	903	15.0
Category *	Low	10	609	1.6	1	165	0.6	1	142	0.7	46	1733	2.7
	Int.	60	1353	4.4	10	479	2.1	3	336	0.9	33	595	5.6
	High	274	3171	8.6	333 §	4489	7.4	340	3494	9.7	265	2297	11.5

py = person-years/100; IR = incidence rate per 100 person-years
* Low risk category as defined in the original development studies:
CHADS2 = 0 points; Framingham = 0–1 points; CHA2DS2-VASc = 0 points; ATRIA 0–5 points
§ Two highest risk categories combined

Discrimination

For each individual outcome, discrimination was modest and c-statistics of the different models were comparable. For example, the c-statistics for all-cause hospitalisation ranged from 0.56 (CHADS) to 0.58 (ATRIA).

Discrimination for mortality outcomes was slightly better than for hospitality outcomes for all models (median of c-statistics of all models 0.65 for all mortality outcomes combined compared to 0.57 for all hospitalisation outcomes combined).

DISCUSSION

In a prospective cohort of community-dwelling anticoagulated patients with atrial fibrillation, rates of hospitalisation and mortality were high and more often due to non-cardiac causes rather than to cardiac causes and exceeding the risk of stroke. We validated four easy-to-apply models to predict clinically relevant outcomes other than stroke in anticoagulated AF patients. Substantial differences were found in the number of patients classified as low risk and their respective risk, discrimination between different models was comparable but rather low for each outcome.

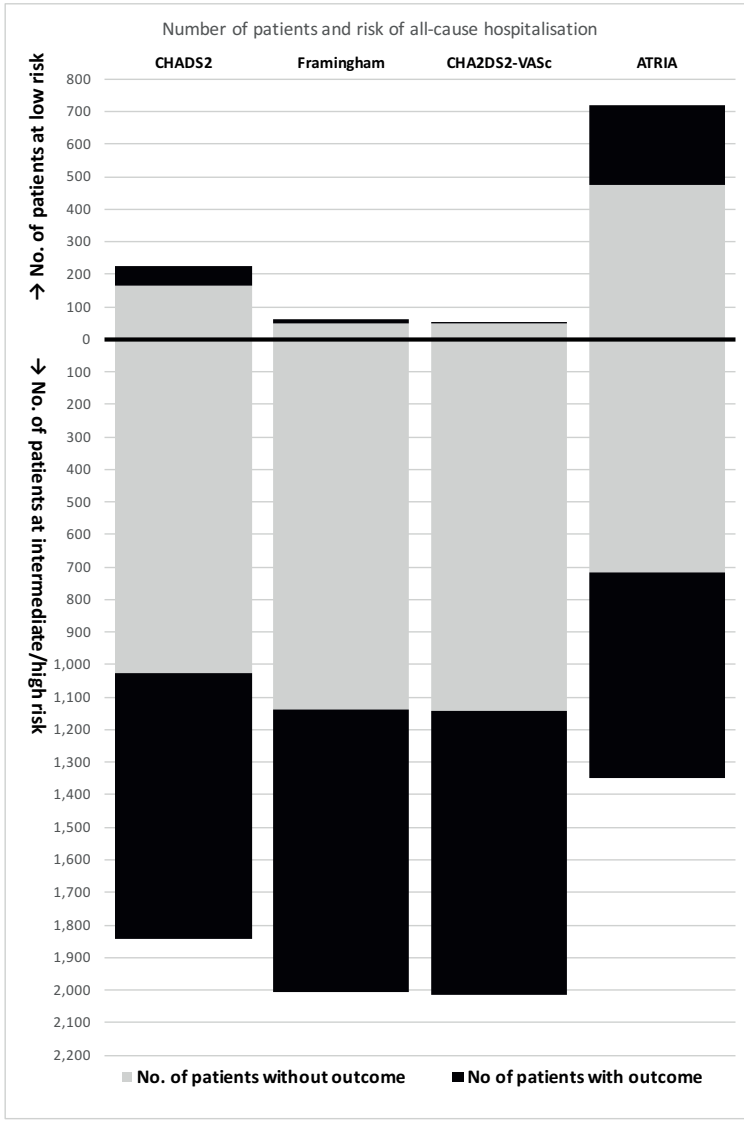


FIGURE 2. Number of patients and all-cause hospitalisations in the 'low risk' and the 'intermediate/high risk' categories for each model

For full appreciation of these results, several points need further consideration. Stroke risk and prevention with oral anticoagulants traditionally form the mainstay in the treatment of atrial fibrillation, with well-addressed concerns about undertreatment.[27] Current literature shows a trend of increasing uptake of oral anticoagulants[28,29] and consequently rates of hospitalisation and mortality now far exceed that of ischaemic stroke. For this study, we only used data of patients

that were already on anticoagulants, and indeed this included over 85% of all AF patients currently enlisted with the participating GPs from this study with an incidence rate of stroke of 1.7 per 100 person-years.

Previous research studied the occurrence of adverse events other than stroke with somewhat conflicting results. For instance, elderly patients in the ORBIT-AF registry and the Euro Heart Survey were more frequently hospitalised than the patients in our study with rates of ~30%/year, and mainly for cardiovascular disease not otherwise specified.[11,30] Conversely, in primary care AF patients in Germany hospitalisation occurred in 18.5% of patients each year, largely similar to our findings.[31] For mortality, similar differences between patient populations are observed. Patients in our cohort died mainly from non-cardiac causes, whereas in the contemporary DOAC trials[32] and in the EORP AF registry in cardiology clinics,[33] death from a cardiac cause was most frequent with IRs around 2.2% and per 100 person-years. These conflicting findings on the comparative risks for cardiac and non-cardiac hospitalisation and mortality in AF patients may be explained by the clinical setting and patient population of each study, further exemplified by a Danish study where particularly in older AF patients non-cardiac hospitalisation outnumbered cardiac hospitalisation[34], similar as in our study. Where in cardiology clinic populations cardiac arrhythmia related admissions and interventions (including e.g. ablation procedures) may (still) predominate, the possibly older and multi-morbid community-dwelling population with chronic AF are at risk of adverse outcomes across multiple organ systems. In our cohort, we predominantly observed hospitalisations for infectious diseases and cancer. This further strengthens recent recommendations, also from the European Society of Cardiology, that managing AF patients requires an *integrated approach* using multidisciplinary teams, likely including at least community care specialists, cardiologists and hospitalists.[6] Considering the growing body of evidence suggesting the interplay of co-morbidities, aging and processes such as inflammation and fibrosis that subsequently interact with AF, its symptoms and its complications,[11,16] this recommendation applies even more so to older chronic AF patients most often predominantly managed in general practice, like the patients included in our cohort.

By targeting relevant adverse events beyond stroke in those patients at highest risk, this integrated approach could be an effective and efficient strategy in managing chronic AF. Herein, risk prediction and identification of patients at increased risk may be considered an important first step. Rather than developing new prediction models to estimate other adverse events beyond stroke risk in individual AF patients, ideally already existing prediction models for stroke can be helpful for this task. It is not uncommon for a prediction model to be used for other purposes

than originally intended, such as predicting fatal cardiovascular events with the Framingham model while developed for non-fatal events only.[35]

This, however, has not extensively been studied yet. Although previous studies have looked into predicting different adverse events, they assessed only one or two prediction models,[36] or included specific populations, e.g. AF patients with heart failure[37] or after atrial fibrillation ablation.[38]

Our study, on the other hand, investigates the value of multiple prediction models based upon a systematic literate review for the prediction of multiple adverse events. It shows that such models can certainly be of value, for three reasons.

First, the prediction models differed substantially in the proportion of AF patients considered at low risk for each outcome and, subsequently, in the observed incidence rates for 'low risk patients'. Although the Framingham and CHA2DS2-VASc model considered only ~3% of AF patients as 'low risk' and seem unfit to identify a reasonable target population for integrated care planning, the ATRIA model considered 35% of patients as 'low risk'. Based on this characteristic it certainly seems a candidate model to be used for predicting adverse events other than stroke.

Second, all models showed similar c-statistics for the several outcomes under study. Although modest at best, these are not different from those commonly reported in studies validating a model for predicting stroke.[20,39] Despite poor discrimination, models for predicting stroke are widely recommended[5–7] and indispensable for tailoring anticoagulant therapy in atrial fibrillation. Provided they identify a realistic proportion of patients as low risk, they may be of equal value for predicting adverse events as well.

Third, the included models mostly consisted of traditional cardiovascular risk factors, though discrimination for non-cardiac outcomes was comparable to cardiac outcomes. It may be hypothesised that a patient's score on each model represents a frailty of adverse outcomes *in general* rather than a cardiovascular risk profile. This clinically relevant insight further strengthens the potential of using stroke risk prediction models in integrated care planning, despite discrimination overall being modest.

Strengths and limitations

Strengths of our study include the large and representative cohort of community-dwelling AF patients using anticoagulation and a follow-up of over two years. Practices covered urban, suburban and rural areas and patients were included irrespective of (co-)treatment by a cardiologist. We furthermore were able to perform a thorough systematic literature review identifying all currently available easy-to-use prediction models for stroke prevention in AF patients.

An important limitation is the lack of biomarker information in our study. Currently under great interest,[40,41] these hold promises for better risk prediction in atrial fibrillation. Our study however aims at risk prediction in general practice, where (novel) biomarkers are not available.

Clinical implication and future considerations

Our results show that in the current era of highly effective anticoagulants, community-dwelling patients with atrial fibrillation are at high risk of adverse events other than ischaemic stroke and that risk prediction using existing models may be an important first step towards integrated AF care. Our findings suggest that the Framingham and CHA2DS2-VASc model may be unable to support physicians in identifying target populations for such care, as AF patients are rarely considered as 'low risk'. The ATRIA seems a promising model classifying meaningful proportions of patients into risk categories, with increasing incidence rates for each relevant adverse outcome. Future steps include the effect of using such a model in integrated care for AF patients. The increasing availability of biomarkers, not only novel tests under scientific evaluation but also established markers available as point-of-care test in general practice, may hold promises for better risk prediction and stratification.

CONCLUSION

In well-anticoagulated community-dwelling patients with atrial fibrillation stroke risk is effectively reduced and thus fairly low, whereas risks of hospitalisation and mortality remain high, importantly mainly for non-cardiac causes. Of the currently available stroke prediction models, the ATRIA model may be considered for predicting adverse events other than stroke.

CONFLICT OF INTEREST

GJG is supported by a VENI grant from The Netherlands Organization for Scientific Research (ZONMW 016.166.030). All funding sources had no role in the design, conduct, analyses, or reporting of the study or in the decision to submit the manuscript for publication. All other authors declared no conflict of interest.

REFERENCES

1 Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, Singer DE. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* 2001; **285**: 2370–5.

2 Heeringa J, van der Kuip DAM, Hofman A, Kors JA, van Herpen G, Stricker BHC, Stijnen T, Lip GYH, Witteman JCM. Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study. *Eur Heart J* Oxford University Press; 2006; **27**: 949–53.

3 Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* Lippincott Williams & Wilkins; 1991; **22**: 983–8.

4 Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. 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.

5 NICE. Atrial fibrillation: the management of atrial fibrillation. (Clinical guideline 180). NICE; 2014.

6 Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, Castella M, Diener H-C, Heidbuchel H, Hendriks J, Hindricks G, Manolis AS, Oldgren J, Popescu BA, Schotten U, Van Putte B, Vardas P, Agewall S, Camm J, Barón-Esquivias G, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *European Heart Journal* 2016; **37**: 2893–962.

7 January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, Conti JB, Ellinor PT, Ezekowitz MD, Field ME, Murray KT, Sacco RL, Stevenson WG, Tchou PJ, Tracy CM, Yancy CW, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol*. 2014. pp. e1–76.

8 Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, Pogue J, Reilly PA, Themeles E, Varrone J, Wang S, Alings M, Xavier D, Zhu J, Diaz R, Lewis BS, Darius H, Diener H-C, Joyner CD, Wallentin L, et al. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2009; **361**: 1139–51.

9 Giugliano RP, Ruff CT, Braunwald E, Murphy SA, Wiviott SD, Halperin JL, Waldo AL, Ezekowitz MD, Weitz JI, Špinar J, Ruzyllo W, Ruda M, Koretsune Y, Betcher J, Shi M, Grip LT, Patel SP, Patel I, Hanyok JJ, Mercuri M, et al. Edoxaban versus Warfarin in Patients with Atrial Fibrillation. *N Engl J Med* 2013; **369**: 2093–104.

10 Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M, Al-Khalidi HR, Ansell J, Atar D, Avezum A, Bahit MC, Diaz R, Easton JD, Ezekowitz JA, Flaker G, Garcia D, Geraldes M, Gersh BJ, Golitsyn S, Goto S, et al. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011; **365**: 981–92.

11 Steinberg BA, Kim S, Fonarow GC, Thomas L, Ansell J, Kowey PR, Mahaffey KW, Gersh BJ, Hylek E, Naccarelli G, Go AS, Reiffel J, Chang P, Peterson ED, Piccini JP. Drivers of hospitalization for patients with atrial fibrillation: Results from the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF). *Am Heart J* 2014; **167**: 735–42. e2.

12 Chamberlain AM, Alonso A, Gersh BJ, Manemann SM, Killian JM, Weston SA, Byrne M, Roger VL. Multimorbidity and the risk of hospitalization and death in atrial fibrillation: A population-based study. *Am Heart J* 2017; **185**: 74–84.
13 Naccarelli GV, Johnston SS, Dalal M, Lin J, Patel PP. Rates and Implications for Hospitalization of Patients ≥ 65 Years of Age With Atrial Fibrillation/Flutter. *The American Journal of Cardiology* 2012; **109**: 543–9.
14 Odutayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG, Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. *BMJ* British Medical Journal Publishing Group; 2016; **354**: i4482.
15 Chimenti C, Russo MA, Carpi A, Frustaci A. Histological substrate of human atrial fibrillation. *Biomed Pharmacother* 2010; **64**: 177–83.
16 Andrade J, Khairy P, Dobrev D, Nattel S. The Clinical Profile and Pathophysiology of Atrial Fibrillation. *Circulation Research* American Heart Association, Inc; 2014; **114**: 1453–68.
17 Skalidis EI, Hamilos MI, Karalis IK, Chlouverakis G, Kochiadakis GE, Vardas PE. Isolated Atrial Microvascular Dysfunction in Patients With Lone Recurrent Atrial Fibrillation. *J Am Coll Cardiol* 2008; **51**: 2053–7.
18 Lamberts H, Wood M, World Organization of National Colleges, Academies, and Academic Associations of General Practitioners/Family Physicians, Party IW. ICPC, international classification of primary care. Oxford University Press, USA; 1987.
19 Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, Reitsma JB, Collins GS. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. *PLoS Med* 2014; **11**: e1001744.
20 van Doorn S, Debray TPA, Kaasenbrood F, Hoes AW, Rutten FH, Moons KGM, Geersing GJ. Predictive performance of the CHA2DS2-VASc rule in atrial fibrillation: a systematic review and meta-analysis. *J Thromb Haemost* 2017; **285**: 2370.
21 Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Ann Intern Med* American College of Physicians; 2015; **162**: W1–73.
22 Hippisley-Cox J, Coupland C, Brindle P. Derivation and validation of QStroke score for predicting risk of ischaemic stroke in primary care and comparison with other risk scores: a prospective open cohort study. *BMJ* British Medical Journal Publishing Group; 2013; **346**: f2573.
23 Lip GYH, Lane DA, Buller H, Apostolakis S. Development of a novel composite stroke and bleeding risk score in patients with atrial fibrillation: the AMADEUS Study. *Chest* 2013; **144**: 1839–47.
24 Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of Clinical Classification Schemes for Predicting Stroke: Results From the National Registry of Atrial Fibrillation. *JAMA* American Medical Association; 2001; **285**: 2864–70.
25 Singer DE, Chang Y, Borowsky LH, Fang MC, Pomernacki NK, Udaltsova N, Reynolds K, Go AS. A new risk scheme to predict ischemic stroke and other thromboembolism in atrial fibrillation: the ATRIA study stroke risk score. *J Am Heart Assoc* 2013; **2**: e000250–0.
26 Wang TJ, Massaro JM, Levy D, Vasan RS, Wolf PA, D'Agostino RB, Larson MG, Kannel WB, Benjamin EJ. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community: the Framingham Heart Study. *JAMA* American Medical Association; 2003; **290**: 1049–56.

27 Ogilvie IM, Newton N, Welner SA, Cowell W, Lip GYH. Underuse of oral anticoagulants in atrial fibrillation: a systematic review. *Am J Med* 2010; **123**: 638–645e4.

28 Gadsbøll K, Staerk L, Fosbøl EL, Sindet-Pedersen C, Gundlund A, Lip GYH, Gislason GH, Olesen JB. Increased use of oral anticoagulants in patients with atrial fibrillation: temporal trends from 2005 to 2015 in Denmark. *European Heart Journal* 2017; **38**: 899–906.

29 Huisman MV, Rothman KJ, Paquette M, Teutsch C, Diener H-C, Dubner SJ, Halperin JL, Ma C, Zint K, Elsaesser A, Bartels DB, Lip GYH, GLORIA-AF Investigators. Antithrombotic Treatment Patterns in Patients with Newly Diagnosed Nonvalvular Atrial Fibrillation: The GLORIA-AF Registry, Phase II. *Am J Med* 2015; **128**: 1306–13.e1.

30 Fumagalli S, Nieuwlaat R, Tarantini F, de Vos CB, Werter CJ, Le Heuzey J-Y, Marchionni N, Crijns HJGM. Characteristics, management and prognosis of elderly patients in the Euro Heart Survey on atrial fibrillation. *Aging Clin Exp Res* 2012; **24**: 517–23.

31 Kirchhof P, Auricchio A, Bax J, Crijns H, Camm J, Diener HC, Goette A, Hindricks G, Hohnloser S, Kappenberger L, Kuck KH, Lip GY, Olsson B, Meinertz T, Priori S, Ravens U, Steinbeck G, Svernhage E, Tijssen J, Vincent A, et al. Outcome parameters for trials in atrial fibrillation: Recommendations from a consensus conference organized by the German Atrial Fibrillation Competence NETwork and the European Heart Rhythm Association. *Europace* 2007; **9**: 1006–23.

32 Gómez-Outes A, Lagunar-Ruíz J, Terleira-Fernández A-I, Calvo-Rojas G, Suárez-Gea ML, Vargas-Castrillón E. Causes of Death in Anticoagulated Patients With Atrial Fibrillation. *J Am Coll Cardiol* 2016; **68**: 2508–21.

33 Lip GYH, Laroche C, Ioachim PM, Rasmussen LH, Vitali-Serdoz L, Petrescu L, Darabantiu D, Crijns HJGM, Kirchhof P, Vardas P, Tavazzi L, Maggioni AP, Boriani G. Prognosis and treatment of atrial fibrillation patients by European cardiologists: One Year Follow-up of the EURObservational Research Programme-Atrial Fibrillation General Registry Pilot Phase (EORP-AF Pilot registry). *Eur Heart J* The Oxford University Press; 2014; **35**: 3365–76.

34 Christiansen CB, Olesen JB, Gislason G, Lock-Hansen M, Torp-Pedersen C. Cardiovascular and non-cardiovascular hospital admissions associated with atrial fibrillation: a Danish nationwide, retrospective cohort study. *BMJ Open* British Medical Journal Publishing Group; 2013; **3**: e001800.

35 Damen JAAG, Hooft L, Schuit E, Debray TPA, Collins GS, Tzoulaki I, Lassale CM, Siontis GCM, Chiocchia V, Roberts C, Schlüssel MM, Gerry S, Black JA, Heus P, van der Schouw YT, Peelen LM, Moons KGM. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ* British Medical Journal Publishing Group; 2016; **353**: i2416.

36 Naccarelli GV, Panaccio MP, Cummins G, Tu N. CHADS2 and CHA2DS2-VASc risk factors to predict first cardiovascular hospitalization among atrial fibrillation/atrial flutter patients. *Am J Cardiol* 2012; **109**: 1526–33.

37 Melgaard L, Gorst-Rasmussen A, Lane DA, Rasmussen LH, Larsen TB, Lip GYH. Assessment of the CHA2DS2-VASc Score in Predicting Ischemic Stroke, Thromboembolism, and Death in Patients With Heart Failure With and Without Atrial Fibrillation. *JAMA* 2015; : 1–9.

38 Jacobs V, May HT, Bair TL, Crandall BG, Cutler M, Day JD, Weiss JP, Osborn JS, Muhlestein JB, Anderson JL, Mallender C, Bunch TJ. The impact of risk score (CHADS2 versus CHA2DS2-VASc) on long-term outcomes after atrial fibrillation ablation. - PubMed - NCBI. *Heart Rhythm* 2015; **12**: 681–6.

39 Quinn GR, Severdija ON, Chang Y, Singer DE. Wide Variation in Reported Rates of Stroke Across Cohorts of Patients With Atrial Fibrillation. *Circulation* 2017; **135**: 208–19.

40 Hijazi Z, Oldgren J, Siegbahn A, Granger CB, Wallentin L. Biomarkers in atrial fibrillation: a clinical review. *European Heart Journal* 2013; **34**: 1475–80.

41 Oldgren J, Hijazi Z, Lindbäck J, Alexander JH, Connolly SJ, Eikelboom JW, Ezekowitz MD, Granger CB, Hylek EM, Lopes RD, Siegbahn A, Yusuf S, Wallentin L. Performance and Validation of a Novel Biomarker-Based Stroke Risk Score for Atrial Fibrillation. Circulation American Heart Association, Inc; 2016; : CIRCULATIONAHA.116.022802.

SUPPLEMENTAL MATERIAL 1. MOST FREQUENTLY PRESCRIBED DRUGS, PER PHARMACOLOGICAL SUBGROUP

ATC Group	n	%
Beta blocking agents	1397	67,6
Agents acting on the renin-angiotensin system	1121	54,2
Diuretics	1007	48,7
Drugs for peptic ulcer and gastro-oesophageal reflux disease	896	43,3
Lipid modifying agents, plain	885	42,8
Antibacterials for systemic use	775	37,5
Calcium channel blockers	511	24,7
Anxiolytics, hypnotics and sedatives	469	22,7
Drugs for obstructive airway diseases	454	22
Corticosteroids, dermatological preparations	433	20,9
Drugs for constipation	431	20,8
Blood glucose lowering drugs, excl. insulins	372	18
Antiarrythmics, class I and III	355	17,2
Platelet aggregation inhibitors, excl. heparin	323	15,6
Opioids	309	14,9
Corticosteroids for systemic use, plain	304	14,7
Other analgesics and antipyretics	259	12,5
Calcium supplements	252	12,2
Vitamins	225	10,9
Antiinflammatory and antirheumatic products, non-steroids	203	9,8
Decongestants and other nasal preparations for topical use	203	9,8
Antifungals for dermatological use	190	9,2
Antidepressants	183	8,8
Drugs used in benign prostatic hypertrophy	175	8,5
Drugs affecting bone structure and mineralisation	150	7,3
Antihistamines for systemic use	134	6,5
Insulins and analogues	130	6,3
Antacids	119	5,8
Iron preparations	95	4,6
Vitamin B12 and folic acid	85	4,1
Heparin	72	3,5
Antipsychotics	56	2,7

SUPPLEMENTAL MATERIAL 2. CHARMS REVIEW AIM

Item	
1. Type of model	Stroke prediction models for patients with atrial fibrillation
2. Intended scope of review	Identify and appraise models to inform physicians' decision making
3. Type of modelling studies	Prediction model development studies
4. Target population to whom the model applies	Adult patients with non-valvular atrial fibrillation
5. Outcomes to be predicted	Stroke, Stroke/TIA, thrombo-embolism or a combination thereof
6. Time span	One year or longer
7. Intended moment of using the model	Any moment in time after initial diagnosis of atrial fibrillation

SUPPLEMENTAL MATERIAL 3. SEARCH STRATEGY

Search string PubMed:
(((Validat* OR Predict*.ti. OR Rule*) OR (Predict* AND (Outcome* OR Risk* OR Model*)) OR ((History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*)) OR (Decision* AND (Model* OR Clinical* OR Logistic Models/)) OR (Prognostic AND (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor* OR Model*))) OR (stratification OR ("ROC Curve"[Mesh]) OR discrimination OR discriminate OR c-statistic OR (c AND statistic) OR (area under the curve) OR auc OR calibration OR indices OR algorithm OR multivariable)) AND (atrial fibrillation)

Search string Embase:
(((validat* OR predict*.ti. OR rule*) OR (predict* AND (outcome* OR risk* OR model*)) OR ((history/exp OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor*) AND (predict* OR model* OR decision* OR identif* OR prognos*)) OR (decision* AND (model* OR clinical* OR (logistic AND models))) OR (prognostic AND ('history'/exp OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor* OR model*))) OR ('stratification'/exp OR (roc AND curve) OR discrimination OR discriminate OR 'c statistic' OR (c AND statistic) OR (area AND under AND the AND curve) OR 'auc'/exp OR 'calibration'/exp OR indices OR 'algorithm'/exp OR multivariable)) AND ('atrial fibrillation'/exp OR 'atrial fibrillation') AND 'article'/it

SUPPLEMENTAL MATERIAL 4. CRITICAL APPRAISAL OF THE INCLUDED STUDIES

Author and year of publication	Gage 2001 [1]	Wang 2003 [2]	Lip 2010 [3]	Singer 2013[4]
Rule	CHADS2	Framingham	CHA2DS2-VASc	ATRIA
Data source	Insurance beneficiaries	Framingham prospective cohort	Euro Heart Survey	Insurance health plan members
Outcome	Stroke	Stroke, or stroke and death	Stroke/Thrombo-embolism	Stroke/Thrombo-embolism
Predictors	Age > 75 (1)	Age 11 categories (with 0–10 points)	Age > 65 or > 75 (with respectively 1 or 2 points)	Age < 65 / 65–74 / 75–84 / > 85 (with respectively 0 / 3 / 5 / 6 points)
	Recent CHF exacerbation (1)	–	LV or RV dysfunction (1)	Chronic heart failure (1)
	Diabetes type 1 or 2 (1)	Diabetes (5)	Diabetes unspecified (1)	Diabetes unspecified (1)
	–	Female (6)	Female (1)	Female (1)
	Previous Stroke/ TIA (2)	Prior Stroke/TIA (6)	Previous stroke/ TIA/TE (2)	Previous Stroke (adjusted points for Age)
	History of hypertension (1)	Systolic blood pressure 5 categories (0–4)	Hypertension (1)	Hypertension (1)
	–	–	Coronary heart disease or peripheral vascular disease (1)	–
	–	–	–	Proteinuria (1)
	–	–	–	eGFR < 45 or ESRD (1)
Patient selection	Beneficiaries with chronic or recurrend AF during hospitalisation	Patients in the original and offspring cohort of the Framingham Heart Study	Patients visiting cardiology clinics participating in the Euro Heart Survey	Healthplan members with outpatient ICD code atrial fibrillation
Treatment received described	Clearly described	Clearly described	Clearly described	Not described
Data collection	Manual record review	During Framingham follow-up visits, or review of inpatient/ outpatient clinic visits	During clinic visit at enrolment and 1 year follow-up	ICD-9 codes, drug prescriptions and lab measurements extracted from relevant databases

Author and year of publication	Gage 2001 [1]	Wang 2003 [2]	Lip 2010 [3]	Singer 2013[4]
Predictors				
candidate predictors	Independent predictors in either the AFI or SPAF rules (based on data from aspirin clinical trials)	Use of hypertensive medication, prior MI, prior congestive HF, prior stroke/ TIA, current smoking, ECG LVH, diabetes, heart murmur suspected of valvular heart disease except mitral stenosis	Based on Birmingham/ NICE stroke risk stratification schema	10 candidate predictors from literature: age, female sex prior ischaemic stroke, diabetes, heart failure, hypertension, CAD, PAD, urine dipstick proteinuria, and low eGFR or ESRD requiring dialysis. In addition, total WBC as an inflammatory marker and an episode of herpes zoster
definition	Relevant and clear definition	Relevant and clear definition	Relevant and clear definition	Relevant, definition based on ICD code
assessed blinded for outcome	Unclear	Unclear but prospective study design	Unclear but prospective study design	Unclear but prospective study design
Outcome				
definition	Hospitalisation for stroke or TIA, ICD code	Panel of 3 members using all relevant information	As diagnosed by a neurologist	Medical records reviewed by 2 physicians
same outcome in all patients	Yes	Yes	Yes	Yes
assessed blinded for predictors	Unclear	No	Unclear	Unclear
Sample size and number of outcomes	1733 patients 71 admitted for stroke, 23 admitted for TIA	705 patients 111 strokes, 485 stroke or death without censoring OAC; 83 strokes, 382 stroke or death with censoring	1084 patients 25 events	In total sample: 10927 patients and 685 TE events In 2/3 derivation sample: 7284 patients and 456 TE events
Missing data				
% missing data	Not reported	13% missing covariate data	31% of cases with missing outcome data	3.5% WBC, 2.8% eGFR, 22.2% proteinuria
handling missing data	N.A.	Patients excluded	Patients excluded	Lab measurements assumed to be normal
Follow-up	Mean 1.2 years, median 1.0	Mean 4.3 years, mean 4.0 years after censoring OAC	1 year	Median 2.4 years

Author and year of publication	Gage 2001 [1]	Wang 2003 [2]	Lip 2010 [3]	Singer 2013[4]
Model development				
type of model	Exponential survival model	Cox proportional hazard model	Logistic regression model	Cox proportional hazard model
predictor selection	Selection beforehand	Age and blood pressure forced into model, other predictors forward selected with $p > 0.1$	Predictors backward selected with $p > 0.1$, variables in the final model if $p < 0.05$	Backward selection with variables in the final model $p < 0.05$ in > 60% of bootstrap samples
events per variable	> 10	> 10	< 10	> 50
correction for optimism	none	none	none	bootstrapping
Model performance	c-statistic stroke risk per score	c-statistic calibration O:E with Hosmer-Lemeshow test stroke risk per score	c-statistic stroke risk per score	c-statistic Goodness-of-fit for calibration Stroke risk per score
Model evaluation				
internal validation	Bootstrapping	Bootstrapping	none	Random split sample 2/3 derivation, 1/3 validation
external validation	none	none	none	External contemporary ATRIA-CVRN cohort
Results	c-statistic 0.82 (95% C.I. 0.80–0.84) Stroke risk per score	c-statistic 0.66 (SD 0.03) for stroke Stroke risk per score	c-statistic 0.61 (95% C.I. 0.51–0.70) Stroke risk per score	c-statistic in derivation 0.74 (95% C.I. 0.72–0.76); c-statistic in internal validation 0.72 (95% C.I. 0.68–0.75) c-statistic in external validation 0.70 (95% C.I. 0.67–0.72)

CHF = congestive heart failure; LV = left ventricle; RV = right ventricle; TE = thrombo-embolism; eGFR = estimated glomerular filtration rate; ESDR = end stage renal disease; MI = myocardial infarction; HF = heart failure; LVH = left ventricular hypertrophy; CAD = coronary artery disease; PAD = peripheral artery disease; WBC = white blood cell count; OAC = oral anticoagulant; N.A. = not applicable; O:E = observed:expectated

SUPPLEMENTAL MATERIAL REFERENCES

1. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of Clinical Classification Schemes for Predicting Stroke: Results From the National Registry of Atrial Fibrillation. *JAMA*. 2001;285:2864–2870.
2. Wang TJ, Massaro JM, Levy D, Vasan RS, Wolf PA, D'Agostino RB, Larson MG, Kannel WB, Benjamin EJ. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community: the Framingham Heart Study. *JAMA*. 2003;290:1049–1056.
3. Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. 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;137:263–272.
4. Singer DE, Chang Y, Borowsky LH, Fang MC, Pomernacki NK, Udaltsova N, Reynolds K, Go AS. A new risk scheme to predict ischemic stroke and other thromboembolism in atrial fibrillation: the ATRIA study stroke risk score. *J Am Heart Assoc*. 2013;2:e000250–e000250.

5

Opportunistic screening for heart failure with natriuretic peptides in patients with atrial fibrillation; results from an analysis of individual participant data of 4 screening studies

Sander van Doorn, Geert-Jan Geersing, Rogier F. Kievit, Yvonne van Mourik, Loes C. Bertens, Evelien E. S. van Riet, Leandra J. Boonman-Winter, Karel G. Moons, Arno W. Hoes, Frans H. Rutten

Submitted

ABSTRACT

Aims: Heart failure (HF) often co-exists in atrial fibrillation (AF) but is often unrecognised due to overlapping symptomatology. Furthermore, AF can cause elevated natriuretic peptide levels, impairing its diagnostic value for HF-detection. Our aims were twofold; first, to assess the prevalence of previously unknown HF in community-dwelling patients with AF, and second to determine the diagnostic value of the natriuretic peptide NTproBNP for HF-screening in AF patients.

Methods and results: Individual participant data from four HF-screening studies in older community-dwelling persons were combined. Presence or absence of HF was in each study established by an expert panel following the criteria of the European Society of Cardiology. We performed a two-stage patient-level meta-analysis to calculate traditional diagnostic indices.

Of the 1,941 individuals included in the 4 studies, 196 (10.1%) had AF at baseline. HF was uncovered in 83 (43%) of these 196 patients with AF, versus 381 (19.7%) in those without AF at baseline. Median NTproBNP levels of AF patients with and without HF were 744 pg/mL and 211 pg/mL, respectively. At the cut-point of 125 pg/mL, sensitivity was 93%, specificity 35%, and positive and negative predictive values 51% and 86%, respectively. Only 23% of all AF patients had an NTproBNP level below the 125 pg/mL cut-point, with still a 13% prevalence of HF in this group.

Conclusion: The prevalence of HF is high among community-dwelling patients with AF. Given its diagnostic accuracy, screening for HF with NTproBNP is inefficient in AF patients and straightforward echocardiography should be considered.

INTRODUCTION

Atrial fibrillation (AF) and heart failure (HF) are both associated with an increased risk of hospitalisations and mortality, and with high healthcare costs.[1,2] They share common causes, may both, at least partly, be considered as accelerated aging of the heart and one disease likely develops in the presence of the other.[3,4] Patients with concomitant heart failure and atrial fibrillation have a worse prognosis than those with either disease in isolation,[5–12] and if occurring simultaneously it may be necessary to adjust anticoagulation therapy,[13,14] and the prescription or dosage of beta-blockers.[15]

Heart failure, however, is frequently unrecognised in patients with AF. Both diseases share key symptoms, e.g. shortness of breath and reduced exercise tolerance. To improve uncovering heart failure in high risk patients with AF, simple diagnostic approaches are needed, in particular in community healthcare settings where extensive imaging tests such as echocardiography are not readily available.

In patients admitted to the hospital for acute dyspnoea, however, amino-terminal pro B-type natriuretic peptides (NTproBNP) had a lower diagnostic capacity to uncover heart failure in patients with AF, than in those without AF,[16] as both may cause elevated serum B-type natriuretic peptide levels due to (episodes of) increased left ventricular wall stress. It is unknown whether NTproBNP can efficiently be used to screen for heart failure in patients with atrial fibrillation or whether a different diagnostic work-up scheme needs to be considered. We were able to use individual participant data from four screening studies for heart failure in general practice to address the current aims: first, we assessed the prevalence of heart failure in community-dwelling older high-risk patients with atrial fibrillation. Next, in this population we investigated whether a simple initial test with NTproBNP can efficiently be used to screen for heart failure.

METHODS

Data source and study population

We used the individual participant data from four screening studies which each focussed on uncovering heart failure in community-dwelling people aged 60 years and over, and at increased risk for heart failure. The studies included, respectively people (1) ≥ 65 years with chronic obstructive pulmonary disease, (2) ≥ 65 years with a recent contact with the general practitioner for shortness of breath, (3) ≥ 65 years with multimorbidity/polypharmacy in combination with shortness of breath or reduced exercise tolerance (assessed by questionnaire), and (4) ≥ 60 years with

type 2 diabetes. These screening studies were performed in the period 2001 to 2010, and described in detail elsewhere.[18–21] All four studies were approved by the medical ethics committee of the University Medical Center Utrecht, the Netherlands, and all participants in each study gave written informed consent.

We combined anonymised individual participant data from each study into a single dataset. Data harmonisation of variables was not needed because the four individual studies had a homogenous design.

Definitions and measurements

All patients underwent a similar diagnostic assessment in the four original studies with a standardised case record form with information on symptoms, smoking status, and medication use, and measurements of height, weight, and blood pressure. The pulse was taken, followed by auscultation of the heart and lungs, palpation of the apex beat in both supine and left lateral position, and measurement of the jugular venous pressure. Finally, inspection of the legs for oedema and/or signs of venous insufficiency was carried out.

A standard 12-lead electrocardiogram (ECG) was recorded and classified according to the Minnesota coding criteria[22] by a single cardiologist blinded to any other diagnostic information.

Serum concentrations of N-terminal pro B-type natriuretic peptide (NTproBNP) were measured with a non-competitive immune-radiometric assay (Roche, Mannheim, Germany).

Tissue Doppler analysis, M-mode, and two-dimensional transthoracic echocardiography were used in accordance with the American Society of Echocardiography guidelines.[23,24] The left ventricular ejection fraction was calculated quantitatively by tracing (Simpson's), or if necessary semi-quantitatively by two-dimensional visual estimate ("eyeballing"). With pulsed wave Doppler of the mitral and pulmonary venous inflow, and tissue Doppler imaging of the mitral annulus motion the following items were measured: the ratio of peak early (E) diastolic filling velocity to peak of atrial (A) contraction (E/A) ratio, the early diastolic mitral annular velocity (e') at both the septal and lateral wall, and the average of both velocities was used to calculate the E/e' ratio.[25]

Outcome definition

In all studies, an expert panel decided on presence or absence of heart failure. In line with the three recent ESC heart failure guidelines: heart failure was considered present if patients had symptoms (and signs) suggestive of heart failure and structural or functional abnormalities with echocardiography in rest.[26–29] Originally, those with heart failure according to the panel were classified as HF with a reduced

ejection fraction (HFrEF) if the LVEF < 45%, HF with a preserved ejection fraction (HFpEF) if i) LVEF ≥ 45%, and ii) E/e' > 13, and/or LA volume index > 34 ml/m^2, and/or increased wall thickness or LV mass, and as isolated right-sided HF if i) LVEF ≥ 45%, and ii) a calculated systolic pulmonary artery pressure > 40 mmHg in the absence of evident left ventricular failure. For the current study, following the most recent guidance from the European Society of Cardiology, we reclassified it into: HFrEF if LVEF < 40%, HF mid-range ejection fraction (HFmrEF) if LVEF 40–49%, and HFpEF if LVEF ≥ 50%.

The diagnosis of HF was made by the expert panel based on all diagnostic test information including medical history, signs and symptoms, imaging tests including echocardiography, ECG and laboratory tests. For two studies[18,19], the results of NTproBNP were not available for the panel to prevent incorporation bias in the analysis of the added diagnostic value. The panels consisted of an experienced GP, and either two cardiologists or one cardiologist and one pulmonologist. In case of disagreement, the final diagnosis was made by the majority vote. Reproducibility of the panel diagnosis was evaluated by re-assessment of a random sample of 10% of diagnosis and showed high kappa values (range 0.74[30] to 0.92.[18])

Data analysis

Missing data was sparse, but before logistic regression analysis multiple imputation for each study separately was applied.[31] We used descriptive statistics expressing baseline characteristics as proportions or median with interquartile range. To evaluate the diagnostic value of NTproBNP, we divided patients into two groups: 1) those with an ECG-confirmed atrial fibrillation and 2) those without atrial fibrillation. For both groups, we calculated the median levels of NTproBNP with interquartile ranges for patients with and without heart failure and tested for differences using independent sample t-tests with NTproBNP on a log-transformed scale.

To evaluate the diagnostic value of NTproBNP we performed two-stage mixed-effects regression meta-analysis.[32] To quantify discrimination of NTproBNP as a continuous variable, we plotted ROC curves and calculated the c-statistic with 95% confidence interval for each study, assessing variation between studies. Consecutively we performed random-effects meta-analysis to combine the four studies.

Next, we used the recommended exclusionary cut-point of 125 pg/mL (14.75 pmol/L)[29] to calculate sensitivity, specificity for each individual study in all patients and in those with and without AF separately, assessing any variation between studies. Using the bivariate approach[33] we obtained a summarised estimate sensitivity and specificity with 95% confidence intervals for all four studies combined. Using the same bivariate approach, we subsequently calculated

positive predictive value (PPV) and negative predictive value (NPV) as described by Leeflang et al.[34] This bivariate approach, a random-effects method, models pairs of (logit transformed) sensitivity and specificity and predictive values from the studies, thereby incorporating any correlation that might exist between these measures. We calculated the proportion of patients with an NTproBNP level below the 125 pg/mL cut-point and the prevalence of heart failure in this group. Additionally, we applied a higher cut-point of 400 pg/mL as recommended by NICE[35] and previously evaluated against the ESC recommended cut-point of 125 pg/mL.[36] All analyses were performed in *R 3.3.0* with packages *mice 2.25* for multiple imputation, *lme4 1.1–12* for first stage mixed-effects regression analysis, *metafor1.9–8* for random-effects meta-analysis of c-statistic and *mada 0.5.7* for bivariate random-effects meta-analysis of sensitivity and specificity, and positive and negative predictive values.

During the conduct and reporting of our study we used items from the PRISMA-IPD[17] statement where applicable.

RESULTS

A total of 1,941 patients were included. In 28 patients atrial fibrillation was established with the electrocardiogram at the baseline assessment, and another 168 patients were already known with ECG-confirmed AF. Table 1 shows the characteristics of the total study population. Patients with atrial fibrillation were, compared to those without AF, older, more often male and more often had a history of prior myocardial infarction, TIA/stroke, COPD or hypertension.

Nearly half of those with AF used oral anticoagulants (all vitamin K antagonists).

Newly detected non-acute heart failure and NTproBNP levels

In total, 464 patients (24.0%) were diagnosed with previously unrecognised heart failure; 72.2% with HFpEF (LVEF ≥ 50%), 11.1% with HFmrEF (LVEF 40–49%), 14.9% with HFrEF (LVEF < 40%), and 1.8% with right-sided HF. In patients with AF, the prevalence of newly detected HF was 42.5% (HFpEF 66%, HFmrEF 11%, HFrEF 15%, right-sided HF 0.07%). These prevalences were considerably higher than in those without AF: 21.9% (HFpEF 73%, HFmrEF 11%, HFrEF 15%, and right-sided HF 0.01%). Among the four studies separately, the lowest prevalence of heart failure was 15%[21] and the highest 36%.[20]

In total, median NTproBNP levels were 744.2 pg/mL (88.0 pmol/L) in patients with AF and newly discovered HF, 194.5 pg/mL (23.0 pmol/L) in newly detected HF cases without AF ($p < 0.001$), 211.4 pg/mL (25.0 pmol/L) in patients with

TABLE 1. Baseline characteristics of 1,941 older people from the community with an increased risk for unrecognised HF (≥ 60 years and type 2 diabetes, ≥ 65 years and chronic obstructive pulmonary disease, ≥ 65 years and shortness of breath, and ≥ 65 and multimorbidity), divided in those with and without AF

	Total n = 1,941 (%)	**with AF n = 196 (%)**	**without AF n = 1,745 (%)**	***p*-value**
Median age (IQR)	72.3 (67.4–77.7)	75.0 (69.0–79.0)	72.0 (67.1–77.0)	
Male sex	964 (49.7)	117 (59.7)	847 (48.5)	0.004
Current smoking	273 (14.1)	21 (10.7)	252 (14.4)	0.19
Comorbidities				
History of MI	169 (8.7)	25 (12.8)	144 (8.3)	0.047
*COPD**	714 (36.8)	81 (41.3)	633 (36.3)	0.19
Hypertension	1104 (56.9)	118 (60.2)	986 (56.5)	0.36
History of TIA/Stroke	196 (10.1)	34 (17.3)	162 (9.3)	< 0.001
Diabetes†	821 (42.3)	84 (42.9)	737 (42.2)	1.00
Cardiovascular drugs				
Loop diuretics	172 (8.9)	43 (21.9)	129 (7.4)	< 0.001
Thiazide diuretics	465 (24.0)	47 (24.0)	418 (24.0)	1.00
MRA (spironolactone)	30 (1.0)	7 (3.6)	23 (1.3)	0.34
ACE-inhibitors/ARB	772 (39.8)	87 (44.4)	685 (39.3)	0.19
Beta-blockers	557 (28.7)	115 (58.7)	442 (25.3)	< 0.001
Digoxin	27 (1.4)	23 (11.7)	4 (0.2)	< 0.001
Vitamin K antagonists	152 (7.8)	95 (48.5)	57 (3.3)	< 0.001
Platelet inhibitors	581 (29.0)	65 (33.2)	516 (29.6)	0.34

AF = atrial fibrillation, MI = myocardial infarction, COPD = chronic obstructive pulmonary disease, MRA = mineralocorticoid receptor antagonists, ACE-inhibitors = angiotensin convertase enzyme inhibitors, ARB = angiotensin II receptor blockers.

* Prevalence of COPD when excluding one study with only patients with COPD.[18]

† Prevalence of diabetes when excluding one study with only patients with diabetes.[19]

AF but without HF, and 93.0 pg/mL (11.0 pmol/L) in those with neither disease ($p < 0.001$). See Figure 1 for the NTproBNP levels, stratified by the four individual studies.

Discriminative diagnostic value of NTproBNP

There was some variation in the c-statistic of NTproBNP among the four screening studies, see Supplemental Material I for the ROC curves and the c-statistics of NTproBNP for diagnosing heart failure in each study separately. Using a two-stage meta-analytical approach, the summary estimate c-statistic of NTproBNP of patients from all studies together was 0.77 (95% CI 0.68 – 0.84). For patients with and without known AF, these numbers were 0.75 (95% CI 0.74 – 0.81) and 0.76 (95% CI 0.67 – 0.83).

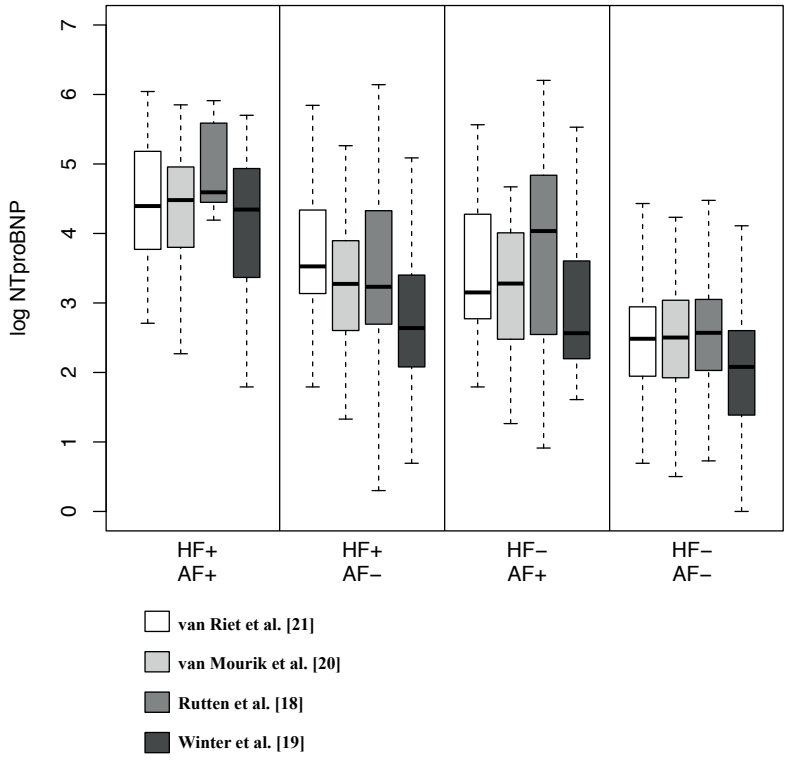


FIGURE 1 NTproBNP in patients with/without HF and AF across studies
log-NTproBNP levels. Whiskers indicate range, boxes indicate median with interquartile range.

NTproBNP cut-point 125 pg/mL and 400 pg/mL

Using the ESC recommended exclusionary cut-point of 125 pg/mL (14.75 pmol/L) [29] for detecting or excluding heart failure, the summary sensitivity and specificity in the total study population were 78% (95% CI 56 – 91%), and 62% (95% CI 52 – 72%), respectively, and the positive predictive value (PPV) 39% (95% C.I. 30 – 48%), and negative predictive value (NPV) 90% (95% CI 76 – 96%). For patients with AF, sensitivity and specificity were 93% (95% CI 78 – 98%), and 35% (95% CI 24 – 48%), respectively. Positive predictive value and negative predictive values were 51% (95% CI 38 – 65%) and 86% (95% CI 65 – 95%), respectively (Table 2). For the values among the four studies separately, see Supplemental Material II.

In total, 53% of patients had an NTproBNP level below the cut-point of 125 pg/mL. For patients with AF this proportion was 23%, and the prevalence of heart failure in this group was 13%. Of patients without AF 53% had an NTproBNP below 125 pg/mL, in which 10% heart failure was uncovered.

TABLE 2. Diagnostic performance measures of NTproBNP using the exclusionary cutpoint of 125 pg/mL.

Group	Sensitivity	Specificity	PPV	NPV	Proportion below cut-point	Prevalence of heart failure in those below cut-point
All patients	78 (56–91)	62 (52–72)	39 (30–48)	90 (76–96)	53 (50–57)	10 (4–21)
AF patients	93 (78–98)	35 (24–48)	51 (38–65)	86 (65–95)	23 (13–29)	13 (3–42)
No AF patients	74 (52–88)	65 (55–73)	36 (27–46)	90 (75–96)	56 (53–60)	10 (4–20)

Numbers are percentages with 95% CI; AF = atrial fibrillation; PPV=positive predictive value; NPV=negative predictive value

Using a previously evaluated higher exclusionary cut-point of 400 pg/mL instead of 125 pg/mL[36] sensitivity and specificity in AF patients were 73% (95% CI 61 – 82%), and 61% (95% CI 49 – 73%), respectively. Positive predictive value and negative predictive values were 59% (95% CI 42 – 74%) and 75% (95% CI 61 – 86%), respectively. The proportion of AF patients with an NTproBNP level below this cut-point of 400 pg/mL increased compared to cut-point 125 pg/mL from 23% to 45%, but the prevalence of missed HF increased as well from 13% to 25%.

DISCUSSION

In this study among 1,941 community-dwelling older people with a high-risk of heart failure, heart failure was uncovered in about 1 in every 4 participants, and nearly 50% in those known with AF. The discriminative value (c-statistic) of NTproBNP for diagnosing heart failure was similar for those with (0.75) and without AF (0.76). Only 23% of all AF patients had an NTproBNP level below the 125 pg/mL cut-point, with still a 13% prevalence of HF in this group.

To the best of our knowledge, this is the first study that evaluated NTproBNP as a tool for selective screening for non-acute heart failure in community-dwelling patients with atrial fibrillation. Two earlier studies showed that in acute dyspneic AF patients in the emergency department setting the c-statistic of NTproBNP for detecting HF was 0.89[16] and 0.91, respectively.[37] The overall diagnostic accuracy for detecting non-acute HF in patients with AF was much lower in patients evaluated in the cardiology outpatient clinic.[38,39] The c-statistic of NTproBNP for detecting left ventricular systolic or diastolic dysfunction or valvular disease was 0.64 [38] and for detecting heart failure 0.78.[39] This is in line with our results (0.75), although both studies selectively included AF patients *suspected* of heart failure, while in our four studies all AF patients were screened. Importantly, these published studies were performed in settings in which echocardiography was

readily available, thus actually lacking the need for evaluation of the diagnostic accuracy of natriuretic peptides. This is different for primary care were access to echocardiography is limited, and where readily available laboratory tests that help to select those requiring referral for echocardiography to definitely confirm HF and determine the type of HF are needed.

Our results suggest that NTproBNP testing for screening for heart failure in high-risk community-dwelling patients with atrial fibrillation is unnecessary given the high prior change of 43% of concurrent HF. With such a high prior risk of HF and the low chance that these patients have a lower NTproBNP value than the exclusionary cut-point of 125 pg/mL (only 23%), NTproBNP does not really help the general practitioner in selecting those needing echocardiography. Instead, in older community patients with AF straightforward echocardiography should be considered.

The underlying mechanism of elevated NTproBNP values is increased wall stress in the left ventricle (LV), and to a far lower degree also the right ventricle and the atria.[40,41] These levels are very high in studies including patients from the emergency department or inpatients, predominantly driven by the abrupt high LV wall stress caused by acute HF, values that are in general much higher than caused by only AF.[16,37,42,43] Also non-acute HF causes a relatively milder left ventricular wall stress, and thus lower NTproBNP levels than in acute HF. Importantly, those with AF have substantially higher levels of NTproBNP than those with sinus rhythm, irrespective of the presence of newly detected slow-onset HF.[38,44]

Several studies tried to assess alternative cut-points for NTproBNP for diagnosing HF in patients with atrial fibrillation presenting at the emergency department. These studies, however, are hampered by the fact that the prior risk of previously unknown HF was very high with 92% and 63%, and thus straightforward echocardiography should always be considered.[16,37] Even at an extremely high NTproBNP cut-points of 3,460 pg/mL the positive predictive value was relatively poor with 79%.[16] In a study in AF patients evaluated in the outpatient setting (prior risk of structural heart disease 73%) a cut-point of 1,764 pg/mL resulted in a sensitivity of 69% and specificity of 77%, a PPV of 89% and an NPV of 50%.[39] Jug et al. reported an 'optimal' cut-point of 1,297 pg/mL[38], however, without mentioning diagnostic accuracy data.

We did not assess an 'optimal' cut-point, but instead assessed the efficiency and proportion of missed cases for the ESC-guideline recommended exclusionary cut-point of 125 pg/mL (14.75 pmol/L) in AF patients and, alternatively, 400 pg/mL. Further increasing the cut-point for patient with AF in our primary care domain would increase the proportion of AF patients with an NTproBNP level below this cut-point, but further increase the proportion of missed cases of heart failure speculating on the likely option that they would not be referred for echocardiography.

Strengths and limitations

The availability of a large dataset, with few missing data, from a representative sample of high-risk patients from the older population from the community, and the complete work-up of participants and applying the same standardised diagnostic procedures and allocation of presence or absence of HF is the major strength of our study.

A limitation is that individual studies did show variation in the prevalence of heart failure and diagnostic measures, and that expressing statistical heterogeneity (e.g. by means of a 95% prediction interval) between only four studies is methodologically not straightforward. We chose to combine all studies nonetheless since they all included clinically similar high risk community-dwelling patients and study methods and materials were similar. Furthermore, information on the type of atrial fibrillation was lacking. Natriuretic peptide levels differ for paroxysmal and persistent atrial fibrillation[45] and the time spent in AF (so-called 'AF burden'). [46] We were unable to take these possible differences into account.

CONCLUSION

In community-dwelling older people with atrial fibrillation, unrecognised HF is very common with nearly 50%, and straightforward echocardiography seems to be the preferred diagnostic strategy to confirm or exclude HF, that is, without prior measurement of NTproBNP levels.

ACKNOWLEDGEMENTS

We would like to thank dr. Thomas P.A. Debray for the valuable input on data analysis.

DISCLOSURE

GJG is supported by a VENI grant from The Netherlands Organization for Scientific Research (ZONMW 016.166.030). All funding sources had no role in the design, conduct, analyses, or reporting of the study or in the decision to submit the manuscript for publication. All other authors declared no conflict of interest.

REFERENCES

1 Writing Group Members, Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, Das SR, de Ferranti S, Després J-P, Fullerton HJ, Howard VJ, Huffman MD, Isasi CR, Jiménez MC, Judd SE, Kissela BM, Lichtman JH, Lisabeth LD, Liu S, et al. Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association. *Circulation* American Heart Association Journals; 2016; 133: e38–60.

2 Chen J, Normand S-LT, Wang Y, Krumholz HM. National and regional trends in heart failure hospitalization and mortality rates for Medicare beneficiaries, 1998-2008. *JAMA* 2011; 306: 1669–78.

3 Maisel WH, Stevenson LW. Atrial fibrillation in heart failure: epidemiology, pathophysiology, and rationale for therapy. *The American Journal of Cardiology* 2003.

4 Geersing G-J, de Groot JA, Reitsma JB, Hoes AW, Rutten FH. The impending epidemic of chronic cardiopulmonary disease and multimorbidity: the need for new research approaches to guide daily practice. *Chest* 2015; 148: 865–9.

5 Ahmed A, Thornton P, Perry GJ, Allman RM, DeLong JF. Impact of atrial fibrillation on mortality and readmission in older adults hospitalized with heart failure. *Eur J Heart Fail* 2004; 6: 421–6.

6 Baldasseroni S, Orso F, Fabbri G, De Bernardi A, Cirrincione V, Gonzini L, Fumagalli S, Marchionni N, Midi P, Maggioni AP. Age-dependent prognostic significance of atrial fibrillation in outpatients with chronic heart failure: data from the Italian Network on Congestive Heart Failure Registry. *Cardiology* 2010; 116: 79–88.

7 Chamberlain AM, Redfield MM, Alonso A, Weston SA, Roger VL. Atrial fibrillation and mortality in heart failure: a community study. *Circ Heart Fail* 2011; 4: 740–6.

8 McManus DD, Hsu G, Sung SH, Saczynski JS, Smith DH, Magid DJ, Gurwitz JH, Goldberg RJ, Go AS. Atrial fibrillation and outcomes in heart failure with preserved versus reduced left ventricular ejection fraction. *J Am Heart Assoc* 2013; 2: e005694.

9 Potpara TS, Polovina MM, Licina MM, Marinkovic JM, Lip GYH. Predictors and prognostic implications of incident heart failure following the first diagnosis of atrial fibrillation in patients with structurally normal hearts: the Belgrade Atrial Fibrillation Study. *Eur J Heart Fail* 2013; 15: 415–24.

10 Rivero-Ayerza M, Scholte Op Reimer W, Lenzen M, Theuns DAMJ, Jordaens L, Komajda M, Follath F, Swedberg K, Cleland JGF. New-onset atrial fibrillation is an independent predictor of in-hospital mortality in hospitalized heart failure patients: results of the EuroHeart Failure Survey. *Eur Heart J* 2008; 29: 1618–24.

11 Senoo K, Lip GYH, Lane DA, Buller HR, Kotecha D. Residual Risk of Stroke and Death in Anticoagulated Patients According to the Type of Atrial Fibrillation: AMADEUS Trial. *Stroke* 2015; 46: 2523–8.

12 Wang TJ, Larson MG, Levy D, Vasan RS, Leip EP, Wolf PA, D'Agostino RB, Murabito JM, Kannel WB, Benjamin EJ. Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: the Framingham Heart Study. *Circulation* Lippincott Williams & Wilkins; 2003; 107: 2920–5.

13 Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, Castella M, Diener H-C, Heidbuchel H, Hendriks J, Hindricks G, Manolis AS, Oldgren J, Popescu BA, Schotten U, Van Putte B, Vardas P, Agewall S, Camm J, Barón-Esquivias G, et al. 2016 ESC Guidelines

for the management of atrial fibrillation developed in collaboration with EACTS. *European Heart Journal* 2016; 37: 2893–962.

14 January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, Conti JB, Ellinor PT, Ezekowitz MD, Field ME, Murray KT, Sacco RL, Stevenson WG, Tchou PJ, Tracy CM, Yancy CW, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol*. 2014. pp. e1–76.

15 Kotecha D, Holmes J, Krum H, Altman DG, Manzano L, Cleland JGF, Lip GYH, Coats AJS, Andersson B, Kirchhof P, Lueder von TG, Wedel H, Rosano G, Shibata MC, Rigby A, Flather MD, Beta-Blockers in Heart Failure Collaborative Group. Efficacy of β blockers in patients with heart failure plus atrial fibrillation: an individual-patient data meta-analysis. *Lancet* 2014; 384: 2235–43.

16 Richards M, Di Somma S, Mueller C, Nowak R, Peacock WF, Ponikowski P, Mockel M, Hogan C, Wu AHB, Clopton P, Filippatos GS, Anand I, Ng L, Daniels LB, Neath S-X, Shah K, Christenson R, Hartmann O, Anker SD, Maisel A. Atrial fibrillation impairs the diagnostic performance of cardiac natriuretic peptides in dyspneic patients: results from the BACH Study (Biomarkers in ACute Heart Failure). *JACC Heart Fail* 2013; 1: 192–9.

17 Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, Tierney JF, PRISMA-IPD Development Group. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *JAMA* 2015; 313: 1657–65.

18 Rutten FH, Moons KGM, Cramer M-JM, Grobbee DE, Zuithoff NPA, Lammers J-WJ, Hoes AW. Recognising heart failure in elderly patients with stable chronic obstructive pulmonary disease in primary care: cross sectional diagnostic study. *BMJ* 2005; 331: 1379–0.

19 Winter LJMB-D, Rutten FH, Cramer MJM, Landman MJ, Liem AH, Rutten GEHM, Hoes AW. High prevalence of previously unknown heart failure and left ventricular dysfunction in patients with type 2 diabetes. *Diabetologia* Springer-Verlag; 2012; 55: 2154–62.

20 van Mourik Y, Moons KG, Bertens LC, Reitsma JB, Hoes AW, Rutten FH. Triage of frail elderly with reduced exercise tolerance in primary care (TREE). a clustered randomized diagnostic study. *BMC Public Health* BioMed Central Ltd; 2012; 12: 385.

21 van Riet EE, Hoes AW, Limburg A, van der Hoeven H, Landman MA, Rutten FH. Strategy to recognize and initiate treatment of chronic heart failure in primary care (STRETCH): a cluster randomized trial. *BMC Cardiovasc Disord* BioMed Central Ltd; 2014; 14: 1.

22 Blackburn H, KEYS A, SIMONSON E, RAUTAHARJU P, PUNSAR S. The electrocardiogram in population studies. A classification system. *Circulation* 1960; 21: 1160–75.

23 Lang RM, Bierig M, Devereux RB, Flachskampf FA, Foster E, Pellikka PA, Picard MH, Roman MJ, Seward J, Shanewise JS, Solomon SD, Spencer KT, Sutton MSJ, Stewart WJ, Chamber Quantification Writing Group, American Society of Echocardiography's Guidelines and Standards Committee, European Association of Echocardiography. Recommendations for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. *Journal of the American Society of Echocardiography : official publication of the American Society of Echocardiography*. 2005. pp. 1440–63.

24 Nagueh SF, Appleton CP, Gillebert TC, Marino PN, Oh JK, Smiseth OA, Waggoner AD, Flachskampf FA, Pellikka PA, Evangelista A. Recommendations for the evaluation of left ventricular diastolic function by echocardiography. *J Am Soc Echocardiogr* 2009; 22: 107–33.

25 Paulus WJ, Tschöpe C, Sanderson JE, Rusconi C, Flachskampf FA, Rademakers FE, Marino P, Smiseth OA, De Keulenaer G, Leite-Moreira AF, Borbély A, Edes I, Handoko ML, Heymans S, Pezzali N, Pieske B, Dickstein K, Fraser AG, Brutsaert DL. How to diagnose diastolic heart failure: a consensus statement on the diagnosis of heart failure with normal left ventricular ejection fraction by the Heart Failure and Echocardiography Associations of the European Society of Cardiology. *Eur Heart J* 2007; 28: 2539–50.

26 McMurray JJV, Adamopoulos S, Anker SD, Auricchio A, Böhm M, Dickstein K, Falk V, Filippatos G, Fonseca C, Gomez Sanchez MA, Jaarsma T, Kober L, Lip GYH, Maggioni AP, Parkhomenko A, Pieske BM, Popescu BA, Rønnevik PK, Rutten FH, Schwitter J, et al. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012. *Eur J Heart Fail* 2012; 14: 803–69.

27 Dickstein K, Cohen Solal A, Filippatos G, McMurray JJV, Ponikowski P, Poole-Wilson PA, Strömberg A, van Veldhuisen DJ, Atar D, Hoes AW, Keren A, Mebazaa A, Nieminen M, Priori SG, Swedberg K, ESC Committee for Practice Guidelines (CPG). ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2008: the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2008 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association of the ESC (HFA) and endorsed by the European Society of Intensive Care Medicine (ESICM). *Eur Heart J*. 2008. pp. 2388–442.

28 Krum H. The Task Force for the diagnosis and treatment of chronic heart failure of the European Society of Cardiology. Guidelines for the diagnosis and treatment of chronic heart failure: full text (update 2005). *Eur Heart J* 2005; 26: 2472–authorreply2473–4.

29 Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, Falk V, González-Juanatey JR, Harjola V-P, Jankowska EA, Jessup M, Linde C, Nihoyannopoulos P, Parissis JT, Pieske B, Riley JP, Rosano GMC, Ruilope LM, Ruschitzka F, Rutten FH, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *European Heart Journal* 2016; 37: 2129–200.

30 van Mourik Y, Bertens LCM, Cramer MJM, Lammers J-WJ, Reitsma JB, Moons KGM, Hoes AW, Rutten FH. Unrecognized heart failure and chronic obstructive pulmonary disease (COPD) in frail elderly detected through a near-home targeted screening strategy. *J Am Board Fam Med* American Board of Family Medicine; 2014; 27: 811–21.

31 Burgess S, White IR, Resche-Rigon M, Wood AM. Combining multiple imputation and meta-analysis with individual participant data. *Stat Med* 2013; 32: 4499–514.

32 Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Stat Med* 2017; 36: 855–75.

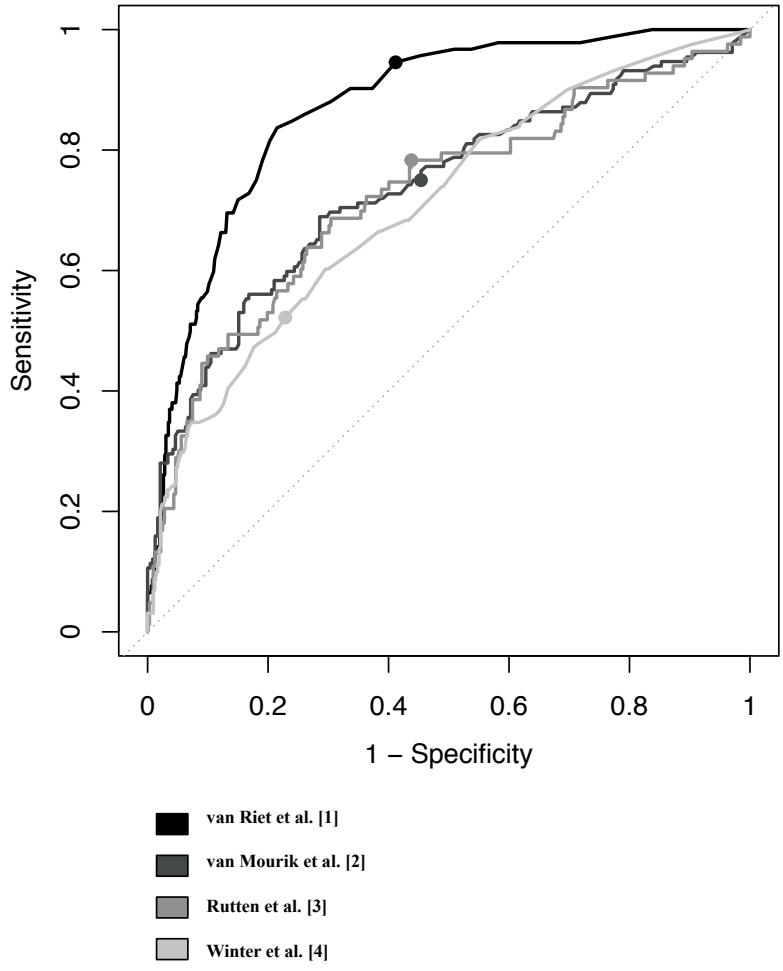
33 Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005; 58: 982–90.

34 Leeflang MMG, Deeks JJ, Rutjes AWS, Reitsma JB, Bossuyt PMM. Bivariate meta-analysis of predictive values of diagnostic tests can be an alternative to bivariate meta-analysis of sensitivity and specificity. *J Clin Epidemiol* 2012; 65: 1088–97.

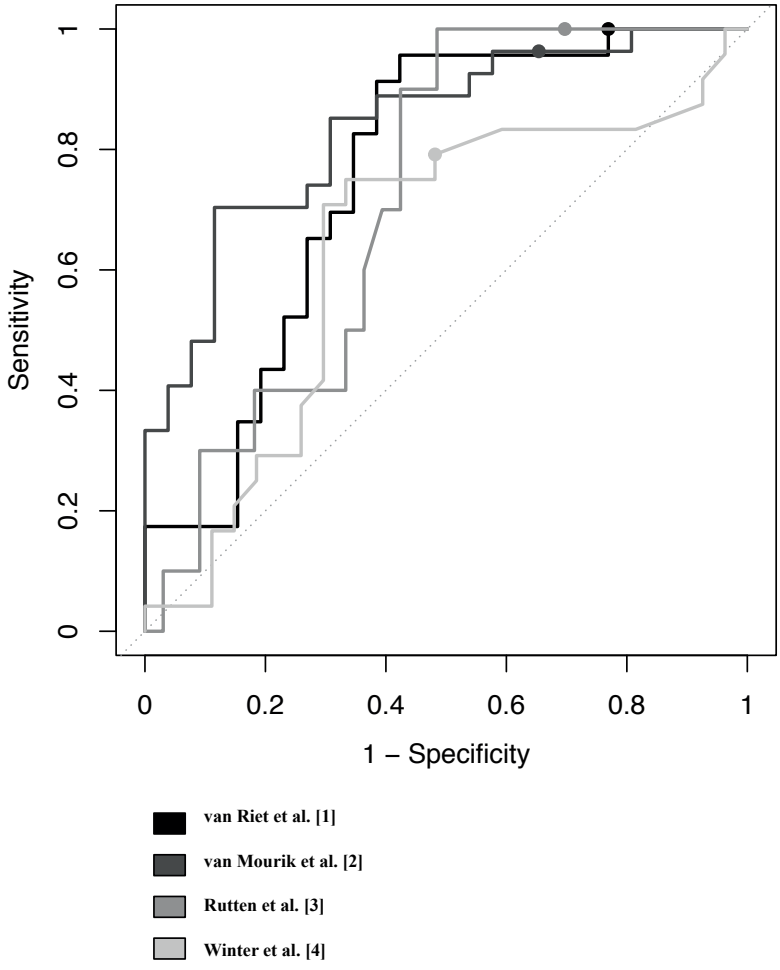
35 National Institute for Health and Clinical Excellence. Chronic Heart Failure. 2010.
36 Taylor CJ, Roalfe AK, Iles R, Hobbs FR, REFER investigators, Barton P, Deeks J, McCahon D, Cowie MR, Sutton G, Davis RC, Mant J, McDonagh T, Tait L. Primary care REFerral for EchocaRdiogram (REFER) in heart failure: a diagnostic accuracy study. *br j gen pract* 2017; 67: e94–e102.
37 Eckstein J, Potocki M, Murray K, Breidthardt T, Ziller R, Mosimann T, Klima T, Hoeller R, Moehring B, Sou SM, Rubini Gimenez M, Morgenthaler NG, Mueller C. Direct comparison of mid-regional pro-atrial natriuretic peptide with N-terminal pro B-type natriuretic peptide in the diagnosis of patients with atrial fibrillation and dyspnoea. *Heart* BMJ Publishing Group Ltd and British Cardiovascular Society; 2012; 98: 1518–22.
38 Jug B, Sebestjen M, Sabovic M, Pohar M, Keber I. Atrial fibrillation is an independent determinant of increased NT-proBNP levels in outpatients with signs and symptoms of heart failure. *Wien Klin Wochenschr* Springer-Verlag; 2009; 121: 700–6.
39 Shelton RJ, Clark AL, Goode K, Rigby AS, Cleland JGF. The diagnostic utility of N-terminal pro-B-type natriuretic peptide for the detection of major structural heart disease in patients with atrial fibrillation. *Eur Heart J* 2006; 27: 2353–61.
40 Ellinor PT, Low AF, Patton KK, Shea MA, MacRae CA. Discordant atrial natriuretic peptide and brain natriuretic peptide levels in lone atrial fibrillation. *J Am Coll Cardiol* 2005; 45: 82–6.
41 Maisel A, Mueller C, Adams K, Anker SD, Aspromonte N, Cleland JGF, Cohen Solal A, Dahlstrom U, DeMaria A, Di Somma S, Filippatos GS, Fonarow GC, Jourdain P, Komajda M, Liu PP, McDonagh T, McDonald K, Mebazaa A, Nieminen MS, Peacock WF, et al. State of the art: Using natriuretic peptide levels in clinical practice. *Eur J Heart Fail* 2008; 10: 824–39.
42 Morello A, Lloyd-Jones DM, Chae CU, van Kimmenade RRJ, Chen AC, Baggish AL, O'Donoghue M, Lee-Lewandrowski E, Januzzi JLJ. Association of atrial fibrillation and amino-terminal pro-brain natriuretic peptide concentrations in dyspneic subjects with and without acute heart failure: results from the ProBNP Investigation of Dyspnea in the Emergency Department (PRIDE) study. *Am Heart J* 2007; 153: 90–7.
43 Linssen GCM, Rienstra M, Jaarsma T, Voors AA, Van Gelder IC, Hillege HL, van Veldhuisen DJ. Clinical and prognostic effects of atrial fibrillation in heart failure patients with reduced and preserved left ventricular ejection fraction. *Eur J Heart Fail* 2011; 13: 1111–20.
44 Freestone B, Gustafsson F, Chong AY, Corell P, Kistorp C, Hildebrandt P, Lip GYH. Influence of atrial fibrillation on plasma von willebrand factor, soluble E-selectin, and N-terminal pro B-type natriuretic peptide levels in systolic heart failure. *Chest* 2008; 133: 1203–8.
45 Chang ICY, Chen LY, Chong JPC, Austin E, Quay CN, Gong L, Mark Richards A, Ling LH. Plasma mid-regional pro-atrial natriuretic peptide and N-terminal pro-brain natriuretic peptide improve discrimination of lone atrial fibrillation. *Int J Cardiol* 2015; 188: 10–2.
46 Plitt DC, Chung EH, Mounsey JP, Schwartz JD, Pursell IW, Gehi AK. Relation of Atrial Fibrillation Burden and N-Terminal Pro-Brain Natriuretic Peptide. The American Journal of Cardiology 2013; 111: 1315–8.

SUPPLEMENTAL MATERIAL 1. ROC CURVES FOR ALL PATIENTS, THOSE WITH AF AND THOSE WITHOUT AF FOR THE INDIVIDUAL STUDIES

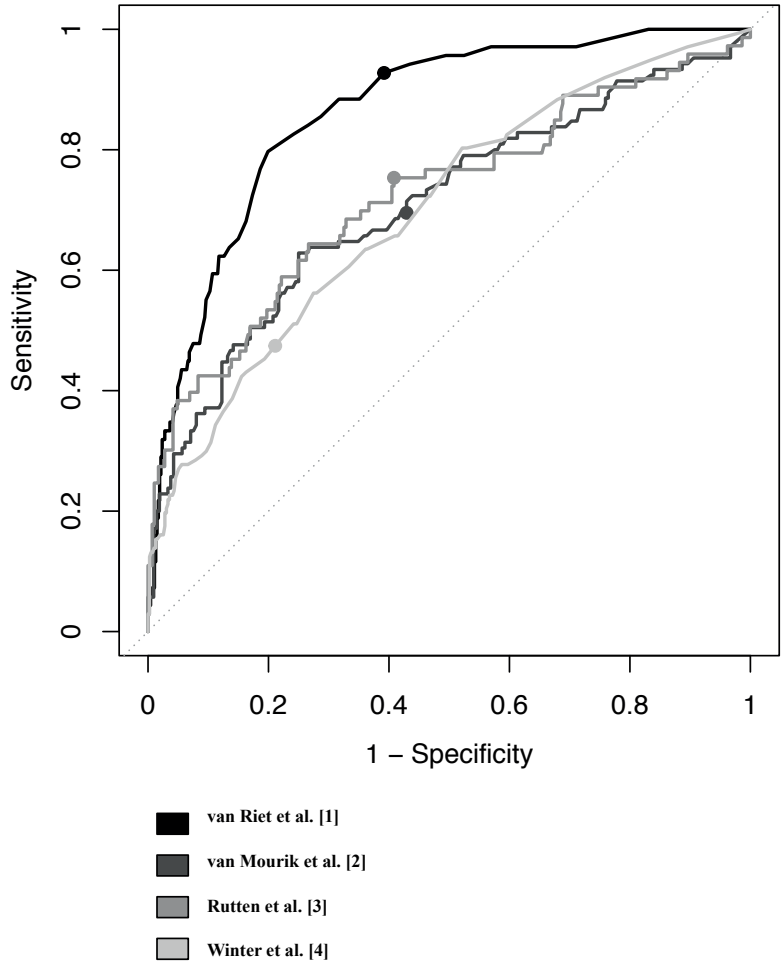
All patients



AF patients


Sensitivity
0
0.2
0.4
0.6
0.8
1
1 – Specificity
van Riet et al. [1]
van Mourik et al. [2]
Rutten et al. [3]
Winter et al. [4]

No AF patients



SUPPLEMENTAL MATERIAL 2. DIAGNOSTIC INDICES IN INDIVIDUAL STUDIES

Cut-point 125 mg/mL

Study		1	2	3	4
Prevalence		48	54	49	31
Sensitivity 125 pg/mL	All	95 (88–98)	76 (68–82)	78 (68–86)	53 (45–60)
	AF	96 (80–99)	96 (82–99)	92 (65–99)	80 (60–91)
	no AF	93 (84–97)	71 (61–78)	75 (64–84)	48 (40–56)
Specificity 125 pg/mL	All	59 (54–63)	55 (49–62)	56 (51–62)	76 (71–80)
	AF	25 (13–43)	35 (20–54)	31 (19–48)	51 (33–68)
	no AF	61 (56–65)	58 (51–64)	59 (54–65)	77 (73–81)
PPV 125 pg/mL	All	30 (25–36)	49 (42–56)	32 (26–38)	45 (38–53)
	AF	53 (39–67)	61 (46–74)	31 (19–48)	59 (42–74)
	no AF	26 (21–32)	46 (38–54)	32 (25–39)	43 (35–51)
NPV 125 pg/mL	All	98 (96–99)	80 (73–85)	91 (86–94)	81 (77–84)
	AF	88 (53–98)	90 (59–98)	92 (65–99)	74 (51–88)
	no AF	98 (96–99)	79 (72–85)	90 (85–94)	81 (77–85)

Numbers indicate percentages (95% CI); PPV = positive predictive value; NPV = negative predictive value

Cut-point 400 mg/mL

Study		1	2	3	4
Prevalence		48	54	49	31
Sensitivity 400 pg/mL	All	48 (38–58)	38 (30–46)	46 (35–56)	23 (17–30)
	AF	70 (49–84)	73 (55–86)	92 (65–99)	70 (50–84)
	no AF	42 (31–53)	29 (21–38)	38 (28–50)	14 (10–21)
Specificity 400 pg/mL	All	93 (91–95)	93 (89–96)	89 (85–92)	97 (95–98)
	AF	65 (46–81)	70 (51–84)	46 (30–62)	69 (50–83)
	no AF	95 (92–96)	96 (93–98)	94 (91–96)	99 (97–99)
PPV 400 pg/mL	All	57 (46–67)	76 (64–85)	52 (41–63)	73 (59–83)
	AF	64 (45–80)	72 (53–85)	37 (22–54)	66 (47–82)
	no AF	53 (40–66)	80 (65–90)	62 (48–75)	79 (60–90)
NPV 400 pg/mL	All	91 (88–93)	73 (68–78)	86 (82–90)	77 (73–80)
	AF	71 (51–85)	71 (52–85)	94 (73–99)	72 (53–86)
	no AF	92 (89–94)	73 (67–78)	86 (82–89)	77 (73–80)

Numbers indicate percentages (95% CI); PPV = positive predictive value; NPV = negative predictive value

SUPPLEMENTAL MATERIAL REFERENCES

1. van Riet EE, Hoes AW, Limburg A, van der Hoeven H, Landman MA, Rutten FH. Strategy to recognize and initiate treatment of chronic heart failure in primary care (STRETCH): a cluster randomized trial. *BMC Cardiovasc Disord* 2014;**14**:1.
2. van Mourik Y, Moons KG, Bertens LC, Reitsma JB, Hoes AW, Rutten FH. Triage of frail elderly with reduced exercise tolerance in primary care (TREE). a clustered randomized diagnostic study. *BMC Public Health* 2012;**12**:385.
3. Rutten FH, Moons KGM, Cramer M-JM, Grobbee DE, Zuithoff NPA, Lammers J-WJ, Hoes AW. Recognising heart failure in elderly patients with stable chronic obstructive pulmonary disease in primary care: cross sectional diagnostic study. *BMJ* 2005;**331**:1379–0.
4. Winter LJMB-D, Rutten FH, Cramer MJM, Landman MJ, Liem AH, Rutten GEHM, Hoes AW. High prevalence of previously unknown heart failure and left ventricular dysfunction in patients with type 2 diabetes. *Diabetologia* 2012;**55**:2154–2162.

6

The effects of misclassification in routine health care databases on the accuracy of prognostic prediction models: A case study of the CHA2DS2-VASc score in atrial fibrillation

S. van Doorn, T.B. Brakenhoff, K.G.M. Moons, F.H. Rutten, A.W. Hoes, R.H.H. Groenwold, G.J. Geersing

Submitted

ABSTRACT

Background: Research on prognostic prediction models frequently uses data from routine healthcare. The disadvantage is the risk of misclassification of predictors which may strongly affect the associations studied. There is no doubt that such misclassification could lead to the derivation of suboptimal prediction models. The extent to which such misclassification affects the validation of existing prediction models is currently unclear.

AIM: A case study to quantify the amount of misclassification in routine care data and its effect on the validation of an existing risk prediction model; the CHA2DS2-VASc clinical prediction model for predicting mortality in patients with atrial fibrillation (AF).

Methods: In a prospective cohort in general practice in the Netherlands, we used computerised retrieved data from the electronic medical records of patients with AF as index predictors. Additionally, manually collected data after scrutinising all complete medical files were used as reference predictors. Comparing the index with the reference predictors, we assessed misclassification in individual predictors by calculating Cohen's kappas and univariable hazard ratios (HR) for mortality. Predictive performance was quantified by the c-statistic and by determining calibration of multivariable models.

Results: In total 2,363 AF patients were included. After a median follow-up of 2.7 (IQR 2.3 – 3.0) years, 368 patients died (incidence rate 6.2 deaths per 100 person-years). Misclassification in individual predictors ranged from substantial (Cohen's kappa 0.56 for prior history of heart failure) to minor (kappa 0.90 for a history of type 2 diabetes). The HR from univariable Cox analysis was 2.1 for heart failure using index predictors, and 1.7 with reference predictors. For type 2 diabetes HRs were 0.81 and 0.83, respectively. The overall model performance was not affected by the misclassification between index and reference predictors of the CHA2DS2-VASc model with a c-statistic of 0.684 for index and 0.681 for reference predictors, and similar calibration.

Conclusion: Even in the presence of substantial predictor misclassification in routine healthcare data, the overall performance of a prediction model was not negatively affected. Although our study should be repeated for other often applied prediction models, our findings suggest that routinely available health care data are a useful source when validating prognostic models, despite misclassification in some of the variables.

INTRODUCTION

Prognostic prediction models aim to estimate the probability that a certain outcome may develop in the future, and in many medical fields they are essential in assisting clinical decision making. Studies on prediction models include development, validation, updating and implementation, and frequently rely on large datasets from routine healthcare.[1] Derived from, for instance, electronic health records or administrative databases, these data offer great potential for clinical research. After a prediction model is developed and its potential usefulness is recognised it is typically validated, possibly using routine health care data, in different health care settings and various countries to justify its application.

Yet, while the validity of routine healthcare data[2] and implications of potential misclassification on studied associations[3–7] are well-addressed *in general*, misclassification in predictors in the context of prognostic research *specifically,* has received little attention. Even though the RECORD statement[2] suggests to assess the accuracy of categorical routine health care variables by comparing them to a reference standard using diagnostic test accuracy measures (i.e. sensitivity, specificity, positive and negative predictive values) or kappa coefficients, it is still unknown whether this approach sufficiently captures the potential bias and/or imprecise inferences that may arise when validating existing prediction models.

Using the well-known CHA2DS2-VASc model as a case study, we aimed to further explore the influence of predictor misclassification on the validation of a prediction model when using routine health care or registry data.

First, we quantified the amount of misclassification present in routine care registry data of a representative sample of patients with atrial fibrillation (AF) in general practice. Second, we assessed the influence of predictor misclassification on the accuracy of the CHA2DS2-VASc model to predict mortality when validated on such data.

METHODS

Clinical setting and the CHA2DS2-VASc prediction rule

For this study, we assessed the prediction of all-cause mortality in patients with atrial fibrillation using the previously developed and validated CHA2DS2-VASc prediction rule. AF is the most common cardiac arrhythmia, with a prevalence of 1–2% in the general population.[8] It is a major risk factor for ischaemic stroke, hence the prediction (and subsequent reduction) of stroke risk is a mainstay in the treatment of AF.[9] Practice guidelines[10–12] recommend the use of a clinical

prediction rule, of which the CHA2DS2-VASc rule is now most commonly recommended and used. This rule was developed in 2010 by Lip et al.[13], as an update to the earlier CHADS2 score,[14] and originally intended to predict either an ischaemic stroke, peripheral embolism, or pulmonary embolism by assigning AF patients points for congestive heart failure (1 point), hypertension (1 point), age above 75 years (2 points), diabetes (1 point) and prior stroke (2 points), age above 65 (1 point), vascular disease (1 point) and female sex (1 point). See Table 1.

TABLE 1. The original CHA2DS2-VASc score[13]

Predictor	Score
Congestive heart failure/LV dysfunction	1
Hypertension	1
Age ≥ 75 years	2
Diabetes mellitus	1
Stroke/TIA/TE	2
Vascular disease (prior myocardial infarction, peripheral artery disease, or aortic plaque)	1
Age 65–74 years	1
Sex category (i.e. female sex)	1

TE = thromboembolism

Index predictors: routine care ICPC codes

We used data from the CAFe study, a large prospective cohort study of patients with AF in general practice in the Netherlands aimed to validate the accuracy of the CHA2DS2-VASc prediction model and to quantify the effect of an automated treatment decision support tool (Trial registration number NTR3741) in a cluster randomised trial. From February 2013 until September 2014, 38 general practices were enrolled. All patients with electrocardiographically confirmed AF were included in the CAFe cohort. Follow-up lasted a minimum of two years. Every three months the electronic patient file of these AF patients was captured into a designated research database, containing diagnosis codes, free text records and test results. In the Netherlands, general practitioners (GPs) are encouraged to record 'diagnosis codes' according to the International Classification of Primary Care (ICPC)[15] during routine care consultations. In the general practices, personal details are registered through linkage to administrative data from the municipal authorities, of which age and sex are captured into the research database. For the remaining predictor values in CHA2DS2-VASc, the corresponding ICPC codes were automatically retrieved and considered as the index predictors. For an overview of the ICPC codes used, see Table 2.

TABLE 2. Automatically extracted ICPC codes for the index predictors in the CHA2DS2-VASc model and the definition of the reference predictors used for manually scrutinising the electronic patient file.

Predictor	ICPC code(s) for index predictors	Definition for reference predictors
Congestive heart failure	K77 heart failure	Signs and symptoms suggestive of heart failure, with structural or functional abnormalities on echocardiography, either with preserved or reduced ejection fraction
Hypertension	K86 Hypertension without organ damage K87 Hypertension with organ damage/secondary hypertension	Repeated systolic blood pressure measurement of 140 mmHg or higher
Age	Age in years	Age in years
Diabetes	T90 Type I and Type II Diabetes	Repeated fasting blood glucose measurement of ≥ 7.0 mmol/L (126 mg/dL) or a non-fasting glucose measurement of ≥ 11.1 mmol/L (200 mg/dL)
Stroke/TIA	K89 TIA K90 Cerebrovascular accident (stroke)	Focal neurological deficit of sudden onset lasting > 24 hours or < 24 hours, respectively
Vascular disease	K74 Angina pectoris K75 Acute myocardial infarction K76 Other chronic ischaemic heart disease K91 Atherosclerosis K92 Other peripheral arterial disease K03 Other pain suspected to originate from the cardiovascular tract	Coronary heart disease: prior myocardial infarction (both ST-elevated myocardial infarction or non-ST-elevated myocardial infarction), angina pectoris or prior percutaneous coronary intervention (PCI) or coronary artery bypass graft surgery (CABG) Peripheral artery disease: symptoms of intermittent claudication with ankle-brachial index ≤ 0.9 or prior surgery or percutaneous intervention on the abdominal or thoracic aorta or lower extremity vessels Previous thrombo-embolism
Sex category	Female sex	Female sex

Reference predictors: manually verified predictors

Except for the predictors 'Age' and 'Sex category', which were obtained from the municipal authorities, the correctness of the routinely recorded ICPC codes corresponding to the remaining CHA2DS2-VASc predictors was manually checked using all available information from the electronic patient file including diagnostic test results, out-of-hours office reports and specialists' letters. Hereto, each patient file was thoroughly scrutinised and the value of each ICPC code corresponding to the predictors in the CHA2DS2-VASc was recorded and used as the reference predictors. The definitions of the reference predictors are shown in Table 2.

For each patient two values for the CHA2DS2-VASc predictors were included in the dataset; one based on the ICPC codes recordings (index) and one based on the manual check of these ICPC codes by scrutinising the complete patient file (reference).

Outcome

Our aim was to study potential misclassification in the prediction variables, not in the outcome. Although the CHA2DS2-VASc was developed to predict either ischaemic stroke, peripheral embolism, or pulmonary embolism, we therefore used all-cause mortality as the outcome for two reasons. First, stroke may be difficult to diagnose, especially stroke as the cause of (unexpected) death. For this case study we therefore used the outcome all-cause mortality as this is not subject to any interpretation. Second, such mortality data may often be captured in municipal authorities, as was the case for the general practices in our study, further avoiding misclassification in the outcome. We manually checked vital status using the electronic patient file. Follow-up was a minimum of two years.

Ethics

All data extracted from the electronic patient files was de-identified by a 'trusted third party'. This study complied with the Data Protection Law in the Netherlands. The medical ethics committee of the University Medical Centre Utrecht, the Netherlands, judged the CAFe study protocol as exempt from review as it was conducted outside the criteria for the Medical Research Involving Human Subjects Act (WMO). Participating general practitioners provided written informed consent.

Data analyses

The following analyses were performed to assess misclassification in the predictors based on routinely recorded ICPC codes (index) and determine the consequences of such misclassification on the prediction of all-cause mortality:

1. We compared the index predictor values with the reference predictor values using Cohen's kappa[16] and calculated sensitivity, specificity, and positive and negative predictive values of the dichotomous index predictors with respect to the reference predictors.
2. For each patient we calculated the CHA2DS2-VASc score using either the index predictors or the reference predictors. We tabulated the two distributions of these CHA2DS2-VASc scores and the discordance. Next, for each CHA2DS2-VASc score based on index predictors and reference predictors we calculated the mortality rate per 100 person-years.
3. We performed univariable Cox proportional hazard analyses to compare the hazard ratios (HRs) of the individual CHA2DS2-VASc predictors from either the index or reference data.
4. We then used Cox proportional hazard models to analyse the predictive performance of both scores. Additionally, we assessed whether the influence of misclassification differed for models *with* or *without* 'Age' and 'Sex category' as

these were assumed not to be misclassified. For both models, the c-index was calculated. For both models, the c-index was calculated.

5. We finally re-fitted the CHA2DS2-VASc rule using the reference predictors, yielding the optimal baseline hazard and hazard ratios. Predictive performance of this reference model was assessed using the c-statistic and the model's calibration for predicting mortality within two years. The re-fitted model was subsequently validated using the index predictors of the patients as input values and again the discrimination and calibration were estimated and compared to the reference model; the difference then occurring can *only* be caused by misclassification.

All analyses were performed in *R 3.32* with the packages *survival 2.40–1* and *rms 5.1–0.*

RESULTS

A total of 2,363 patients with atrial fibrillation were included in the cohort. The median age was 77 (IQR 68 – 84) years, and 52.3% were male. During a follow-up of 5,901 person years (median 2.7, IQR 2.3 – 3.0) years, in total 368 patients died (crude incidence rate 6.2/100 person-years), mostly from non-cardiac causes (74%).

Misclassification in individual predictor values

There was substantial variation in the amount of misclassification between the index predictors, see Table 3. For instance, the prevalence of (a history of) heart failure according to the ICPC codes was 28.1%, whereas by manually checking all available information in the electronic patient file the prevalence was 18.3% (Cohen's kappa 56.1). The prevalence of other index and reference predictors were more comparable; e.g. for hypertension 60.8% and 59.9% (kappa 70.9), respectively, and for diabetes 24.3% and 22.5% (kappa 89.7), respectively. As a result, sensitivity was lowest for heart failure (55%) and highest for diabetes (89%). Specificity ranged from 83% (hypertension) to 99% (diabetes).

TABLE 3. Prevalence of individual ICPC codes (index predictors) and manually verified diagnoses (reference predictors) and measures of misclassification

	ICPC codes (index predictors	Manually verified diagnoses (reference predictors)	kappa	Sensitivity	Specificity	PPV	NPV
Heart failure	28.1	18.3	56.1	54.5	95.7	83.3	84.3
Hypertension	60.8	59.9	70.9	87.8	83.3	89.1	81.6
Diabetes	24.3	22.5	89.7	88.6	98.8	95.8	96.4
Stroke	18.7	16.4	75.5	74.8	97.1	85.5	94.4
Vascular disease	34.6	26	60.4	63.1	93.7	84.2	82.7

ICPC = International Classification of Primary Care; PPV = positive predictive value; NPV = negative predictive value

CHA2DS2-VASc scores and observed mortality

With respect to the reference predictors, the index predictors assigned patients the correct CHA2DS2-VASc scores in between 40.7% (for score 7) and 85.0% (for score 0); see Figure 1. The median CHA2DS2-VASc score using index data was 4.0 (IQR 2 – 5), for the reference data this was 3.0 (IQR 2 – 5).

Table 4 shows the incidence rates of all-cause mortality for each CHA2DS2-VASc score calculated with index and reference predictors.

TABLE 4. Incidence rate of all-cause mortality for each CHA2DS2-VASc score as calculated with ICPC codes (index predictors) or manually verified diagnoses (reference predictors).

	ICPC codes (index predictors)				**Manually verified diagnoses (reference predictors)**			
Score	*No. of patients (%)*	*No. of events*	*py*	*IR*	*No. of patients (%)*	*No. of events*	*py*	*IR*
1	124 (5.3)	2	338	0.6	125 (5.3)	2	346	0.6
2	194 (8.2)	2	541	0.4	203 (8.6)	4	567	0.7
3	307 (13.0)	29	892	3.3	344 (14.6)	37	994	3.7
4	356 (15.1)	48	1041	4.6	417 (17.7)	54	1208	4.5
5	404 (17.2)	67	1186	5.6	431 (18.3)	83	1274	6.5
6	292 (12.4)	86	887	9.7	254 (10.8)	89	795	11.2
7	187 (7.9)	70	590	11.9	139 (5.9)	60	441	13.6
8	82 (3.5)	37	262	14.1	54 (2.3)	27	187	14.4
9	33 (1.4)	26	127	20.5	16 (0.7)	10	60	16.7
1	8 (0.3)	1	23	4.3	4 (0.2)	2	14	14.3

py = person-years/100; IR = incidence rate (number of events/100 person-years)

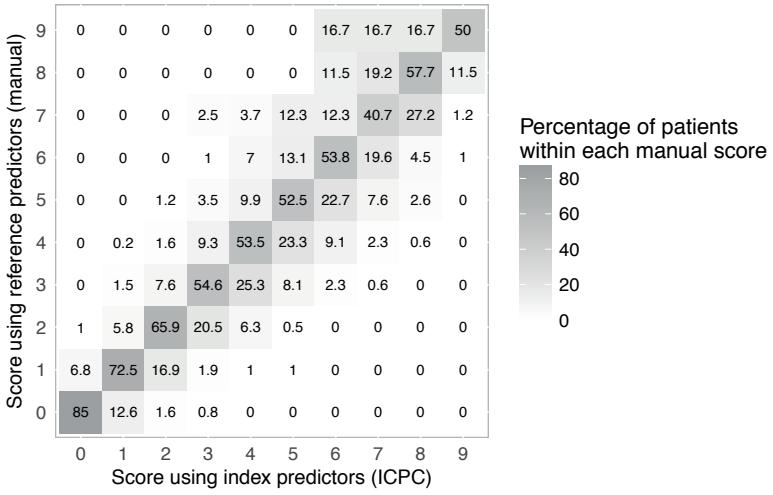


FIGURE 1. CHA2DS2-VASc scores assigned by index predictors and reference predictors
The concordance of CHA2DS2-VASc scores as calculated using the index predictors (x-axis) and as calculated using the reference predictors (y-axis). Numbers are percentages.

Univariable analyses

The hazard ratios (HRs) of the individual CHA2DS2-VASc predictors were higher for the index predictors compared to the reference predictors, except for 'Vascular disease' (Table 5).

The CHA2DS2-VASc score as a continuous covariate resulted in the same hazard ratio of 1.4 per 1 point increase in the score for both type of predictors (index or reference). The c-statistics of these models were 0.682 (95% CI 0.653 – 0.712) and 0.685 (95% CI 0.655 – 0.715), respectively.

Multivariable analyses

In a multivariable model *without* age and sex the HRs showed only minor differences except, again, for heart failure. These differences attenuated when including age and sex in the model, both obtained from the municipal authorities and therefore devoid of misclassification.

The c-statistic of the model developed using the reference predictors was 0.715 (95% CI 0.571 – 0.629). From this model, we used the baseline hazard and coefficients for validation with the index predictors as input values, resulting in a c-statistic of 0.600 (95% CI 0.571 – 0.629). Even though there was a slight underestimation of the probability of survival across most risk deciles when using routine health care data, differences in calibration were minimal (Figure 2).

TABLE 5. Hazard ratios from univariable and multivariable Cox proportional hazard analyses using the ICPC codes (index predictors) or manually verified diagnoses (reference predictors).

5a. Hazard ratios (95% CI) from univariable analyses

	ICPC codes (index predictors)	Manually verified diagnoses (reference predictors)
Heart failure	3.1 (2.9 – 3.3)	2.5 (2.2 – 2.7)
Hypertension	1.1 (0.9 – 1.4)	1.1 (0.9 – 1.3)
Diabetes	1.6 (1.3 – 1.8)	1.6 (1.3 – 1.8)
Stroke	2.0 (1.8 – 2.3)	2.0 (1.8 – 2.2)
Vascular disease	1.7 (1.6 – 1.9)	1.9 (1.7 – 2.1)
Score	1.4 (1.4 – 1.5)	1.4 (1.4 – 1.5)

5b. Hazard ratios (95% CI) from multivariable analyses, without (left-hand side) and with (right-hand side) 'Age' included in the model

	ICPC codes (index predictors)	Manually verified diagnoses (reference predictors)	ICPC codes (index predictors)	Manually verified diagnoses (reference predictors)
Heart failure	2.7 (2.5 – 3.0)	2.2 (1.9 – 2.4)	2.1 (1.9 – 2.3)	1.7 (1.5 – 1.9)
Hypertension	1.0 (0.7 – 1.2)	1.0 (0.8 – 1.2)	0.8 (0.6 – 1.0)	0.8 (0.6 – 1.1)
Diabetes	1.3 (1.1 – 1.6)	1.3 (1.0 – 1.5)	1.4 (1.2 – 1.6)	1.3 (1.1 – 1.5)
Stroke	1.8 (1.6 – 2.0)	1.9 (1.6 – 2.1)	1.6 (1.4 – 1.9)	1.6 (1.4 – 1.9)
Vascular disease	1.3 (1.1 – 1.5)	1.5 (1.3 – 1.8)	1.1 (0.9 – 1.4)	1.4 (1.2 – 1.6)
Age 65 – 75 years			2.5 (1.8 – 3.1)	2.5 (1.9 – 3.1)
Age ≥ 75 years			6.3 (5.7 – 6.9)	6.8 (6.3 – 7.4)
Sex category			1.0 (0.8 – 1.2)	1.1 (0.9 – 1.3)

Even though there was an underestimation of the probability of survival across all risk deciles when using routine health care data, especially in the lower range of predicted probabilities, differences in calibration were minimal (Figure 2).

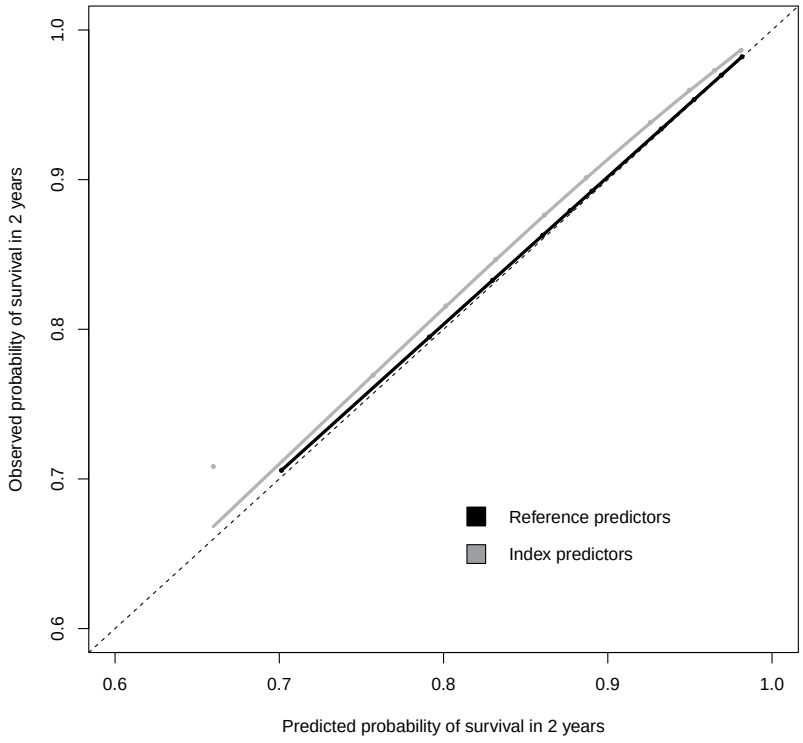


FIGURE 2. Calibration plot of the CHA2DS2-VASc model.
Calibration plot showing deciles of observed and predicted probabilities of survival from the CHA2DS2-VASc model developed using the reference predictors, and validated using the baseline hazard and coefficients for validation with the index predictors as input values.

DISCUSSION

We illustrated the impact of potential misclassification in routine healthcare data when such data was used as predictors in a prognostic prediction model. In our validation of the CHA2DS2-VASc rule in patients with atrial fibrillation, we found substantial misclassification in the predictor values from routinely collected general practice diagnosis codes, but this did not affect the accuracy of the model to predict mortality.

In recent years, the availability of data routinely collected during healthcare delivery has grown substantially,[17] whereas in the past epidemiologic research often was dependent on dedicated prospective cohorts.[18] With the availability of faster computers and software programs, everyday health care data, possibly linked to other data sources has a great potential for large-scale observational clinical studies. Indeed, in the field of AF for instance studies evaluating popula-

tions with over 100,000 AF patients are becoming the new standard, rather than an exception.[19–21] Importantly though, these studies mostly rely on diagnosis disease codes (e.g. ICD-10 codes, READ codes or ICPC coding) as generated during daily clinical practice. Following studies on investigating the completeness of morbidity coding[22] or the methods and reporting of validity assessment[23], the quality of these data has been questioned. While these studies certainly contribute to knowledge on the validity of routine healthcare *data itself,* it does not provide full insight in the validity of applying such data in prediction model studies. This is important, because the number of prediction models used in everyday practice is rapidly increasing.[24–26]

To the best of our knowledge, our study is the first to quantify the influence of predictor misclassification in routine healthcare data on the results of a study validating a clinical prediction model.

For full appreciation of our findings, several remarks should be made. First, several processes leading to misclassification in data from routine healthcare can be hypothesised. At the most basic level, simple coding mistakes such as typing errors or choosing the wrong diagnosis code may lead to the inadvertent presence or absence of a diagnosis code. Furthermore, if an initially suspected disease (e.g. heart failure, or coronary heart disease) is not confirmed after future diagnostic testing, the diagnosis code needs active removal from the electronic patient file or it will lead to 'false positives'. Practitioners conversely may also omit diagnosis codes for certain diseases frequently occurring and managed concomitantly. For instance, recording 'hypertension' and 'coronary heart disease' (both included in the CHA2DS2-VASc model) together as 'cardiovascular disease' may cause 'false negatives' in the index predictors. After manually scrutinising the patient file we were able to incorporate the correct diagnoses into the reference predictors.

Second, a further cause for misclassification may be suboptimal diagnostic criteria for a certain disease. We found substantial variation in the validity of data from routine healthcare where, for instance, 'a history of heart failure' showed notable misclassification. It can be difficult to diagnose heart failure, especially in absence of echocardiography as is often the case in general practice. Indeed it has been shown that heart failure is often over-diagnosed in general practice, similarly as in our study.[27] Diabetes, on the contrary, is predominantly diagnosed in general practice based on well-defined diagnostic criteria, and showed very limited misclassification. When using routine care data in epidemiological research, potential difficulties in diagnosis of diseases and thus variation between data sources in the variables under study (e.g. electronic patient records or administrative databases) should be considered.[28]

Third, as a result of misclassification in predictors the total CHA2DS2-VASc score for a given patient differed substantially between data sources. This may have large implications if a cut-point is applied as is the case with the CHA2DS2-VASc score.[10] Well-defined specific treatment recommendations apply for those with a score of 0, 1 or ≥ 2 and miscalculation by only one point will impact the proportion of patients eligible for anticoagulant treatment. As an illustration, of the patients in whom such treatment was indicated (CHA2DS2-VASc score ≥ 2) based on index predictors, nearly 20% had a score 0–1 based on reference predictors and thus no strict indication for treatment. Likewise, validation studies of prediction rules commonly report the observed risk per score, and, for example, for a CHA2DS2-VASc score of 1 based on index predictors, this risk was nearly twice as high as the same score based on the reference predictors.

Fourth, while misclassification in individual predictors was substantial, the discrimination and calibration of full models containing all predictors of CHA2DS2-VASc was comparable between routinely collected index data and the reference data. The misclassification in the former, thus, seem to 'average out' in multivariable analyses. Our results suggest that while a data source shows low performance on the 'traditional' measures of accuracy (kappa, sensitivity/specificity and predictive values), one may still observe valid estimates when validating a multivariable prediction model.

Lastly, we observed distinct differences in a model with and without 'Age' and 'Sex' as a covariate. Presumably, such variables may be obtained from the municipality with (near) perfect classification. Studies relying on routine care data, either prognostic studies or otherwise, may benefit from including such types of variables to attenuate the influence of misclassification. For models not including such predictors, overall model performance may be more prone to misclassification.

Strengths and limitations

Strengths of our study include the large sample size of manually collected data that served as the reference standard against which routine care data was compared. We verified the disease status, predictor values and outcomes in over 2,000 health records. Manually scrutinising electronic patients files is a resource intensive process, and we believe this amount approaches what may be considered the maximum feasible amount. Furthermore, we could collect clinical data from general practice, but also could include specialists' letters with diagnoses and test results from secondary care. Consequently, we were able to study an often-used clinical prediction rule without any missing data.

A limitation of our study is that, irrespective of clear definitions for manually checking the predictors, some information (e.g. description of signs and symptoms

in free text fields) leave room for different interpretation. The final judgment was made by the researcher, based essentially on the same data that was used by the GP to record the initial ICPC diagnosis code. We did not subject patients to any new clinical assessment. As such, some misclassification might also have occurred in our reference data. Furthermore, we only evaluated a single prediction model. How our results apply to other prognostic prediction models should be the focus of future research. In addition, our study used all-cause mortality as the outcome while the CHA2DS2-VASc rule was specifically designed to predict stroke risk. While this avoided misclassification in the outcome, the influence of misclassification on the performance of its intended purpose requires further research.

Future considerations

Our results provide evidence that misclassification in routine healthcare data can be substantial and that several aspects (e.g. the risk of the outcome with a certain score) of the validation of a clinical prediction rule may be influenced, while other aspects (such as discrimination and calibration of the prognostic score) may not. Future studies should focus on the influence of different misclassification patterns on validation performance, e.g. in a simulation study. In addition, when data on true predictor status is available, this can be used to correct for misclassification in routine health care data.[5] Insight is needed in the amount of reference data necessary to ensure reliable prediction model performance. This can advise researchers on the efforts required to obtain any reference data (e.g. the proportion of patient files that needs manual checking). Ultimately, future research on these topics can further inform applied researchers on when routine health care data can reliably be used to evaluate prediction models.

CONCLUSION

In this case study of CHA2DS2-VASc, we observed that even in the presence of substantial predictor misclassification in routine healthcare data, the overall performance of a prediction model was not negatively affected. Although our study should be repeated for other often applied prediction models, our findings suggest that routinely available health care data are a useful source when validating prognostic models, despite misclassification in some of the variables.

REFERENCES

1. Riley RD, Ensor J, Snell KIE, Debray TPA, Altman DG, Moons KGM, Collins GS. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ* 2016; **353**: i3140.
2. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, Elm von E, Langan SM, Committee RW. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Med* 2015; **12**: e1001885.
3. Buonaccorsi JP. Measurement Error. CRC Press; 2010.
4. Wayne AF. Measurement error models. New Yorkjohn Wiley; 1987.
5. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. Measurement Error in Nonlinear Models. CRC Press; 2006.
6. Stefanski LA. Measurement error models. *J Am Stat Assoc* 2000.
7. Guolo A. Robust techniques for measurement error correction: a review. *Stat Methods Med Res* 2008; **17**: 555–80.
8. Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, Singer DE. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* 2001; **285**: 2370–5.
9. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991; **22**: 983–8.
10. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, Castella M, Diener H-C, Heidbuchel H, Hendriks J, Hindricks G, Manolis AS, Oldgren J, Popescu BA, Schotten U, Van Putte B, Vardas P, Agewall S, Camm J, Barón-Esquivias G, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J* 2016; **37**: 2893–962.
11. January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, Conti JB, Ellinor PT, Ezekowitz MD, Field ME, Murray KT, Sacco RL, Stevenson WG, Tchou PJ, Tracy CM, Yancy CW, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol*. 2014. pp. e1–76.
12. NICE. Atrial fibrillation: the management of atrial fibrillation. (Clinical guideline 180). NICE; 2014.
13. Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. 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.
14. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of Clinical Classification Schemes for Predicting Stroke: Results From the National Registry of Atrial Fibrillation. *JAMA* 2001; **285**: 2864–70.
15. Lamberts H, Wood M, World Organization of National Colleges, Academies, and Academic Associations of General Practitioners/Family Physicians, Party IW. ICPC, international classification of primary care. Oxford University Press, USA; 1987.
16. Cohen J. A coefficient of agreement for nominal scales. *Educ psychol meas* 1960.

17. de Lusignan, S., Teasdale, S., Little, D., Zapp, J., Zuckerman, A., Bates, D. W., & Steele, A. (2004). Comprehensive computerised primary care records are an essential component of any national health information strategy: report from an international consensus conference. (Vol. 12, pp. 255–264). Presented at the Informatics in primary care.
18. DAWBER TR, Kannel WB. An epidemiologic study of heart disease: the Framingham study. *Nutr Rev* 1958; **16**: 1–4.
19. Aspberg S, Chang Y, Atterman A, Bottai M, Go AS, Singer DE. Comparison of the ATRIA, CHADS2, and CHA2DS2-VASc stroke risk scores in predicting ischaemic stroke in a large Swedish cohort of patients with atrial fibrillation. *Eur Heart J* 2016; **37**: 3203–10.
20. Chao T-F, Liu C-J, Tuan T-C, Chen S-J, Wang K-L, Lin Y-J, Chang S-L, Lo L-W, Hu Y-F, Chen T-J, Chiang C-E, Chen S-A. Comparisons of CHADS2 and CHA2DS2-VASc scores for stroke risk stratification in atrial fibrillation: Which scoring system should be used for Asians? *Heart Rhythm* 2016; **13**: 46–53.
21. Nielsen PB, Larsen TB, Skjøth F, Overvad TF, Lip GYH. Stroke and thromboembolic event rates in atrial fibrillation according to different guideline treatment thresholds: A nationwide cohort study. *Sci Rep* 2016; **6**: 27410.
22. Jordan K, Porcheret M, Croft P. Quality of morbidity coding in general practice computerized medical records: a systematic review. *Fam Pract* 2004; **21**: 396–412.
23. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010; **69**: 4–14.
24. Bouwmeester W, Zuithoff NPA, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, Altman DG, Moons KGM. Reporting and methods in clinical prediction research: a systematic review. *PLoS Med* 2012; **9**: 1–12.
25. Steyerberg EW, Moons KGM, van der Windt DA, Hayden JA, Perel P, Schroter S, Riley RD, Hemingway H, Altman DG, for the PROGRESS Group. Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. *PLoS Med* 2013; **10**: e1001381.
26. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med* 2006; **144**: 201–9.
27. Valk MJ, Mosterd A, Broekhuizen BD, Zuithoff NP, Landman MA, Hoes AW, Rutten FH. Overdiagnosis of heart failure in primary care: a cross-sectional study. *br j gen pract* 2016; **66**: e587–92.
28. Siregar S, Pouw ME, Moons KGM, Versteegh MIM, Bots ML, van der Graaf Y, Kalkman CJ, van Herwerden LA, Groenwold RHH. The Dutch hospital standardised mortality ratio (HSMR) method and cardiac surgery: benchmarking in a national cohort using hospital administration data versus a clinical database. *Heart* 2014; **100**: 702–10.

7

General discussion

STROKE RISK PREDICTION OF PATIENTS WITH ATRIAL FIBRILLATION IN GENERAL PRACTICE

Atrial fibrillation (AF) is the most common cardiac arrhythmia. Its prevalence in the general population is 1–2% and rises steeply with age up to over 15% in patients 85 years and older.[1,2] One in four middle-aged adults will eventually develop AF.[3]

Patients with AF are at high risk of a kaleidoscope of adverse outcomes.[4] Most importantly, the risk of ischaemic stroke is increased on average five-fold,[5] to well over 5% each year in patients aged 75 years and older depending on risk factors.[6] Clearly, stroke prevention therefore is a key element in the treatment of AF. Also, common concurrent diseases such as heart failure need early detection and adequate treatment.

Patients with AF are frequently treated in general practice, where the population is often characterised by a high age, frequent multimorbidity,[7,8] and polypharmacy.[9,10] General practitioners need to manage both AF and any co-occurring (chronic) conditions, each with potential complications and treatment interactions. Atrial fibrillation therefore poses tremendous challenges to physicians, patients and the healthcare system.

This thesis aimed to address some of the major challenges, most notable for the primary care setting. The main findings include:

- In a meta-analysis of 19 validation studies, it was shown that there is large unexplained heterogeneity in stroke risks for different scores on the CHA2DS2-VASc prediction model (Chapter 1);
- A survey among general practitioners revealed that uncertainty about how to manage patients who have sinus rhythm after an episode of AF, and about who is responsible for anticoagulation are important reasons for non-adherence to guidelines on stroke prevention (Chapter 2);
- A cluster randomised trial showed that an automated decision support is ineffective in optimising anticoagulant therapy and preventing stroke in general practice (Chapter 3);
- In a prospective cohort study we showed that anticoagulated patients with atrial fibrillation are at high risk of hospitalisation and mortality, predominantly for non-cardiac reasons, (Chapter 4);
- Using a meta-analysis of individual patient data from four screening studies, we were able to assess a high prevalence of previously unknown heart failure in patients with AF and concluded that NTproBNP was of limited value for heart failure screening in AF, (Chapter 5);

– Predictor misclassification in routine healthcare data has limited impact on the validation of existing prognostic prediction models, in this case the CHA2DS2-VASc model (Chapter 6).

RISK PREDICTION IN ATRIAL FIBRILLATION

Anticoagulants – either vitamin K antagonists (VKAs) or direct oral anticoagulants (DOAC) – are very effective in preventing ischaemic stroke in AF, [11–15] and *on average* their benefit of a 66% relative reduction in stroke risk to patients with atrial fibrillation is indisputable. However, treatment decisions for *individual* patients can be less of a clear-cut case. Stroke risk differs among individual patients with AF as does risk of (major) bleeding complications inherently caused by oral anticoagulants. Essentially, both risks determine the individual harm-benefit ratio of anticoagulants, but may also affect the choice for the type (VKA or DOAC) and the dose intensity (normal or reduced) of anticoagulants. Weighing of stroke risk versus bleeding risk therefore is key in the management of atrial fibrillation. While recognising that predicting *bleeding risk* is of equal importance, in this general discussion I will focus on *stroke risk* prediction in patients with atrial fibrillation. Notably, the possible implication of the findings in this thesis on research aiming to improve stroke risk prediction in AF patients will be discussed.

To predict stroke risk in individuals with atrial fibrillation, in 2010 Lip et al. published the CHA2DS2-VASc model as developed in a cohort of AF patients visiting cardiology clinics.[16] In the very same year, before studies on external validation of the model were published, CHA2DS2-VASc was included in the Guidelines for the management of atrial fibrillation of the European Society of Cardiology,[17] first as a complement to, and later as a replacement[18] of the original CHADS2 score developed by Gage et al. in 2001.[19] This swift introduction of CHA2DS2-VASc in the 2010 ESC guideline was followed by other practice guidelines, including those from NICE in the United Kingdom[20], the American Heart Association[21] and primary care organisations.[22,23] This effectively made CHA2DS2-VASc the most used prediction model driving anticoagulant treatment decisions.

The results of the meta-analysis presented in Chapter 1 of this thesis, however, showed that the exact stroke risk predicted by the CHA2DS2-VASc model remains uncertain, resulting in two partly complementary problems. First, patients currently believed to have a *stroke risk low enough* to withhold anticoagulation (based in a CHA2DS2-VASc score 0 or 1) may still have a stroke risk well over 2% per year that merits stroke prevention therapy. Second, patients considered *at high risk of stroke* (i.e. those with a CHA2DS2-VASc score ≥ 2) are recommended anticoagulant

treatment while some have a stroke risk even as low as 1%. These problems likely cause preventable strokes and bleeds respectively, and hence call for improvement in stroke risk prediction and subsequent optimisation of preventive treatment. Below, I will propose two distinct pathways to improve stroke risk prediction:

Improving stroke risk prediction by re-evaluating the *existing* individual (weights of the) CHA2DS2-VASc predictors including an evaluation of possible interaction terms, thus updating the existing rule to new data at hand; Evaluating *additional* novel stroke risk predictors, including clinical parameters, biomarkers, and/or imaging parameters.

Next, I propose the necessary steps future research should take to address these possible pathways of optimising stroke risk prediction in AF patients.

1. Updating the existing CHA2DS2-VASc rule

An important first pathway towards improved prediction of stroke in patients with atrial fibrillation using the CHA2DS2-VASc model includes: i) re-evaluating the individual CHA2DS2-VASc predictors using more (recent) patient data, and ii) evaluation whether addition of interaction terms between these predictors improves the performance of the score.

1.1. Re-evaluating the individual CHA2DS2-VASc predictors

Since the development of the first stroke prediction models[19,24,25] that ultimately led to the development of CHA2DS2-VASc, the disease entities of some of the predictors changed in subsequent decades. As a result, the predictive weights associated with each may also have changed. It is now widely recognised that 'heart failure' for example may be categorised in heart failure with *preserved, mid-range* or *reduced* ejection fraction.[26,27] Prognosis of the latter improved due to changes in treatment.[28,29] As a result, heart failure as a predictor for stroke in patients with AF likely changed as well. Several recent studies could indeed no longer could an association between (well-treated) heart failure and risk of stroke in AF,[30–33] though others still did.[6,34]. Such contradictions are found also for other predictors in CHA2DS2-VASc. Both diabetes and vascular disease were associated with stroke in AF patients recruited in hospitals with CHA2DS2-VASc score 1 in Taiwan,[34] while this was not the case in comparable patients in neighbouring Hong Kong.[35] Future research should confirm or reject such contemporary diseases as true predictors for stroke. In addition, the disease *severity* and *subtypes* should be considered. For example, the duration of diabetes seems to influence the risk of stroke,[36] as may the type of diabetes (type 1 versus type 2). The severity and control of hypertension in anticoagulated AF patients influences stroke risk.[37,38] As another example, the predictor 'vascular disease' included in CHA2DS2-VASc

consists of a heterogeneous group of conditions. Peripheral artery disease seems to be a stronger predictor than coronary artery disease[39] and the definition of the latter varies across validation studies.[40–42] Re-evaluating and updating the definition of these predictors in future research may improve stroke predictions of CHA2DS2-VASc.

As a consequence, the individual weights of the predictors in CHA2DS2-VASc are in need for a re-evaluation. Currently all predictors in the CHA2DS2-VASc model are assigned 1 or 2 points to the total score. Given the aforementioned examples of changing disease definitions over time as well as the impact of disease severity of the different predictors, this very likely represents a simplification of the clinical reality of stroke risk in AF patients. Indeed, several studies have shown substantially higher hazard ratios (HRs) for predicting stroke in AF of diabetes and high age compared to the other predictors,[6,34] while others found remarkably similar HRs.[32,33]. Re-evaluating and subsequent tailoring the weights of the individual CHA2DS2-VASc predictors may result in a score that better reflects the total risk of stroke in a patient with AF.

Furthermore, the predictor 'Age' needs considerable attention in this respect. Currently the age category is broad, assigning the same points to patients aged 65 and 74 years old. Adjusting the age categories may improve risk prediction as for example in Asians it has been suggested that a relevant increased risk of stroke can be observed already from 50 years of age.[43,44] In addition, inclusion of age in the model as a continuous predictor should be considered as it is well-known that categorisation of a continuous measurement leads to loss of predictive information.[45]

1.2 Evaluation of interactions between CHA2DS2-VASc predictors

The next step is the evaluation whether the addition of interactions between existing predictors improves the predictive accuracy of CHA2DS2-VASc. It is currently recognised that 'Female sex' increases stroke risk mostly in the presence of additional risk factors, notably age. Women with a CHA2DS2-VASc score 1 (for 'Sex class' only) are, contrary to males with score 1, considered to be at low risk of stroke, i.e. lower than men.[46] However, this 'interaction' effect of female sex may differ for each of the individual CHA2DS2-VASc predictors. This is illustrated by data from a large Danish cohort[6]. In male AF patients, diabetes as a sole risk factor is associated with a HR of 4.5 (95% CI 2.0 – 9.9) while this is 1.4 (95% CI 0.2 – 10.0) in females. Conversely, vascular disease in males has a HR of 1.0 (95% CI 0.3 – 3.10) while for females this is 3.7 (95% CI 1.2 – 11.9). Although confidence limits are wide and differences are not statistically significant, a 63 year-old male AF patient with diabetes seems at higher risk of stroke than a 63 year-old female

with diabetes and AF, whereas his individual CHA2DS-VASc score paradoxically is lower than hers (score of 1 versus 2, respectively).

Furthermore, other possible interaction terms may be investigated, notably age. High age in itself is a strong predictor of stroke, but any interplay in stroke risk with other predictors in CHA2DS2-VASc is not accounted for. This hypothesis is supported again by data from Olesen et al.[6] It showed that heart failure or hypertension are stronger predictors of stroke in older patients compared to younger patients, while this was not the case for diabetes.

2. Evaluation of additional stroke risk predictors

The second pathway to improved prediction of stroke is the identification of novel predictors for stroke other than those already included in CHA2DS2-VASc: e.g., clinical parameters, biomarkers, and echocardiographic parameters.

2.1 Clinical parameters

Several clinical parameters not included in CHA2DS2-VASc are associated with stroke risk in AF patients, of which renal failure seems most important.[31,47,48] Renal failure has been included in the ATRIA[49] and also in the QStroke stroke prediction model,[50] though adding renal failure as a predictor to CHA2DS2-VASc showed no incremental value. Friberg et al.[51] found no significant improvement in c-statistic nor in the net reclassification index (NRI, expressing the improvements in risk predictions)[52] of the CHA2DS2-VASc score if assigning 1 or 2 points for end stage renal disease. Similarly, both Banerjee et al.[53] and Roldan et al.[54] added 1 or 2 points to the CHA2DS2-VASc score for different categories of estimated glomerular filtration rate and observed no significant change in c-statistic nor in NRI. Both c-statistic and NRI, however, are criticised for their inability to express the incremental value of a predictor when added to a prediction model.[55–57] After updating CHA2DS2-VASc the value of renal value in stroke risk prediction needs confirmation using appropriate measures, such as decision curve analyses.[58]

Obstructive sleep apnoea syndrome (OSAS) is a well-known risk factor for stroke,[59] and often concurrent in patients with atrial fibrillation.[60] Its value as a predictor of stroke in AF has been evaluated in a retrospective study in AF patients referred for polysomnography. OSAS was detected in 85% of patients with atrial fibrillation.[61] In those with OSAS, risk of stroke was on average 25% compared to 8% in those without (adjusted odd ratio of 3.7, 95% CI 1.2 to 10.6). This relation was also confirmed in a case-control study by Mansukhani et al,[62] showing that OSAS patients with a history of stroke more often suffered from AF than OSAS patients without previous stroke (adjusted OR 8.3, 95% CI 3.1 – 23.4).

On the other hand, based on health insurance data of Taiwanese patients, Chang et al[63] did not find a difference in the prevalence of OSAS between AF patients with and without stroke, and no change in c-statistic nor NRI when OSAS was added as a predictor to CHA2DS2-VASc. Further research across different care settings is needed to provide definite evidence for OSAS as a stroke risk factor in AF.

Also, potential risk differences between people with different ethnic backgrounds should be taken into account when predicting stroke in patients with atrial fibrillation. In patients with AF, Lip et al.[64] summarised consistent rates of stroke in community-based cohort studies in China, Japan, Singapore and Taiwan of ~14%, but a far lower rate in South Korea of 2.8%. Importantly, however, while in Chinese studies warfarin use was extremely low (0.5 – 2.7%), this was not reported for the study in South Korea thus hampering straightforward comparison. Using a large routine healthcare database, though, Shen et al.[65] confirmed stroke risk differences between ethnic groups showing that compared to Caucasian ethnicity, African-American ethnicity was associated with an adjusted HR for stroke risk of 1.6 (95% CI 1.3 – 1.9). In a similar data source, stroke rate in Caucasians was 14.8 per 1000 person-years, compared to 29.3 per 1000 person-years in African Americans.[66] Such findings underline the importance of considering ethnicity also to explain heterogeneity across validation studies of CHA2DS2-VASc[67,68] and a possibility for improving stroke risk prediction. In this respect, Kabra et al.[69] evaluated the inclusion of ethnic background in CHA2DS2-VASc. African-American ethnicity increased stroke risk as compared to Caucasian ethnicity (hazard ratio 1.4, 95% CI 1.3 – 1.4 after adjusting for CHA2DS2-VASc score and anticoagulation, consistent with that found by Shen et al.) and was even found to be a stronger predictor than heart failure, hypertension, diabetes mellitus, or a history of vascular disease. Importantly, however, the c-statistic did not change when assigning 1 additional point for African-American background in CHA2DS2-VASc, calibration improved. Further research is needed to confirm the improvement of CHA2DS2-VASc by incorporating ethnic background as a predictor and, similar to the approaches described above, possible interaction of ethnicity with the existing CHA2DS2-VASc variables need evaluation.

2.2 Biomarkers

The role of biomarkers has been under increasing attention in recent years.[70] Several blood parameters have been identified as predictors for stroke in AF patients, albeit with different underlying principles.

C*ardiac* biomarkers are released under influence of myocyte stress or injury. The B-type natriuretic peptide (BNP or NTproBNP) – well-known for its diagnostic value in suspected heart failure (see also Chapter 6 of this thesis) – was identified

as an independent stroke risk predictor in anticoagulated patients in the trials investigating the effectiveness of direct oral anticoagulants (adjusted HR 2.4 for the highest compared to the lowest quartile group in two studies),[71,72] as was troponin (HR 2.0 for high vs. low in three studies).[71,73,74] Also more novel 'cardiac' markers have been evaluated, e.g. growth differentiation factor 15, galectin-3 and amyloid proteins,[75,76] though their value in predicting stroke in patients with AF has not yet been confirmed.

It is believed that atrial fibrillation is linked to *inflammation*[77,78] and different markers for inflammatory response have been evaluated. C-reactive protein (CRP) is associated with several adverse outcomes in patients with AF, including myocardial infarction[79] and mortality.[80,81] Studies observed a correlation between CRP and stroke risk factors (Spearman's correlation for increasing CHADS2 score 0.15, $p < 0.001$)[81] and thrombotic state (Spearman's correlation for fibrinogen 0.42 ($p < 0.001$) and for Von Willebrand Factor (vWF) 0.38 ($p < 0.001$), respectively).[82] In the RE-LY trial,[79] the highest CRP quartile showed an adjusted HR of 1.7 (95% CI 1.3 – 2.2) compared to the lowest quartile for the composite endpoint of ischaemic stroke, systemic embolism, myocardial infarction, pulmonary embolism, and vascular death. In the same study, another marker for inflammation, interleukin-6 (IL-6), was associated with a HR of 2.5 (95% CI 1.9 – 3.3), and a comparable HR of 2.4 (95% CI 1.8 – 3.1) was found in a prospective cohort of inpatients with atrial fibrillation.[83] The addition of such measurements to CHA2DS2-VASc has yet to be investigated.

Finally, markers for *coagulation and endothelial (dys)function* seem promising additions to CHA2DS2-VASc for predicting stroke. Especially D-dimer, a well-established biomarker in the diagnosis of thromboembolism, has repeatedly been identified as a predictor for stroke in anticoagulated AF patients. In a small prospective cohort, D-dimer as a continuous measurement showed a HR of 1.4 (95% CI 1.0 – 1.9, unit of increment not reported) and the addition of 1 or 2 points for a D-dimer level ≥ 0.5ug/mL increased the c-statistic of CHADS2 from 0.78 to 0.83 or 0.85, respectively.[84] In a prospective cohort in Japan, the risk of stroke, TIA or peripheral embolism in the absence of risk factors was high at 3.8% per year for AF patients with a D-dimer level > 0.15 ug/mL.[85] Finally, in the Aristotle trial[86] a high D-dimer level was associated with stroke/systemic embolism (HR 1.7, 95% CI 1.1 – 2.6) and the c-statistic of CHADS2 improved from 0.646 to 0.655. The vWF, though not commonly measured in routine practice, also seems an independent risk factor for stroke. Although no independent association was found by Conway et al.[87] for vWF levels (mean level in the study population 145 IU/dL) below 131 IU/dL compared to 158 IU/dL, the HR was 3.7 (95% CI 2.0 – 4.5) for an optimal cut-point of 6 pg/mL (mean level in the study population 4 pg/mL, no IU/dL

reported) in the study by Pinto et al.[83]. Furthermore, a level > 194 IU/dL showed a HR of 2.2 (95% CI 1.5 – 3.2) in a prospective tertiary hospital based study (median vWF level in the study population 190 IU/dL) by García-Fernández et al.[88] It was shown that the c-statistic of CHA2DS2-VASc remained unchanged after addition of vWF, but future research in different settings need to confirm this finding.

2.3 Echocardiographic parameters

In addition to clinical parameters and blood measurements, cardiac imaging has been suggested to improve stroke risk prediction. Transthoracic echocardiography is a simple imaging technique and widely available in secondary care. It is recommended for every patient with atrial fibrillation for the assessment of structural heart disease (e.g. valvular disease), and ventricular size and function (systolic and diastolic).[89] It may additionally be considered for improving stroke risk prediction as well.[20]

A simple measurement of the left atrial size was first identified as a stroke risk factor some 25 years ago. The SPAF investigators[90] found an adjusted relative risk (RR) of 1.6 (95% CI 1.0 – 2.5) for atrial enlargement on the occurrence of thromboembolism. For comparison, in the same study a history of hypertension or thrombo-embolism were both associated with an adjusted RR of 1.9. Later prospective studies confirmed an increasing stroke risk with increasing atrial size in patients with atrial fibrillation. In a small sample of AF patients without risk factors for stroke at baseline, increased risk of cardiovascular events after 15 years of follow-up was shown if the left atrial size was large (adjusted HR 4.5, 95% CI 1.6 – 12.7).[91]

Besides atrial size, Saha et al.[92] showed that poor atrial function (e.g. low 'strain') was related to a high CHADS2 score (OR 0.9, 95% CI 0.8 – 1.0). Addition of both measurements to CHADS2 increased its performance in predicting cardiac events or death (p-value for chi-square test < 0.003). In a cross-sectional study, two measurements for left atrial strain were significantly lower in AF patients with a history of stroke as compared to patients without (OR 0.8 and 0.02, respectively).[93] Large contemporary longitudinal studies to assess the value of adding atrial function to a stroke prediction model such as CHA2DS2-VASc, however, are missing.

As many cardiac thrombi causing stroke in patients with atrial fibrillation originate from the left atrial appendage (LAA), thrombus formation in LAA, its morphology and function is of particular interest in stroke risk prediction. In the SPAF trial,[94] risk of stroke or thromboembolism was higher in 38 (10%) patients with a thrombus in the LAA as compared to those without thrombus (unadjusted RR 2.7, p = 0.04). Additionally, in AF patients at high risk of stroke the risk of

cardio-embolic events was increased 2.5-fold if the flow velocity as a measure of LAA function was reduced[95]

Other echocardiographic findings, or even different imaging modalities including MRI may in the future be of value in refining prediction of stroke in atrial fibrillation.[96,97] Currently, though, the value of such measurements in predicting stroke is unknown and the availability of such measurements is limited in daily practice.

DIRECTIONS FOR FUTURE RESEARCH

To improve the prediction of stroke in patients with atrial fibrillation, future research should take the necessary steps along the pathways described above. These steps may include conducting original studies but, importantly, could also rely on existing data.

The IPD meta-analysis of existing studies

For updating the CHA2DS2-VASc model by re-evaluating the individual predictors and any possible interaction, an analysis using individual patient data (IPD) of previously conducted studies in AF could be performed. In Chapter 1 of this thesis I identified multiple high quality validation studies that can provide the means to very efficiently address many, if not all, of the shortcomings of CHA2DS2-VASc in its current form.

Such an IPD meta-analysis offers several advantages. First, combining individual patients data form existing validation studies of CHA2DS2-VASc in a meta-analysis will result in a large sample size that allows for more robust modelling. In addition to a high number of outcome events that allows for the prediction of *subtypes* of outcomes (e.g. only stroke; the combination of stroke and TIA; or all thrombo-embolic events), interactions between CHA2DS2-VASc predictors that occur infrequently can efficiently be evaluated. Second, better standardisation can be achieved for the predictors.[98] For instance, IPD can provide insight in stroke risk differences for patients with *uncontrolled* hypertension as compared to *controlled* (i.e. treated) hypertension for future definitions of proven stroke risk predictors or even allow for modelling blood pressure as a continuous measurement, with possibly inclusion of interaction terms for treatment with hypertension. Third, the development of an update of CHA2DS2-VASc can directly be validated in the IPD dataset.[99] This allows for evaluating the performance of a model across different settings of care. The meta-analysis in Chapter 1 showed a different performance of CHA2DS2-VASc in studies recruiting patients from the general population compared to those

recruiting hospital patients, and such differences in populations, subgroups and settings can be taken into account when using IPD meta-analysis.

This makes an extensive analysis of existing individual patient data of previously conducted studies an essential first step and, importantly, it makes any new prospective validation study of CHA2DS2-VASc redundant.

Routine care data

Another promising and highly efficient source of data is routine healthcare.[98] Patient data, either collected during routine clinical practice such as electronic patient files or collected for administrative purposes such as hospital discharge records or health insurance databases, is increasingly stored digitally and in standardised manner (e.g. using ICD-10, READ or ICPC codes). Future research can benefit from this trend in finding stroke risk predictors other than those included in CHA2DS2-VASc. Exposure status of candidate predictors as mentioned above (e.g. renal failure or OSAS) may be obtained from routine care databases, as does the occurrence of stroke. Of course, the validity of such data sources is then of utmost importance and methodologic research should focus on ways to assess and improve data validity and quality.[100–102] Chapter 6 of my thesis provides an example of such practice and shows substantial misclassification in predictor values between diagnosis codes (ICPC) collected during routine care, in a comparison with values manually collected by scrutinising the complete patient file. While this study concludes that for *validating* a prediction model the influence of this misclassification is only minor, for *updating* a model such as CHA2DS2-VASc this may be different.

Biobanks

One specific source of readily available data for investigating stroke risk prediction are 'biobanks'. Here, large amounts of biomaterial such as blood samples are stored for future (scientific) use. Combined with information on clinical parameters such as the presence of atrial fibrillation and the occurrence of stroke, this biomaterial provides great opportunities to identify biomarkers as stroke predictors. Ideally, such biobanks are further linked to prospective cohorts or routine care data sources for the collection of data on multiple (types of) potential stroke predictors and outcomes at once. One local example includes the Utrecht Cardiovascular Cohort[103] that records standardised information on patients with cardiovascular disease or major cardiovascular risk factors referred to the University Medical Center Utrecht for diagnostic testing or treatment. The combination of both biomarker measurements and standardised routine care data offers extremely valuable input for studies aiming to improve stroke risk prediction in atrial fibrillation.

Original studies

Individual patient data, routine care data and existing biobanks, however, may not be able to provide *all* necessary answers for the issues described above. Although these sources can be used to both generate and test hypotheses, original studies are still needed in the future. Examples include original case-control or case-cohort designs for efficient identification of stroke risk factors not measured in existing validation studies or during routine healthcare. Prospective cohorts and even clinical trials such as the one described in Chapter 3 of this thesis may be needed to validate future prediction models and assess the impact of their use in daily practice.

Clearly though, the most important and most promising first steps that will lead to the urgently needed improvement in stroke risk prediction rely on data that is readily available. All that is needed is the willingness of researchers to share their individual patient data, the efforts of clinicians and other care providers to accurately record data during routine practice and a large dose of commitment and cooperation.

OPTIMISING STROKE RISK PREDICTION VERSUS APPLICABILITY IN DAILY PRACTICE

Ultimately, optimised stroke risk prediction should lead to optimised patient outcome. As a means to stroke prevention, therefore, prediction models first and foremost need to be easily applicable in daily clinical practice. Assigning different weights to predictors, using continuous measurements, or including interaction terms will certainly complicate stroke risk models. Furthermore, the measurement of biomarkers or echocardiographic parameters may require time and equipment that is not available at the point of care. The competing interests of optimising stroke risk prediction versus applicability in daily practice need continuous attention. Strategies applying more complicated stroke prediction models *only* in patients identified as low to intermediate stroke risk by traditional (simple) models, i.e. those in who uncertainty on anticoagulant treatment decisions remains, may be considered. Furthermore, technological advances may be used to improve the applicability of complicated prediction models. Examples include web-based calculators such as deployed by the Qstroke model[50], automated stroke risk predictions based on data from routine healthcare such as evaluated in Chapter 3 of my thesis, or mobile phone applications.

CLOSING REMARKS

Atrial fibrillation brings many challenges to physicians, but also to scientists in search for ways to improve patient care. The prediction of stroke is a major part of AF care and this thesis aimed to address some important issues. However, stroke prediction is only the beginning. I have not touched upon the topic of predicting the risk of bleeding, or how to incorporate such risks in the treatment with anticoagulants. Furthermore, as shown in Chapter 4 of this thesis, patients with AF are at risk of many adverse events other than stroke and it is still unclear how to reduce these risks. Continuous efforts are needed to clarify these issues and, ultimately, improve the entire spectrum of care for patients with atrial fibrillation.

REFERENCES

1 Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, Singer DE. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* 2001; 285: 2370–5.

2 Heeringa J, van der Kuip DAM, Hofman A, Kors JA, van Herpen G, Stricker BHC, Stijnen T, Lip GYH, Witteman JCM. Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study. *Eur Heart J* 2006; 27: 949–53.

3 Lloyd-Jones DM, Wang TJ, Leip EP, Larson MG, Levy D, Vasan RS, D'Agostino RB, Massaro JM, Beiser A, Wolf PA, Benjamin EJ. Lifetime Risk for Development of Atrial Fibrillation. *Circulation* 2004; 110: 1042–6.

4 Odutayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG, Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. *BMJ* 2016; 354: i4482.

5 Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991; 22: 983–8.

6 Olesen JB, Lip GYH, Hansen ML, Hansen PR, Tolstrup JS, Lindhardsen J, Selmer C, Ahlehoff O, Olsen A-MS, Gislason GH, Torp-Pedersen C. Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study. *BMJ* 2011; 342: d124.

7 van den Akker M, Buntinx F, Metsemakers JFM, Roos S, Knottnerus JA. Multimorbidity in General Practice: Prevalence, Incidence, and Determinants of Co-Occurring Chronic and Recurrent Diseases. *J Clin Epidemiol* 1998; 51: 367–75.

8 LaMori JC, Mody SH, Gross HJ, daCosta DiBonaventura M, Patel AA, Schein JR, Nelson WW. Burden of comorbidities among patients with atrial fibrillation. *Ther Adv Cardiovasc Dis* 2013; 7: 53–62.

9 Payne RA, Avery AJ, Duerden M, Saunders CL, Simpson CR, Abel GA. Prevalence of polypharmacy in a Scottish primary care population. - PubMed - NCBI. *Eur J Clin Pharmacol* 2014; 70: 575–81.

10 Proietti M, Raparelli V, Olshansky B, Lip GYH. Polypharmacy and major adverse events in atrial fibrillation: observations from the AFFIRM trial. *Clin Res Cardiol* 2015.

11 Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, Pogue J, Reilly PA, Themeles E, Varrone J, Wang S, Alings M, Xavier D, Zhu J, Diaz R, Lewis BS, Darius H, Diener H-C, Joyner CD, Wallentin L, et al. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2009; 361: 1139–51.

12 Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, Breithardt G, Halperin JL, Hankey GJ, Piccini JP, Becker RC, Nessel CC, Paolini JF, Berkowitz SD, Fox KAA, Califf RM, ROCKET AF Investigators. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med* 2011; 365: 883–91.

13 Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M, Al-Khalidi HR, Ansell J, Atar D, Avezum A, Bahit MC, Diaz R, Easton JD, Ezekowitz JA, Flaker G, Garcia D, Geraldes M, Gersh BJ, Golitsyn S, Goto S, et al. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011; 365: 981–92.

14 Giugliano RP, Ruff CT, Braunwald E, Murphy SA, Wiviott SD, Halperin JL, Waldo AL, Ezekowitz MD, Weitz JI, Špinar J, Ruzyllo W, Ruda M, Koretsune Y, Betcher J, Shi M, Grip

LT, Patel SP, Patel I, Hanyok JJ, Mercuri M, et al. Edoxaban versus Warfarin in Patients with Atrial Fibrillation. *N Engl J Med* 2013; 369: 2093–104.

15 Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Ann Intern Med* 2007; 146: 857–67.

16 Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. 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.

17 European Heart Rhythm Association, European Association for Cardio-Thoracic Surgery, Camm AJ, Kirchhof P, Lip GYH, Schotten U, Savelieva I, Ernst S, Van Gelder IC, Al-Attar N, Hindricks G, Prendergast B, Heidbuchel H, Alfieri O, Angelini A, Atar D, Colonna P, De Caterina R, De Sutter J, Goette A, et al. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). *Eur Heart J*. 2010. pp. 2369–429.

18 Camm AJ, Lip GYH, De Caterina R, Savelieva I, Atar D, Hohnloser SH, Hindricks G, Kirchhof P, ESC Committee for Practice Guidelines (CPG), Kirchhof P, Kolh P. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation. Developed with the special contribution of the European Heart Rhythm Association. *Eur Heart J*. 2012. pp. 2719–47.

19 Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of Clinical Classification Schemes for Predicting Stroke: Results From the National Registry of Atrial Fibrillation. *JAMA* 2001; 285: 2864–70.

20 NICE. Atrial fibrillation: the management of atrial fibrillation. (Clinical guideline 180). NICE; 2014.

21 January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, Conti JB, Ellinor PT, Ezekowitz MD, Field ME, Murray KT, Sacco RL, Stevenson WG, Tchou PJ, Tracy CM, Yancy CW, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol*. 2014. pp. e1–76.

22 Hobbs FR, Taylor CJ, Jan Geersing G, Rutten FH, Brouwer JR, on behalf of the European Primary Care Cardiovascular Society (EPCCS) SPAF working group. European Primary Care Cardiovascular Society (EPCCS) consensus guidance on stroke prevention in atrial fibrillation (SPAF) in primary care. *Eur J Prev Cardiol* 2015; : 2047487315571890.

23 Dutch College of General Practitioners Guideline Development Group for Atrial fibrillation. Guideline Atrial fibrillation (second partial revision). *Huisarts Wet* 2013: 392–401.

24 Atrial Fibrillation Investigators. Risk factors for stroke and efficacy of antithrombotic therapy in atrial fibrillation. Analysis of pooled data from five randomized controlled trials. *Arch Intern Med* 1994; 154: 1449–57.

25 Stroke Prevention in Atrial Fibrillation Investigators. Risk factors for thromboembolism during aspirin therapy in patients with atrial fibrillation: The stroke prevention in atrial fibrillation study. *J Stroke Cerebrovasc Dis*1995; 5: 147–57.

26 Remme WJ, Swedberg K, Task Force for the Diagnosis and Treatment of Chronic Heart Failure, European Society of Cardiology. Guidelines for the diagnosis and treatment of chronic heart failure. *Eur Heart J* 2001. pp. 1527–60.

27 Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, Falk V, González-Juanatey JR, Harjola V-P, Jankowska EA, Jessup M, Linde C, Nihoyannopoulos P, Parissis JT, Pieske B, Riley JP, Rosano GMC, Ruilope LM, Ruschitzka F, Rutten FH, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 2016; 37: 2129–200.

28 Jhund PS, MacIntyre K, Simpson CR, Lewsey JD, Stewart S, Redpath A, Chalmers JWT, Capewell S, McMurray JJV. Long-term trends in first hospitalization for heart failure and subsequent survival between 1986 and 2003: a population study of 5.1 million people. *Circulation* 2009; 119: 515–23.

29 Christiansen MN, Kober L, Weeke P, Vasan RS, Jeppesen JL, Smith JG, Gislason GH, Torp-Pedersen C, Andersson C. Age-Specific Trends in Incidence, Mortality, and Comorbidities of Heart Failure in Denmark, 1995 to 2012. *Circulation* 2017; 135: 1214–23.

30 Stroke Risk in Atrial Fibrillation Working Group. Independent predictors of stroke in patients with atrial fibrillation: a systematic review. *Neurology* 2007; 69: 546–54.

31 Friberg L, Rosenqvist M, Lip GYH. Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: the Swedish Atrial Fibrillation cohort study. *Eur Heart J* 2012; 33: 1500–10.

32 van den Ham HA, Klungel OH, Singer DE, Leufkens HGM, van Staa TP. Comparative Performance of ATRIA, CHADS2, and CHA2DS2-VASc Risk Scores Predicting Stroke in Patients With Atrial Fibrillation: Results From a National Primary Care Database. *J Am Coll Cardiol* 2015; 66: 1851–9.

33 Aspberg S, Chang Y, Atterman A, Bottai M, Go AS, Singer DE. Comparison of the ATRIA, CHADS2, and CHA2DS2-VASc stroke risk scores in predicting ischaemic stroke in a large Swedish cohort of patients with atrial fibrillation. *Eur Heart J* 2016; 37: 3203–10.

34 Chao T-F, Liu C-J, Wang K-L, Lin Y-J, Chang S-L, Lo L-W, Hu Y-F, Tuan T-C, Chen T-J, Lip GYH, Chen S-A. Should Atrial Fibrillation Patients With 1 Additional Risk Factor of the CHA2DS2-VASc Score (Beyond Sex) Receive Oral Anticoagulation? *J Am Coll Cardiol* 2015; 65: 635–42.

35 Huang D, ANGUO L, YUE W-S, YIN L, Tse H-F, Siu C-W. Refinement of ischemic stroke risk in patients with atrial fibrillation and CHA2 DS2 -VASc score of 1. *Pacing Clin Electrophysiol* 2014; 37: 1442–7.

36 Overvad TF, Skjøth F, Lip GYH, Lane DA, Albertsen IE, Rasmussen LH, Larsen TB. Duration of Diabetes Mellitus and Risk of Thromboembolism and Bleeding in Atrial Fibrillation Nationwide Cohort Study. *Stroke* 2015; 46: 2168–74.

37 Rao MP, Halvorsen S, Wojdyla D, Thomas L, Alexander JH, Hylek EM, Hanna M, Bahit MC, Lopes RD, De Caterina R, Erol C, Goto S, Lanas F, Lewis BS, Husted S, Gersh BJ, Wallentin L, Granger CB, Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation (ARISTOTLE) Steering Committee and Investigators. Blood Pressure Control and Risk of Stroke or Systemic Embolism in Patients With Atrial Fibrillation: Results From the Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation (ARISTOTLE) Trial. *J Am Heart Assoc* 2015; 4: e002015.

38 Vemulapalli S, Hellkamp AS, Jones WS, Piccini JP, Mahaffey KW, Becker RC, Hankey GJ, Berkowitz SD, Nessel CC, Breithardt G, Singer DE, Fox KAA, Patel MR. Blood pressure

control and stroke or bleeding risk in anticoagulated patients with atrial fibrillation: Results from the ROCKET AF Trial. *Am Heart J* 2016; 178: 74–84.

39 Anandasundaram B, Lane DA, Apostolakis S, Lip GY. The impact of atherosclerotic vascular disease in predicting a stroke, thromboembolism and mortality in atrial fibrillation patients: a systematic review. *J Thromb Haemost* 2013; 11: 975–87.

40 McAlister FA, Wiebe N, Jun M, Sandhu R, James MT, McMurtry MS, Hemmelgarn BR, Tonelli M. 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: 243–52.

41 Nielsen PB, Larsen TB, Skjøth F, Overvad TF, Lip GYH. Stroke and thromboembolic event rates in atrial fibrillation according to different guideline treatment thresholds: A nationwide cohort study. *Sci Rep* 2016; 6: 27410.

42 Okumura K, Inoue H, Atarashi H, Yamashita T, Tomita H, Origasa H. Validation of CHA2DS2-VASc and HAS-BLED Scores in Japanese Patients With Nonvalvular Atrial Fibrillation. *Circ J* 2014; 78: 1593–9.

43 Chao TF, Lip GY, Liu CJ, Tuan TC, Chen SJ, Wang KL, Lin YJ, Chang SL, Lo LW, Hu YF, Chen TJ, Chiang CE, Chen SA. Validation of a Modified CHA2DS2-VASc Score for Stroke Risk Stratification in Asian Patients with Atrial Fibrillation. *Stroke* 2016; 47: 2462–9.

44 Chan PH, Lau C-P, Tse H-F, Chiang C-E, Siu C-W. CHA2DS2-VASc Recalibration With an Additional Age Category (50-64 Years) Enhances Stroke Risk Stratification in Chinese Patients With Atrial Fibrillation. *Can J Cardiol*2016.

45 Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Ann Intern Med* 2015; 162: W1–73.

46 Kirchhof P, Nabauer M, Gerth A, Limbourg T, Lewalter T, Goette A, Wegscheider K, Treszl A, Meinertz T, Oeff M, Ravens U, Breithardt G, Steinbeck G. Impact of the type of centre on management of AF patients: surprising evidence for differences in antithrombotic therapy decisions. *Thromb Haemost* 2011; 105: 1010–23.

47 Olesen JB, Fauchier L, Lane DA, Taillandier S, Lip GYH. Risk factors for stroke and thromboembolism in relation to age among patients with atrial fibrillation: the Loire Valley Atrial Fibrillation Project. *Chest* 2012; 141: 147–53.

48 Go AS, Fang MC, Udaltsova N, Chang Y, Pomernacki NK, Borowsky L, Singer DE, ATRIA Study Investigators. Impact of proteinuria and glomerular filtration rate on risk of thromboembolism in atrial fibrillation: the anticoagulation and risk factors in atrial fibrillation (ATRIA) study. *Circulation* 2009; 119: 1363–9.

49 Singer DE, Chang Y, Borowsky LH, Fang MC, Pomernacki NK, Udaltsova N, Reynolds K, Go AS. A new risk scheme to predict ischemic stroke and other thromboembolism in atrial fibrillation: the ATRIA study stroke risk score. *J Am Heart Assoc* 2013; 2: e000250–0.

50 Hippisley-Cox J, Coupland C, Brindle P. Derivation and validation of QStroke score for predicting risk of ischaemic stroke in primary care and comparison with other risk scores: a prospective open cohort study. *BMJ* 2013; 346: f2573.

51 Friberg L, Benson L, Lip GYH. Balancing stroke and bleeding risks in patients with atrial fibrillation and renal failure: the Swedish Atrial Fibrillation Cohort study. *Eur Heart J* 2015; 36: 297–306.

52 Leening MJG, Vedder MM, Witteman JCM, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Ann Intern Med* 2014; 160: 122–31.
53 Banerjee A, Fauchier L, Vourc'h P, Andres CR, Taillandier S, Halimi JM, Lip GYH. Renal impairment and ischemic stroke risk assessment in patients with atrial fibrillation: the Loire Valley Atrial Fibrillation Project. *J Am Coll Cardiol* 2013; 61: 2079–87.
54 Roldan V, Marin F, Manzano-Fernandez S, Fernandez H, Gallego P, Valdes M, Vicente V, Lip GY. Does chronic kidney disease improve the predictive value of the CHADS2 and CHA2DS2-VASc stroke stratification risk scores for atrial fibrillation? *Thromb Haemost* 2013; 109: 956–60.
55 Hilden J, Gerds TA. A note on the evaluation of novel biomarkers: do not rely on integrated discrimination improvement and net reclassification index. *Stat Med* 2014; 33: 3405–14.
56 Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Pepe MS. Net reclassification indices for evaluating risk prediction instruments: a critical review. *Epidemiology* 2014; 25: 114–21.
57 Hilden J. Commentary: On NRI, IDI, and "good-looking" statistics with nothing underneath. *Epidemiology* 2014; 25: 265–7.
58 Vickers AJ. Decision analysis for the evaluation of diagnostic tests, prediction models and molecular markers. *Am Stat* 2008; 62: 314–20.
59 Javaheri S, Barbe F, Campos-Rodriguez F, Dempsey JA, Khayat R, Javaheri S, Malhotra A, Martinez-Garcia MA, Mehra R, Pack AI, Polotsky VY, Redline S, Somers VK. Sleep Apnea: Types, Mechanisms, and Clinical Cardiovascular Consequences. *J Am Coll Cardiol* 2017; 69: 841–58.
60 Holmqvist F, Guan N, Zhu Z, Kowey P, Allen L. Obstructive sleep apnea and atrial fibrillation: findings from ORBIT-AF. *Am Heart J* 2015. doi:10.1016/j.ahj.2014.12.024
61 Yaranov DM, Smyrlis A, Usatii N, Butler A, Petrini JR, Mendez J, Warshofsky MK. Effect of obstructive sleep apnea on frequency of stroke in patients with atrial fibrillation. *Am J Cardiol* 2015; 115: 461–5.
62 Mansukhani MP, Calvin AD, Kolla BP, Brown RD, Lipford MC, Somers VK, Caples SM. The association between atrial fibrillation and stroke in patients with obstructive sleep apnea: a population-based case-control study. *Sleep Med* 2013; 14: 243–6.
63 Chang C-C, Chiu C-C, Chiang C-H, Huang C-C, Chan W-L, Huang *P*-H, Chen Y-C, Chen T-J, Chung C-M, Lin S-J, Chen J-W, Leu H-B. Obstructive sleep apnea and the risk of ischemic stroke in patients with atrial fibrillation. *In journal cardiol* 2015; 181: 144–6.
64 Lip GYH, Brechin CM, Lane DA. The global burden of atrial fibrillation and stroke: a systematic review of the epidemiology of atrial fibrillation in regions outside North America and Europe. *Chest* 2012; 142: 1489–98.
65 Shen AY-J, Yao JF, Brar SS, Jorgensen MB, Wang X, Chen W. Racial/Ethnic differences in ischemic stroke rates and the efficacy of warfarin among patients with atrial fibrillation. *Stroke* 2008; 39: 2736–43.
66 Shroff GR, Solid CA, Herzog CA. Atrial fibrillation, stroke, and anticoagulation in Medicare beneficiaries: trends by age, sex, and race, 1992-2010. *J Am Heart Assoc* 2014; 3: e000756.
67 Quinn GR, Severdija ON, Chang Y, Singer DE. Wide Variation in Reported Rates of Stroke Across Cohorts of Patients With Atrial Fibrillation. *Circulation* 2017; 135: 208–19.
68 van Doorn S, Debray TPA, Kaasenbrood F, Hoes AW, Rutten FH, Moons KGM, Geersing GJ. Predictive performance of the CHA2DS2-VASc rule in atrial fibrillation: a systematic review and meta-analysis. *J Thromb Haemost* 2017; 15: 1065–77.

69 Kabra R, Girotra S, Vaughan Sarrazin M. Refining Stroke Prediction in Atrial Fibrillation Patients by Addition of African-American Ethnicity to CHA2DS2-VASc Score. *J Am Coll Cardiol* 2016; 68: 461–70.

70 Hijazi Z, Oldgren J, Siegbahn A, Granger CB, Wallentin L. Biomarkers in atrial fibrillation: a clinical review. *Eur Heart J* 2013; 34: 1475–80.

71 Hijazi Z, Oldgren J, Andersson U, Connolly SJ, Ezekowitz MD, Hohnloser SH, Reilly PA, Vinereanu D, Siegbahn A, Yusuf S, Wallentin L. Cardiac biomarkers are associated with an increased risk of stroke and death in patients with atrial fibrillation: a Randomized Evaluation of Long-term Anticoagulation Therapy (RE-LY) substudy. *Circulation* 2012; 125: 1605–16.

72 Hijazi Z, Wallentin L, Siegbahn A, Andersson U, Christersson C, Ezekowitz J, Gersh BJ, Hanna M, Hohnloser S, Horowitz J, Huber K, Hylek EM, Lopes RD, McMurray JJV, Granger CB. N-terminal pro-B-type natriuretic peptide for risk assessment in patients with atrial fibrillation: insights from the ARISTOTLE Trial (Apixaban for the Prevention of Stroke in Subjects With Atrial Fibrillation). *J Am Coll Cardiol* 2013; 61: 2274–84.

73 Hijazi Z, Siegbahn A, Andersson U, Granger CB, Alexander JH, Atar D, Gersh BJ, Mohan P, Harjola VP, Horowitz J, Husted S, Hylek EM, Lopes RD, McMurray JJV, Wallentin L, on behalf of the ARISTOTLE Investigators. High-sensitivity troponin I for risk assessment in patients with atrial fibrillation: insights from the Apixaban for Reduction in Stroke and other Thromboembolic Events in Atrial Fibrillation (ARISTOTLE) trial. *Circulation* 2014; 129: 625–34.

74 Hijazi Z, Hijazi Z, Wallentin L, Wallentin L, Siegbahn A, Siegbahn A, Andersson U, Andersson U, Alexander JH, Alexander JH, Atar D, Atar D, Gersh BJ, Gersh BJ, Hanna M, Hanna M, Harjola VP, Harjola V-P, Horowitz JD, Horowitz JD, et al. High-sensitivity troponin T and risk stratification in patients with atrial fibrillation during treatment with apixaban or warfarin. *J Am Coll Cardiol* 2014; 63: 52–61.

75 Szymanski FM, Lip GYH, Filipiak KJ, Platek AE, Hrynkiewicz-Szymanska A, Opolski G. Stroke Risk Factors Beyond the CHA_2DS_2-VASc Score: Can We Improve Our Identification of "High Stroke Risk" Patients With Atrial Fibrillation? *Am J Cardiol* 2015; 116: 1781–8.

76 Hijazi Z, Oldgren J, Siegbahn A, Wallentin L. Application of Biomarkers for Risk Stratification in Patients with Atrial Fibrillation. *Clin Chem* 2017; 63: 152–64.

77 Frustaci A, Chimenti C, Bellocci F, Morgante E. Histological substrate of atrial biopsies in patients with lone atrial fibrillation. *Circulation* 1997.

78 Chung MK, Martin DO, Sprecher D, Wazni O. C-reactive protein elevation in patients with atrial arrhythmias. *Circulation* 2001.

79 Aulin J, Siegbahn A, Hijazi Z, Ezekowitz MD, Andersson U, Connolly SJ, Huber K, Reilly PA, Wallentin L, Oldgren J. Interleukin-6 and C-reactive protein and risk for death and cardiovascular events in patients with atrial fibrillation. *Am Heart J* 2015; 170: 1151–60.

80 Hermida J, Lopez FL, Montes R, Matsushita K, Astor BC, Alonso A. Usefulness of High-Sensitivity C-Reactive Protein to Predict Mortality in Patients With Atrial Fibrillation (from the Atherosclerosis Risk In Communities [ARIC] Study). *The Am J Cardiol* 2012; 109: 95–9.

81 Lip GYH, Patel JV, Hughes E, Hart RG. High-sensitivity C-reactive protein and soluble CD40 ligand as indices of inflammation and platelet activation in 880 patients with nonvalvular atrial fibrillation: relationship to stroke risk factors, stroke risk stratification schema, and prognosis. *Stroke* 2007; 38: 1229–37.

82 Conway DSG, Buggins P, Hughes E, Lip GYH. Relationship of interleukin-6 and C-Reactive protein to the prothrombotic state in chronic atrial fibrillation. *J Am Coll Cardiol* 2004; 43: 2075–82.

83 Pinto A, Tuttolomondo A, Casuccio A, Di Raimondo D, Di Sciacca R, Arnao V, Licata G. Immuno-inflammatory predictors of stroke at follow-up in patients with chronic non-valvular atrial fibrillation (NVAF). *Clin Sci* 2009; 116: 781–9.

84 Sadanaga T, Kohsaka S, Ogawa S. D-dimer levels in combination with clinical risk factors can effectively predict subsequent thromboembolic events in patients with atrial fibrillation during oral anticoagulant therapy. *Cardiology* 2010; 117: 31–6.

85 Nozawa T, Inoue H, Hirai T, Iwasa A, Okumura K, Lee J-D, Shimizu A, Hayano M, Yano K. D-dimer level influences thromboembolic events in patients with atrial fibrillation. *Int J Cardiol* 2006; 109: 59–65.

86 Christersson C, Wallentin L, Andersson U, Alexander JH, Ansell J, De Caterina R, Gersh BJ, Granger CB, Hanna M, Horowitz JD, Huber K, Husted S, Hylek EM, Lopes RD, Siegbahn A. D-dimer and risk of thromboembolic and bleeding events in patients with atrial fibrillation - observations from the ARISTOTLE trial. *J Thromb Haemost* 2014; : n/a–n/a.

87 Conway DSG, Pearce LA, Chin BSP, Hart RG, Lip GYH. Prognostic value of plasma von Willebrand factor and soluble *P*-selectin as indices of endothelial damage and platelet activation in 994 patients with nonvalvular atrial fibrillation. *Circulation* 2003; 107: 3141–5.

88 García-Fernández A, Roldan V, Rivera-Caravaca JM, Hernandez-Romero D, Valdes M, Vicente V, Lip GYH, Marin F. Does von Willebrand factor improve the predictive ability of current risk stratification scores in patients with atrial fibrillation? *Sci Rep* 2017; 7: 41565.

89 Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, Castella M, Diener H-C, Heidbuchel H, Hendriks J, Hindricks G, Manolis AS, Oldgren J, Popescu BA, Schotten U, Van Putte B, Vardas P, Agewall S, Camm J, Barón-Esquivias G, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J* 2016; 37: 2893–962.

90 Stroke Prevention in Atrial Fibrillation Investigators. Predictors of thromboembolism in atrial fibrillation: II. Echocardiographic features of patients at risk. The Stroke Prevention in Atrial Fibrillation Investigators. *Ann Intern Med* 1992; 116: 6–12.

91 Osranek M, Bursi F, Bailey KR, Grossardt BR, Brown RD, Kopecky SL, Tsang TS, Seward JB. Left atrial volume predicts cardiovascular events in patients originally diagnosed with lone atrial fibrillation: three-decade follow-up. *Eur Heart J* 2005; 26: 2556–61.

92 Saha SK, Anderson PL, Caracciolo G, Kiotsekoglou A, Wilansky S, Govind S, Mori N, Sengupta PP. Global left atrial strain correlates with CHADS2 risk score in patients with atrial fibrillation. *J Am Soc Echocardiogr* 2011; 24: 506–12.

93 Shih J-Y, Tsai W-C, Huang Y-Y, Liu Y-W, Lin C-C, Huang Y-S, Tsai L-M, Lin L-J. Association of decreased left atrial strain and strain rate with stroke in chronic atrial fibrillation. *J Am Soc Echocardiogr* 2011; 24: 513–9.

94 Echocardiography TSPIAFICO. Transesophageal echocardiographic correlates of thromboembolism in high-risk patients with nonvalvular atrial fibrillation. The Stroke Prevention in Atrial Fibrillation Investigators Committee on Echocardiography. *Ann Intern Med* 1998; 128: 639–47.

95 Goldman ME, Pearce LA, Hart RG, Zabalgoitia M. Pathophysiologic correlates of thromboembolism in nonvalvular atrial fibrillation: I. Reduced flow velocity in the left atrial appendage. *J Am Soc Echocardiogr* 1999.

96 Providencia R, Trigo J, Paiva L, Barra S. The role of echocardiography in thromboembolic risk assessment of patients with nonvalvular atrial fibrillation. *J Am Soc Echocardiogr* 2013; 26: 801–12.
97 Calenda BW, Fuster V, Halperin JL, Granger CB. Stroke risk assessment in atrial fibrillation: risk factors and markers of atrial myopathy. *Nat Rev Cardiol* 2016; 13: 549–59.
98 Riley RD, Ensor J, Snell KIE, Debray TPA, Altman DG, Moons KGM, Collins GS. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ* 2016; 353: i3140.
99 Debray TPA, Riley RD, Rovers MM, Reitsma JB, Moons KGM, group CIM-AM. Individual Participant Data (IPD) Meta-analyses of Diagnostic and Prognostic Modeling Studies: Guidance on Their Use. *PLoS Med* 2015; 12: e1001886.
100 Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, Elm von E, Langan SM, Committee RW. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Med* 2015; 12: e1001885.
101 Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010; 69: 4–14.
102 Jordan K, Porcheret M, Croft P. Quality of morbidity coding in general practice computerized medical records: a systematic review. *Fam Pract* 2004; 21: 396–412.
103 Asselbergs FW, Visseren FL, Bots ML, de Borst GJ, Buijsrogge MP, Dieleman JM, van Dinther BG, Doevendans PA, Hoefer IE, Hollander M, de Jong PA, Koenen SV, Pasterkamp G, Ruigrok YM, van der Schouw YT, Verhaar MC, Grobbee DE. Uniform data collection in routine clinical practice in cardiovascular patients for optimal care, quality control and research: The Utrecht Cardiovascular Cohort. *Eur J Prev Cardiol* 2017; 24: 840–7.

Addenda

Summary

INTRODUCTION

Atrial fibrillation (AF) is the most common cardiac arrhythmia. The prevalence in the general population is 1–2%, rising to over 15% in elderly over 85 years of age. [1,2] It is well-known as a major risk factor for ischaemic stroke, but AF is also associated with increased risk of hospitalisation,[3] mortality[4] and a multitude of other adverse events.[5]

In this thesis, we address the management of these risks in patients with AF, particularly in general practice. We focus on three domains: the prediction of ischaemic stroke in AF; preventive treatment with anticoagulants; and adverse events other than stroke, including hospitalisation, mortality and heart failure.

STROKE RISK PREDICTION IN ATRIAL FIBRILLATION

In stroke prevention in AF, the benefits of the highly effective anticoagulants need to be balanced against their associated risk of bleeding complications. For this, practice guidelines on the management of AF recommend predicting stroke risk in every individual patient with AF, and subsequently anticoagulation only for those at high risk of stroke. In **Chapter 1** of this thesis, we evaluate the performance of the CHA2DS2-VASc prediction model as developed in 2010 by Lip et al.[6] Although widely used for predicting stroke risk in patients with AF, external validation studies of CHA2DS2-VASc show ambiguous and conflicting results.

We performed systematic literature review of existing studies validating CHA2DS2-VASc in AF patients not (yet) anticoagulated and a random effects meta-analysis of estimates of stroke risk per score. To explore between-study heterogeneity in stroke risk, we stratified studies to the clinical setting in which patient enrolment started, and performed meta-regression.

In total 19 studies were evaluated with over two million person-years of follow-up. In studies recruiting AF patients in hospitals, stroke risk for a score of zero, one and two were 0.4% (approximate 95% prediction interval (PI) 0.2 to 3.2%), 1.2% (95% PI 0.1 – 3.8%) and 2.2% (95% PI 0.03 – 7.8%), respectively. This was consistently higher than studies recruiting patients from the open general population, with risks of 0.2% (95% PI 0.0 – 0.9%), 0.7% (0.3 – 1.2%) and 1.5% (95% PI 0.4 – 3.3%) for score zero to two respectively. Heterogeneity as reflected by the wide prediction intervals could not be fully explained by meta-regression.

We conclude that studies validating CHA2DS2-VASc demonstrate high heterogeneity in predicted stroke risks for different scores. As such, it may have difficulties in tailoring anticoagulant treatment adequately in AF patients at intermediate risk

of stroke (roughly those with a score of 1 or 2). Differences in stroke risks between studies recruiting from hospitals or from the general population indicate that possible case-mix differences between populations should be taken into account in clinical decision-making but further uncertainty remains.

ANTICOAGULANT TREATMENT IN ATRIAL FIBRILLATION

While their benefits for prevention of stroke in AF are evident, anticoagulants are often underused. In a meta-analysis in 2010, Ogilvie et al.[7] found many studies to report underuse of anticoagulants in 30% of patients at high risk of stroke.

In **Chapter 2** of this thesis, we evaluate the reasons for non-adherence to practice guidelines on stroke prevention in AF in general practice.

In an observational cross-sectional study in Dutch general practice, the management of 440 AF patients from 19 general practices was analysed, and reasons for non-adherence to the CHADS2 score,[8] the precursor of CHA2DS2-VASc, per individual patient assessed by asking the GP.

The median age of the included patients was 76.0 (interquartile range (IQR) 16.0) years, 55% were male, and 61.6% received cooperative care from the cardiologist. Undertreatment according to the CHADS2 and CHA2DS2-VASc decision rules occurred in 93 (21.1%) and 104 (23.6%) patients, respectively. Overtreatment occurred in 84 (19.1%) and 29 (3.4%) patients, respectively. The main reasons mentioned by GPs for non-adherence per individual undertreated case was i) sustained sinus rhythm after an episode of AF, and ii) the cardiologist was considered responsible for the anticoagulation. We hypothesise that these barriers may be overcome by clearer evidence-based recommendations on the management of patients with (presumed) sinus rhythm after AF, and better agreement on responsibility of initiation and monitoring of anticoagulation therapy between GPs and cardiologists. Interestingly, adverse effects or contra-indications to drugs, or patient's preferences were seldom mentioned as reasons for undertreatment.

As a next step, in order to improve the anticoagulant management of patients with AF, we undertook a large cluster-randomised trial assessing the effectiveness of automated CHA2DS2-VASc based decision support on stroke prevention in patients with AF in general practice, that we describe in **Chapter 3**

We randomised a total of 38 practices. In the index practices, GPs were provided with an automatically generated CHA2DS2-VASc based anticoagulant treatment recommendation. The GPs in the reference practices provided care as usual. The primary outcome was the incidence of ischaemic stroke, TIA and/or thromboem-

bolism (TE). Secondary outcomes were bleeding and the proportion of patients on guideline recommended anticoagulant treatment.

In total, 1129 AF patients were included in the 19 index practices and 1226 AF patients in the 19 reference practices. The median age was 77 (IQR 68 – 75) years, and 48% were male. The median CHA2DS2-VASc score was 3.0 (IQR 2.0 – 5.0). Underuse of anticoagulants as recommended for patients with CHA2DS2-VASc score ≥ 2 was 6.6%. After a median follow-up of 2.7 (IQR 2.3 – 3.0) years, the incidence rate per 100 person-years of ischaemic stroke/TIA/TE was 1.96 in the index group and 1.42 in the reference group (hazard ratio (HR) 1.3, 95% C 0.8 – 2.1). No difference was observed in the rate of bleeding (0.79 versus 0.82), or in the underuse (7.2% versus 8.2%) or overuse (8.0% versus 7.9%) of anticoagulation.

We conclude that in this study in patients with AF in general practice, underuse of anticoagulants is relatively low. Nevertheless, providing general practitioners with CHA2DS2-VASc based decision support does not seem to result in a reduction in stroke incidence, nor does it affect bleeding risk or anticoagulant over- or underuse.

RISK OF HOSPITALISATION, MORTALITY AND HEART FAILURE IN PATIENTS WITH ATRIAL FIBRILLATION

In the subsequent two chapters of this thesis, we focus on the increased risk in AF patients of adverse events other than stroke alone. Their clinical importance in increasingly recognised, as exemplified by recent practice guidelines recommending integrated AF care aiming to prevent all clinically relevant outcomes, including hospitalisation and mortality. In **Chapter 4** we explore whether models typically used to predict stroke can also identify anticoagulated patients at high risk of these outcomes.

In a cohort of 2,068 anticoagulated AF patients in general practice, we described the risks of cardiac and non-cardiac hospitalisation and mortality, systematically searched for existing prediction models developed for stroke prediction in AF, and validated these models for predicting hospitalisation and mortality by calculating the observed risk per score as incidence rates (IRs) and c-statistic.

During a median follow-up of 2.7 (IQR 2.2 – 3.0) years, the IR per 100 person-years was 22.1 for hospitalisations and 6.7 for all-cause mortality. Non-cardiac events outnumbered cardiac events (IRs 15.7 versus 7.6 per 100 person-years for hospitalisation, $p < 0.001$ and 5.0 versus 1.7, $p < 0.001$ for mortality). The proportion of patients considered at 'low risk' by each model ranged from 3% (CHA2DS2-VASc)

to 35% (ATRIA). The median c-statistic of all models was 0.57 for hospitalisation outcomes, and 0.65 for mortality outcomes.

We conclude that in well-anticoagulated community-dwelling AF patients stroke risk is exceeded by risks of hospitalisation and mortality, importantly mainly for non-cardiac causes. The ATRIA model may be considered most suitable for predicting adverse events other than stroke.

Heart failure (HF) is another important adverse condition that frequently co-occurs in patients with AF, and vice versa. They share common causes and one disease may likely develop or progress in presence of the other. However, heart failure in patients with AF is often unrecognised due to overlapping symptomatology. Furthermore, AF can cause elevated natriuretic peptide levels, impairing its diagnostic value for HF-detection. In **Chapter 5** we therefore describe the prevalence of previously unknown HF in community-dwelling patients with AF, and the diagnostic value of the natriuretic peptide NTproBNP for HF-screening in AF patients.

We combined individual participant data from four HF-screening studies in older community-dwelling persons. The presence or absence of HF was in each study established by an expert panel following the criteria of the European Society of Cardiology. We performed a two-stage patient-level meta-analysis to calculate traditional diagnostic indices.

Of the 1,941 individuals included in the fur studies, 196 (10.1%) had AF at baseline. HF was uncovered in 83 (43%) of these 196 patients with AF, versus 381 (19.7%) in those without AF at baseline. Median NTproBNP levels of AF patients with and without HF were 744 pg/mL and 211 pg/mL, respectively. At the cut-point of 125 pg/mL, sensitivity was 93%, specificity 35%, and positive and negative predictive values 51% and 86%, respectively. Only 23% of all AF patients had an NTproBNP level below the 125 pg/mL cut-point, with still a 13% prevalence of HF in this group.

We concluded that the prevalence of HF is high among community-dwelling patients with AF. Given its diagnostic accuracy, screening for HF with NTproBNP is inefficient in AF patients and straightforward echocardiography should be considered.

MISCLASSIFICATION IN ROUTINE CARE DATA

Much research evaluated in this thesis, like in many other fields in medical science, relies on data collected during routine clinical practice. Therefore, finally, in **Chapter 6** of this thesis, we focus on potential misclassification in such routine

care data, specifically in studies on prognostic prediction models. Misclassification of predictors may strongly affect the associations studied. There is no doubt that this could lead to the derivation of suboptimal prediction models. The extent to which such misclassification affects the validation of existing prediction models is currently unclear.

We therefore quantified the amount of misclassification in routine care data and its effect on the validation of an existing risk prediction model in a case study using the CHA2DS2-VASc clinical prediction model for predicting mortality in patients with atrial fibrillation.

In a prospective cohort in general practice in the Netherlands, we used computerised retrieved data from the electronic medical records of patients known with AF as index predictors. Additionally, manually collected data after scrutinising all complete medical files were the reference predictors. Comparing the index with the reference predictors, we assessed misclassification in individual predictors by calculating Cohen's kappas and univariable hazard ratios (HR) for mortality. We quantified predictive performance by calculating the c-statistic and by determining calibration of multivariable models.

In total 2,363 AF patients were included. After a median follow-up of 2.7 (IQR 2.3 – 3.0) years, 368 patients died (incidence rate 6.2 deaths per 100 person-years). Misclassification in individual predictors ranged from substantially (Cohen's kappa 0.56 for prior history of heart failure) to minor (kappa 0.90 for a history of type 2 diabetes). The (HR) from univariable Cox analysis was 2.1 for heart failure using index predictors, and 1.7 with reference predictors. For type 2 diabetes HRs were 0.81 and 0.83, respectively. The overall model performance was not affected by the misclassification between index and reference predictors of the CHA2DS2-VASc model with a c-statistic of 0.684 for index and 0.681 for reference predictors, and similar calibration.

We conclude that even in the presence of substantial predictor misclassification in routine healthcare data, the overall performance of a prediction model was not negatively affected. Although our study should be repeated for other often applied prediction models, our findings suggest that routinely available health care data indeed are a useful source when validating prognostic models, despite misclassification in some of the variables.

In the **General Discussion** concluding this thesis, we focus on stroke risk prediction as a crucial part of the management of AF. We identify two pathways along which more accurate prediction of stroke may be achieved.

First, the CHA2DS2-VASc prediction model needs updating. Each predictor in the model, as well as the assigned weight, needs to be re-evaluated. Furthermore, any possible interaction terms need consideration.

Second, additional predictors may be added to the CHA2DS2-VASc model. These include clinical parameters, of which renal failure is the most promising predictors; biomarkers derived from the blood, either with a cardiac origin, associated with inflammation, or with coagulation or endothelial (dys)function; and echocardiographic markers of the size and function of the left atrium and/or left atrial appendage.

We conclude that individual patient data from existing studies, readily available data from routine clinical practice, and the growing availability of biomaterial in designated biobanks provide great opportunities for efficiently improving stroke risk prediction models. The applicability of future more sophisticated, and undoubtedly more complicated, models need careful attention in future research and clinical practice. Because ultimately, better stroke prediction should lead to better anticoagulant treatment and in the end better prevention of stroke in patients with atrial fibrillation.

REFERENCES

1 Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, Singer DE. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* 2001; **285**: 2370–5.

2 Heeringa J, van der Kuip DAM, Hofman A, Kors JA, van Herpen G, Stricker BHC, Stijnen T, Lip GYH, Witteman JCM. Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study. *Eur Heart J* Oxford University Press; 2006; **27**: 949–53.

3 Steinberg BA, Kim S, Fonarow GC, Thomas L, Ansell J, Kowey PR, Mahaffey KW, Gersh BJ, Hylek E, Naccarelli G, Go AS, Reiffel J, Chang P, Peterson ED, Piccini JP. Drivers of hospitalization for patients with atrial fibrillation: Results from the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF). *Am Heart J* 2014; **167**: 735–42. e2.

4 Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of Atrial Fibrillation on the Risk of Death. *Circulation* 1998; **98**: 946–52.

5 Odutayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG, Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. *BMJ* 2016; **354**: i4482.

6 Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. 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.

7 Ogilvie IM, Newton N, Welner SA, Cowell W, Lip GYH. Underuse of oral anticoagulants in atrial fibrillation: a systematic review. *Am J Med* 2010; **123**: 638–645e4.

8 Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of Clinical Classification Schemes for Predicting Stroke: Results From the National Registry of Atrial Fibrillation. *JAMA*; 2001; **285**: 2864–70.

Samenvatting

INLEIDING

Atriumfibrilleren (AF) is de meest voorkomende hartritmestoornis. De prevalentie in de algemene bevolking is 1–2%, en stijgt tot meer dan 15% in patiënten van 85 jaar en ouder.[1,2] AF is niet alleen een bekende risico factor voor ischaemisch CVA, het is ook geassocieerd met een verhoogd risico op hospitalisatie,[3] mortaliteit[4] en een verscheidenheid aan andere nadelige gevolgen.[5]

In dit proefschrift bespreken we de behandeling van deze risico's, in het bijzonder voor AF patiënten in de eerste lijn. We richten ons daarbij op drie domeinen: het voorspellen van ischaemisch CVA bij AF; de preventieve behandeling met antistolling; en nadelige gevolgen van AF naast CVA zoals hospitalisatie, mortaliteit en hartfalen.

HET RISICO OP CVA VOORSPELLEN BIJ ATRIUMFIBRILLEREN

Bij de preventie van CVA bij AF moeten de voordelen van de zeer effectieve anticoagulantia afgewogen worden tegen het risico op bloedingscomplicaties. Richtlijnen voor de behandeling van AF adviseren daarom om bij iedere individuele AF patiënt het risico op CVA te voorspellen, en vervolgens alleen in het geval van een hoog CVA risico anticoagulantia voor te schrijven. In **Hoofdstuk 1** van dit proefschrift evalueren we de waarde van de CHA2DS2-VASc beslisregel zoals ontwikkeld in 2010 door Lip et al.[6] Hoewel deze regel veelvuldig wordt gebruikt om het risico op CVA bij patiënten met AF te voorspellen, zijn de resultaten in externe validatie onderzoeken twijfelachtig en tegenstrijdig.

Wij deden daarom een systematische literatuur review van externe validatie onderzoeken van CHA2DS2-VASc in AF patiënten die (nog) geen anticoagulantia gebruikten. Vervolgens deden we een random effects meta-analyse van het voorspelde risico op CVA per score. Om heterogeniteit in kaart te brengen stratificeerden wij onderzoeken naar het domein waar de patiëntinclusie startte, en pasten meta-regressie toe.

In totaal werden 19 onderzoeken met een follow-up van meer dan twee miljoen persoonsjaren geëvalueerd. In onderzoeken die AF patiënten in het ziekenhuis includeerden was het CVA risico voor score nul, één en twee respectievelijk 0,4% (95% predictie interval (PI) 0,2 tot 3,2%), 1,2% (95% PI 0,1 – 3,8%) en 2,2% (95% PI 0,03 – 7,8%). Dit was consequent hoger dan onderzoeken die patiënten in de algemene bevolking includeerden, met risico's van 0,2% (95% PI 0,0 – 0,9%), 0,7% (95% PI 0,3 – 1,2%) en 1,5% (95% PI 0,4 – 3,3%) voor respectievelijk score nul tot twee. Heterogeniteit zoals uitgedrukt door het brede predictie interval kon niet volledig worden verklaard door meta-regressie.

We concluderen dat in externe validatie onderzoeken het voorspelde risico op CVA voor verschillende scores van CHA2DS2-VASc veel heterogeniteit vertoont. Besluiten rond de behandeling met anticoagulantia van patiënten met een intermediaire risico op CVA (rond score 1 tot 2) blijven daarmee ingewikkeld. Verschillen in CVA risico's tussen onderzoeken die patiënten includeerden in ziekenhuizen of in de algemene bevolking tonen aan dat mogelijke case-mix verschillen moeten worden meegewogen bij deze behandelbeslissingen, maar enige onzekerheid zal blijven bestaan.

BEHANDELING MET ANTICOAGULANTIA IN ATRIUMFIBRILLEREN

Hoewel de voordelen van anticoagulantia bij het voorkómen van CVA evident zijn, is er vaak sprake van onderbehandeling. In een meta-analyse in 2010 beschreven Ogilvie et al.[7] dat in vele onderzoeken gemiddeld 30% van de AF patiënten met een hoog risico op CVA dikwijls niet met anticoagulantia werd behandeld. In **Hoofdstuk 2** van dit proefschrift bestuderen wij daarom de redenen voor het afwijken van de richtlijn voor CVA preventie bij atriumfibrilleren in de eerste lijn.

In een observationeel cross-sectioneel onderzoek in de Nederlandse eerste lijn werd de behandeling van 440 AF patiënten in 19 huisartspraktijken geanalyseerd, en werden de redenen voor het afwijken van de CHADS2 beslisregel[8], de voorloper van CHA2DS2-VASc, voor iedere patiënt bij de huisarts uitgevraagd.

De mediane leeftijd van de geïncludeerde patiënten was 76,0 (interkwartiele range (IQR) 67 – 83) jaar, 55% was man en 62% was onder mede-behandeling van een cardioloog. Er was sprake van onderbehandeling volgens de CHADS2 en CHA2DS2-VASc beslisregel in respectievelijk 93 (21,1%) en 104 (23,6%) patiënten. Overbehandeling kwam voor in respectievelijk 84 (19,1%) en 29 (3,4%) patiënten. Als belangrijkste redenen om af te wijken van de richtlijn noemden huisartsen i) blijvend sinusritme na een periode van AF, en ii) de cardioloog werd verantwoordelijk geacht voor de behandeling met anticoagulantia. Mogelijk kunnen deze drempels worden weggenomen door wetenschappelijk onderbouwde, duidelijke aanbevelingen voor de behandeling van patiënten met (verondersteld) sinus ritme na een periode van AF, en daarnaast betere afstemming tussen huisarts en cardioloog over de verantwoordelijkheden rond het starten en vervolgen van anticoagulantia. Opvallend genoeg werden bijwerkingen of contra-indicaties voor anticoagulantia, of de voorkeur van de patiënt zelden genoemd als reden voor onderbehandeling.

Om de behandeling met anticoagulantia van patiënten met AF in de eerste lijn vervolgens te optimaliseren, verrichtten we een groot cluster-gerandomiseerd onderzoek naar de effectiviteit van een geautomatiseerde beslisondersteuning voor CVA preventie op basis van CHA2DS2-VASc, waarvan de resultaten staan beschreven in **Hoofdstuk 3**

We randomiseerden 38 huisartspraktijken. Huisartsen in de index praktijken ontvingen een automatisch opgesteld behandeladvies voor anticoagulantia op basis van CHA2DS2-VASc. Huisartsen in de referentie groep leverden gebruikelijke zorg. De primaire uitkomst was de incidentie van ischaemisch CVA, TIA en/of thrombo-embolie (TE). Secundaire uitkomsten waren bloeding en het percentage patiënten dat volgens de richtlijn met anticoagulantia werd behandeld.

In totaal werden 1.129 AF patiënten geïncludeerd in 19 index praktijken en 1.226 AF patiënten in 19 referentie praktijken. De mediane leeftijd was 77 (IQR 68 – 75) jaar, 48% was man. De mediane CHA2DS2-VASc score was 3.0 (IQR 2.0 – 5.0). Onderbehandeling met antistolling zoals geadviseerd voor patiënten met een CHA2DS2-VASc score ≥ 2 kwam voor in 6,6% van de populatie. Na een mediane follow-up van 2,7 (IQR 2,3 – 3,0) jaar was de incidentie-dichtheid ('incidence rate' IR) per 100 persoonsjaren van ischaemisch CVA/TIA/TE 1,96 in de index groep en 1,42 in de referentie groep (hazard ratio (HR) 1,3; 95% C 0,8 – 2,1). Er werd geen verschil gevonden in het risico op bloeding (HR 0,79 versus 0,82), of in onderbehandeling (7,2% versus 8,2%) of overbehandeling (8,0% versus 7,9%) met anticoagulantia.

Wij concluderen dat in deze studie in AF patiënten in de eerste lijn onderbehandeling met anticoagulantia relatief laag is. Een behandeladvies voor huisartsen op basis van CHA2DS2-VASc leidt niet tot een daling in de IR van CVA, en heeft geen invloed op het risico op bloeding, of op overbehandeling en onderbehandeling met anticoagulantia.

RISICO OP HOSPITALISATIE, MORTALITEIT EN HARTFALEN IN PATIËNTEN MET ATRIUMFIBRILLEREN

In de volgende hoofdstukken van dit proefschrift richten we ons op het verhoogde risico op andere negatieve uitkomsten dan CVA in patiënten met AF. Het belang hiervan wordt in toenemende mate onderkend, met als voorbeeld de geïntegreerde AF zorg die wordt aanbevolen in recente behandelrichtlijnen met als doel alle negatieve uitkomsten te voorkomen, inclusief hospitalisatie en mortaliteit. In **Hoofdstuk 4** onderzoeken wij of beslisregels die bedoeld zijn voor het voorspellen

van het risico op CVA, patiënten met anticoagulantia en een hoog risico op deze uitkomsten zouden kunnen identificeren.

In een cohort van 2.068 AF patiënten in de eerste lijn die anticoagulantia gebruikten beschreven wij het risico op cardiale en niet-cardiale hospitalisatie en mortaliteit, zochten wij systematisch naar bestaande beslisregels voor het voorspellen van CVA in AF, en valideerden deze regels voor het voorspellen van hospitalisatie en mortaliteit door het berekenen van het geobserveerde risico per score als incidentie dichtheid en de c-statistic.

Gedurende een mediane follow-up van 2,7 (IQR 2,2 – 3,0) jaar was de IR per 100 persoonsjaren 22,1 voor hospitalisatie en 6,7 voor mortaliteit ongeacht de oorzaak. Niet-cardiale uitkomsten kwamen vaker voor dan cardiale uitkomsten (IR 15,7 versus 7,6 per 100 persoonsjaren voor hospitalisatie, $p < 0{,}001$ en 5,0 versus 1,7, $p < 0{,}001$ voor mortaliteit). Het percentage patiënten dat door elk van de beslisregels werd beschouwd als 'laag risico' varieerde van 3% (CHA2DS2-VASc) tot 35% (ATRIA). De mediane c-statistic van alle modellen was 0,57 voor hospitalisatie en 0,65 voor mortaliteit.

Wij concluderen dat in AF patiënten die anticoagulantia gebruiken in de eerste lijn, het risico op hospitalisatie en mortaliteit het risico op CVA overstijgt, voornamelijk vanwege niet-cardiale oorzaken. De ATRIA beslisregel zou overwogen kunnen worden om andere negatieve uitkomsten dan CVA te voorspellen.

Ook hartfalen (HF) is een belangrijke aandoening die vaak voorkomt onder patiënten met AF, en vice versa. Ze delen dezelfde oorzaken en de ene aandoening kan ontstaan of verergeren in aanwezigheid van de ander. Door overlappende symptomen wordt HF echter vaak niet herkend in patiënten met AF. Daarnaast zorgt AF voor verhoging van natriuretische peptiden in het bloed waarmee de diagnostische waarde hiervan wordt beperkt. In **Hoofdstuk 5** beschrijven wij daarom de prevalentie van voorheen onbekend HF in AF patiënten in de eerste lijn, en de diagnostische waarde van de natriuretische peptide NTproBNP voor HF screening in AF patiënten.

Wij combineerden individuele patiënten data van vier onderzoeken naar HF screening onder oudere patiënten in de eerste lijn. De aan- of afwezigheid van HF werd in elke studie vastgesteld door een panel van experts aan de hand van de criteria van de European Society of Cardiology. Wij verrichtten een meta-analyse van data op individueel patiëntniveau in twee stappen om de gebruikelijke diagnostische uitkomstmaten te berekenen.

Van de 1.941 patiënten uit de vier studies, hadden 196 (10,1%) AF bij de baseline meting. HF werd vastgesteld bij 83 (43%) van deze 196 patiënten met AF, versus bij 381 (19,7%) patiënten zonder AF. De mediane NTproBNP waarden in AF patiënten met en zonder HF waren respectievelijk 744 pg/mL en 211 pg/mL. Slechts 23%

van alle AF patiënten had een NTproBNP waarde lager dan het afkappunt 125 pg/mL, en 13% in deze groep had evengoed HF.

We concluderen dat de prevalentie van HF in AF patiënten in de eerste lijn hoog is. Gezien de diagnostische waarde is het gebruik van NTproBNP voor HF screening bij patiënten met AF niet efficiënt en zou echocardiografie overwogen moeten worden.

MISCLASSIFICATIE IN ROUTINE ZORG DATA

Veel onderzoek dat in dit proefschrift aan bod komt, net als veel onderzoek in andere sectoren in de medische wetenschap, maakt gebruik van gegevens die tijdens routinematige gezondheidszorg zijn verzameld. Daarom richten wij ons tenslotte in **Hoofdstuk 6** op mogelijke misclassificatie in deze routine zorg data, in het bijzonder in onderzoek naar prognostische beslisregels. Misclassificatie in predictoren kan onderzoeksresultaten sterk beïnvloeden. Dit leidt ongetwijfeld tot de ontwikkeling van suboptimale beslisregels, maar in hoeverre misclassificatie de validatie van een bestaande beslisregel beïnvloedt is vooralsnog onduidelijk.

Wij kwantificeerden daarom de hoeveelheid misclassificatie in routine zorg data en het effect ervan op de validatie van een bestaande beslisregel, met als voorbeeld de CHA2DS2-VASc klinische beslisregel voor het voorspellen van mortaliteit in patiënten met atriumfibrilleren. In een prospectief cohort in de Nederlandse eerste lijn gebruikten we automatisch verkregen gegevens uit het elektronisch dossier van patiënten met AF als index predictoren. Handmatig verzamelde data na het bestuderen van het complete medisch dossier dienden vervolgens als referentie predictoren. We beoordeelden misclassificatie door het vergelijken van de index en referentie predictoren met het uitrekenen van Cohen's kappa en univariabele hazard ratios voor mortaliteit. We kwantificeerden de voorspellende waarde van multi-variabele modellen door het uitrekenen van de c-statistic en de calibratie. In totaal werden 2.363 AF patiënten geïncludeerd. Gedurende een mediane follow-up van 2,7 (IQR 2,3 – 3,0) jaar overleden 368 patiënten (IR 6,2 per 100 persoons-jaren). Misclassificatie in individuele predictoren varieerde van aanzienlijk (Cohen's kappa 0,56 voor een voorgeschiedenis van hartfalen) tot beperkt (kappa 0,90 voor een voorgeschiedenis van diabetes type 2). De HR uit univariabele Cox analyse was 2,1 voor hartfalen gebruik makend van index predictoren, en 1,7 gebruik makend van referentie predictoren. Voor diabetes type 2 waren de HRs respectievelijk 0,81 en 0,83. De voorspellende waarde van de volledige CHA2DS2-VASc beslisregel werd niet beïnvloed door misclassificatie, met een c-statistic van 0.684 voor index predictoren en van 0,681 voor referentie predictoren, met gelijke calibratie.

We concluderen dat zelfs bij aanzienlijke misclassificatie in predictoren in routine zorg data, de waarde van een beslisregel niet negatief wordt beïnvloed. Hoewel ons onderzoek herhaald zal moeten worden met andere veelgebruikte beslisregels lijkt op basis van onze resultaten dat ongeacht misclassificatie in sommige variabelen, routine zorg data een waardevolle bron zijn bij het valideren van klinische beslisregels.

De **General Discussion** waarmee dit proefschrift afsluit, behandelt het voorspellen van het risico op CVA als cruciaal onderdeel van de behandeling van AF. We bediscussiëren twee trajecten die kunnen leiden tot het nauwkeuriger voorspellen van CVA.

Ten eerste zal de CHA2DS2-VASc beslisregel herzien moeten worden. Elke predictor in de beslisregel, alsmede het toegekend aantal punten moet opnieuw worden geëvalueerd. Daarnaast moeten mogelijke interactietermen worden overwegen.

Ten tweede zouden nieuwe predictoren aan CHA2DS2-VASc toegevoegd kunnen worden. Hieronder vallen klinische kenmerken, waarvan nierinsufficiëntie de meest veelbelovende is; biomarkers die gemeten kunnen worden in het bloed, afkomstig uit het hart, of gerelateerd aan ontsteking, coagulatie of endotheel (dys) functie; en echocardiografische metingen van de grootte en functie van het linker atrium en/of het linker hartoor.

We concluderen dat individuele patiënten data uit bestaande onderzoeken, ruimschoots voorhanden zijnde routine zorg data, en de toenemende beschikbaarheid van biomateriaal in speciale biobanken uitgelezen mogelijkheden bieden om het voorspellen van CVA efficiënt te verbeteren. De toepasbaarheid van toekomstige geavanceerde, en daardoor ongetwijfeld complexere, beslisregels verdient verdere aandacht in onderzoek en in de klinische praktijk. Uiteindelijk moet namelijk een betere voorspelling van het risico op CVA leiden ook tot betere behandeling met anticoagulantia en ten slotte een betere preventie van CVA bij patiënten met atriumfibrilleren.

REFERENTIES

1 Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, Singer DE. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* 2001; 285: 2370–5.

2 Heeringa J, van der Kuip DAM, Hofman A, Kors JA, van Herpen G, Stricker BHC, Stijnen T, Lip GYH, Witteman JCM. Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study. *Eur Heart J* 2006; 27: 949–53.

3 Steinberg BA, Kim S, Fonarow GC, Thomas L, Ansell J, Kowey PR, Mahaffey KW, Gersh BJ, Hylek E, Naccarelli G, Go AS, Reiffel J, Chang P, Peterson ED, Piccini JP. Drivers of hospitalization for patients with atrial fibrillation: Results from the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF). *Am Heart J* 2014; 167: 735–42. e2.

4 Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of Atrial Fibrillation on the Risk of Death. *Circulation* 1998; 98: 946–52.

5 Odutayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG, Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. *BMJ* 2016; 354: i4482.

6 Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. 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.

7 Ogilvie IM, Newton N, Welner SA, Cowell W, Lip GYH. Underuse of oral anticoagulants in atrial fibrillation: a systematic review. *Am J Med* 2010; 123: 638–645e4.

8 Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of Clinical Classification Schemes for Predicting Stroke: Results From the National Registry of Atrial Fibrillation. *JAMA* 2001; 285: 2864–70.

Dankwoord

Dit proefschrift had niet kunnen bestaan zonder de mensen die mij de afgelopen jaren geweldig hebben geholpen, of er zijn geweest op de juiste tijd en op de juiste plaats.

Prof.dr. Moons, geachte promotor, beste Carl. Jouw enthousiasme is even oneindig als aanstekelijk. Dank voor al het hardop denken, je goeie ideeën en je krachtige besluiten. Een promovendus, en misschien wel helemaal een verlegen huisarts als ik, kan zich in een promotor niets beters wensen dan jouw passie, ambitie en academische vaardigheden.

Prof.dr. Hoes, geachte promotor, beste Arno. De momenten dat wij elkaar zagen waren spaarzaam, maar voor jou zondermeer genoeg om het onderzoek direct naar een hoger plan te tillen. Ik bewonder hoe jij jouw aandacht over zoveel zaken verdeelt, en op de cruciale momenten toch precies tot de kern van elk onderzoek kunt doordringen. Dank voor je scherpe blik, je feedback en je humor.

Dr. Rutten, geachte co-promotor, beste Frans. Zonder jouw inhoudelijke kennis was dit onderzoek nooit op stoom geraakt, en zonder jouw relativeringsvermogen was het allang uit de bocht gevlogen. Dank voor de tijd die jij altijd nam, de lange gesprekken, alertheid en kritische geest.

Geachte dr. Geersing, lieve Geert-Jan. Vanaf het moment dat we elkaar ontmoetten in jouw praktijk in Amsterdam, heb jij mij meegenomen langs het pad van de wetenschap. Jij wees mij niet alleen de gebaande paden, maar toonde ook mooie huisartsgeneeskundige kronkelwegen en wetenschappelijke vergezichten. Dank voor alle steun en vertrouwen, voor al het geduld en inspiratie. Ik ben blij dat dit proefschrift maar een tussenstop is tijdens onze samenwerking, en dat het eindpunt nog even niet in zicht is.

Geachte dr. Oudega, beste Ruud. Tijdens de begindagen van CAFe ben je van onmetelijke waarde geweest. Dank voor je onvermoeibare talent zaken van de grond te krijgen en aan de gang te houden.

Geachte prof.dr. Schutgens, prof.dr. Damoiseaux, prof.dr. Kapelle, prof.dr. van Gelder, prof.dr. Huisman, beste leden van de beoordelingscommissie. Veel dank voor het kritisch lezen en beoordelen van mijn proefschrift.

Veel dank aan de co-auteurs van de artikelen in dit proefschrift. Beste Floor Hartman-Weide, Femke Kaasenbrood, Rogier Kievit, Caitriona O'Flynn en Timo

Brakenhoff, geachte dr. van Mourik, dr. Bertens, dr. van Riet en dr. Boonman-Winter, dank voor de samenwerking. Geachte dr. Debray, beste Thomas, jouw hulp en begeleiding bij het schrijven van ons artikel was van onmetelijke waarde. Ik heb heel veel van je kunnen leren, en veel aan je mogen vragen wanneer de statistiek mij (weer) te ingewikkeld werd.

Het CAFe onderzoek dankt haar succes aan de geweldige medewerking van de deelnemende huisartsen. Veel dank aan de huisartspraktijken Aletta, Bilthoven, Binnenstad, Bonekamp & van der Meer, van Bodegom, Bosboomstraat, Brouwerd & van de Sanden, Carnegiedreef, de Angstel, de Bilt, de Brug, de Dame, Detmoldstraat, Eedenburgh, de Egelantier, Galecop, Grijzenhout-Ballieux, van Hijfte, de Julius Gezondheidscentra, Kievitstraat, de Lichtenberch, Maliesingel, Mariahoek, Nije Veste, Oog in Al, Overvecht, de Reiger, Tuindorp, Tuinwijk en Tuna-Tüzün.

Veel dank aan alle stagiairs die hebben meegeholpen aan het CAFe onderzoek. Beste Ceren, Floor, Ilse, Maartje, Maria, Sophie en Stephanie, dank voor alle inzet. Beste Annerien, je hebt een ongelooflijke prestatie geleverd bij de dataverzameling. Je was absoluut de juiste persoon op de juiste tijd, dank voor je geweldige inspanningen.

Beste Julia Velikopolskaia en Erwin Stehouwer, datamanagement UMC Utrecht, en Hugo Smeets, JHN, dank voor al jullie hulp tijdens de technische en organisatorische hobbels die we samen hebben moeten nemen. Ik kon telkens weer bij jullie terecht in tijden van nood en er was telkens weer energie om aan een oplossing te werken.

Geachte prof.dr. Scholten, beste Rob, geachte dr. Hooft, beste Lotti. Met veel plezier heb ik met jullie samengewerkt tijdens de zoektocht naar de zin (en onzin) van diagnostische testen bij pijn op de borst. Ik heb er ontzettend van geleerd, sterke verhalen gehoord en gelachen. Ik hoop dat we in de toekomst nog eens samen kunnen werken.

Lieve Andrea, Ineke, Karin en Peter; lieve assistentes, praktijkondersteuners en alle andere collega's bij Gezondheidscentra Maarssenbroek, dank voor alle steun en jullie enorme flexibiliteit de afgelopen jaren. Ook op de momenten dat ik Spechtenkamp haastig moest verlaten voor het onderzoek bleven jullie geïnteresseerd en geduldig. We zijn een prachtig team samen.

Dankjewel 'VTE' collega's Henrike, Janneke, Annelieke, Carline, Linda en Rosanne, en dankjewel kamergenoten in het Julius Centrum voor de samenwerking én de afleiding. Beste Faas, dank voor alle mooie jaren die we samen hebben doorgebracht tijdens de opleiding tot huisarts, de master epidemiologie en het onderzoek. Ik ben zeer vereerd dat jij mij als paranimf wilt bijstaan bij de laatste loodjes van mijn promotie.

Amber, Annegriet, Judith en Leonard, lieve collega's, dank voor de reflectie en de goeie vragen tijdens onze intervisie samen. Corine, dank voor al het brainstormen.

Lieve Chris, in voor- en tegenspoed ben jij een veilige haven in Utrecht, dankjewel. Lieve Mark, dank voor jouw interesse en onweerstaanbare wijsheid. Lieve Sjoert, geen ritme is ons te onregelmatig. Ook vanuit de verte, dankjewel. Lieve Arend, goed eten en goede muziek in goed gezelschap leidt tot goede gedachten. Dank voor je inspiratie.

Lieve Zorah, wat kennen we elkaar kort, maar wat kennen we elkaar goed. Dankjewel voor jouw mooie blik op de wereld, en al je geduld en steun tijdens de laatste fase van mijn promotie.

Tot slot gaat mijn allergrootste dank uit naar mijn familie. Lieve Annelies, jouw stijl en jouw verwondering vormen schitterende inspiratie. Dank voor alle geweldige momenten samen. Lieve Marjolein, mijn hele leven heb ik op jou mogen steunen als grote zus. Altijd was je er voor advies, of voor een luisterend oor. Ik kan me daarom geen betere paranimf voorstellen. Dankjewel voor je talent en je hulp.

Lieve ouders, papa en mama, dank voor jullie onvoorwaardelijke liefde en betrokkenheid. Dankzij de kansen en het vertrouwen dat ik van jullie heb gekregen ligt er nu dan toch echt een proefschrift. Zonder jullie steun in goede en in slechte tijden was het er nooit van gekomen. Dankjewel.

About the author

Sander van Doorn was born in Amsterdam on May 22nd, 1984. He received his medical training at the VU University Medical Center Amsterdam during which he followed the Honors Program, conducting research at the department of Adolescent and Child Psychiatry under the supervision of prof. dr. Th.A.H. Doreleijers.

In 2009-2010, Sander worked as a resident in Cardiology at Noordwest Ziekenhuisgroep (formerly Medisch Centrum Alkmaar) and in Psychiatry at Arkin, Amsterdam.

In 2010 he started his training in General Practice (GP) at the Julius Center for Health Sciences and Primary Care of the University Medical Center Utrecht, the Netherlands. From 2012 he combined this training with a PhD project on risk management in atrial fibrillation, supervised by prof. dr. K.G.M. Moons and prof. dr. A.W. Hoes. In 2015 he received a master's degree in Clinical Epidemiology at Utrecht University.

Currently Sander lives in Amsterdam and works as a GP at Gezondheidscentra Maarssenbroek. He spends his free time making music and fermenting foods.