BMJ

RESEARCH

Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study

Jonas Bjerring Olesen, research fellow,[1] Gregory Y H Lip, professor,[2] Morten Lock Hansen, research fellow,[1] Peter Riis Hansen, research director,[1] Janne Schurmann Tolstrup, research director,[3] Jesper Lindhardsen, research fellow,[1] Christian Selmer, research fellow,[1] Ole Ahlehoff, research fellow,[1] Anne-Marie Schjerning Olsen, research fellow,[1] Gunnar Hilmar Gislason, research director,[1] Christian Torp-Pedersen, professor[1]

[1]Department of Cardiology, Copenhagen University Hospital Gentofte, 2900 Hellerup, Denmark
[2]University of Birmingham Centre for Cardiovascular Sciences, City Hospital, Birmingham B18 7QH, UK
[3]National Institute of Public Health, 1399 Copenhagen K, Denmark

Correspondence to: J Olesen
jo@heart.dk

Cite this as: *BMJ* 2011;342:d124
doi:10.1136/bmj.d124

ABSTRACT

Objectives To evaluate the individual risk factors composing the $CHADS_2$ (Congestive heart failure, Hypertension, Age≥75 years, Diabetes, previous Stroke) score and the CHA_2DS_2-VASc (CHA_2DS_2-Vascular disease, Age 65-74 years, Sex category) score and to calculate the capability of the schemes to predict thromboembolism.

Design Registry based cohort study.

Setting Nationwide data on patients admitted to hospital with atrial fibrillation.

Population All patients with atrial fibrillation not treated with vitamin K antagonists in Denmark in the period 1997-2006.

Main outcome measures Stroke and thromboembolism.

Results Of 121 280 patients with non-valvular atrial fibrillation, 73 538 (60.6%) fulfilled the study inclusion criteria. In patients at "low risk" (score=0), the rate of thromboembolism per 100 person years was 1.67 (95% confidence interval 1.47 to 1.89) with $CHADS_2$ and 0.78 (0.58 to 1.04) with CHA_2DS_2-VASc at one year's follow-up. In patients at "intermediate risk" (score=1), this rate was 4.75 (4.45 to 5.07) with $CHADS_2$ and 2.01 (1.70 to 2.36) with CHA_2DS_2-VASc. The rate of thromboembolism depended on the individual risk factors composing the scores, and both schemes underestimated the risk associated with previous thromboembolic events. When patients were categorised into low, intermediate, and high risk groups, C statistics at 10 years' follow-up were 0.812 (0.796 to 0.827) with $CHADS_2$ and 0.888 (0.875 to 0.900) with CHA_2DS_2-VASc.

Conclusions The risk associated with a specific risk stratification score depended on the risk factors composing the score. CHA_2DS_2-VASc performed better than $CHADS_2$ in predicting patients at high risk, and those categorised as low risk by CHA_2DS_2-VASc were truly at low risk for thromboembolism.

INTRODUCTION

Patients with atrial fibrillation have a substantial risk of stroke, which is modified by the presence or absence of several risk factors.[1 2] These risk factors have been used to develop thromboembolic risk stratification schemes, which have somewhat arbitrarily divided the risk of thromboembolism into low, intermediate, and high risk strata.[3] Given the limitations of oral anticoagulation treatment with vitamin K antagonists, such risk stratification allows clinicians to target patients at "high risk" for treatment with vitamin K antagonists. For the intermediate risk category, guidelines recommend treatment with vitamin K antagonists or aspirin, and aspirin is recommended for the low risk category.

Schemes for stratifying the risk of stroke have been largely derived from non-anticoagulated arms of clinical trial cohorts, in which many potential thromboembolic risk factors were not recorded. In these historical trials, less than 10% of patients screened were randomised, and over the past 15-20 years the evolution of risk schemes has not improved their predictive value for patients at high risk.[4] More recent data in patients at intermediate risk show that vitamin K antagonists are superior to aspirin in reducing the risk of thromboembolism and adverse events,[5-7] and aspirin does not reduce the risk of thromboembolism in atrial fibrillation patients at "low risk".[8] Thus, a paradigm shift has been proposed whereby greater efforts are made to identify "truly low risk" patients who may not need any antithrombotic treatment, whereas all others could be considered for oral anticoagulation.[8-10]

The most commonly used scheme for stratifying the risk of stroke is the $CHADS_2$ (Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, previous Stroke/transient ischaemic attack (doubled risk weight)) score.[11] Various limitations of this score have been discussed, including classification of a large proportion of patients as being at "intermediate risk" and its omission of many potential thromboembolic risk factors.[10] The 2006 ACC/AHA/ESC guideline listed these potential additional risk factors as being "less validated or weaker risk factors," including female sex, age 65-74 years, coronary artery disease, and

Table 1 | Baseline characteristics of patients. Values are numbers (percentages)

Characteristics	Patients discharged with non-valvular AF (n=121 280)	Patients who survived 7 days (n=118 243)	Patients who did not receive VKA or heparin (study population) (n=73 538)
Heart failure	22 759 (18.8)	22 082 (18.7)	13 126 (17.9)
Hypertension	48 171 (39.7)	47 224 (39.9)	25 060 (34.1)
Age ≥75 years	65 512 (54.0)	63 292 (53.5)	43 864 (59.7)
Age 65-74 years	29 367 (24.2)	28 817 (24.4)	14 544 (19.8)
Diabetes mellitus	11 072 (9.1)	10 754 (9.1)	6 496 (8.8)
Previous thromboembolism*	23 528 (19.4)	22 291 (18.9)	13 368 (18.2)
Vascular disease	20 305 (16.7)	19 568 (16.6)	12 873 (17.5)
Female sex	56 490 (46.6)	54 869 (46.4)	37 651 (51.2)
$CHADS_2$ score:			
0	26 139 (21.6)	25 863 (21.9)	16 406 (22.3)
1	38 024 (31.4)	37 225 (31.5)	23 730 (32.3)
2	28 249 (23.3)	27 540 (23.3)	16 393 (22.3)
3	18 198 (15.0)	17 477 (14.8)	10 846 (14.8)
4	8 178 (6.7)	7 760 (6.6)	4 745 (6.5)
5	2 210 (1.8)	2 110 (1.8)	1 260 (1.7)
6	282 (0.2)	268 (0.2)	158 (0.2)
CHA_2DS_2-VASc score:			
0	10 125 (8.4)	10 065 (8.5)	6 369 (8.7)
1	14 526 (12.0)	14 376 (12.2)	8 203 (11.2)
2	22 115 (18.2)	21 726 (18.4)	12 771 (17.4)
3	27 834 (23.0)	27 152 (23.0)	17 371 (23.6)
4	22 676 (18.7)	21 995 (18.6)	13 887 (18.9)
5	14 213 (11.7)	13 639 (11.5)	8 942 (12.2)
6	6 927 (5.7)	6 586 (5.6)	4 244 (5.8)
7	2 327 (1.9)	2 194 (1.9)	1 420 (1.9)
8	467 (0.4)	443 (0.4)	285 (0.4)
9	70 (0.1)	67 (0.1)	46 (0.1)
Drugs:			
α adrenergic blocker	1 729 (1.4)	1 681 (1.4)	1 005 (1.4)
Non-loop diuretic	37 292 (30.8)	36 319 (30.7)	21 695 (29.5)
Vasodilator	3 769 (3.1)	3 659 (3.1)	2 329 (3.2)
β blocker	50 370 (41.5)	49 611 (42.0)	26 160 (35.6)
Calcium channel blocker	35 235 (29.1)	34 539 (29.2)	18 966 (25.8)
Renin-angiotensin system inhibitor	33 445 (27.6)	32 731 (27.7)	16 868 (22.9)
Loop diuretic	47 676 (39.3)	46 340 (39.2)	27 602 (37.5)
Statin	13 629 (11.2)	13 372 (11.3)	6 919 (9.4)
Antiplatelet drug	38 007 (31.3)	37 047 (31.3)	25 503 (34.7)
Digoxin	60 661 (50.0)	59 547 (50.4)	31 418 (42.7)
Amiodarone	3 879 (3.2)	3 825 (3.2)	1 874 (2.6)

AF=atrial fibrillation; VKA=vitamin K antagonist.
*Includes peripheral artery embolism, transient ischaemic attack, ischaemic stroke, and pulmonary embolism.

thyrotoxicosis.[12] Since 2006, stronger evidence has accumulated that these additional risk factors (with the exception of thyrotoxicosis) should be considered in assessing thromboembolic risk and would be of value in identifying those patients at truly low risk.[10 13] The additional risk factors have been expressed in the CHA_2DS_2-VASc (Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, previous Stroke/transient ischaemic attack, Vascular disease, Age 65-74 years, Sex category; age≥75 years and previous stroke carry doubled risk weight) score, which has been proposed to complement the $CHADS_2$ score.[13] In the original validation study from the Euro-Heart survey, CHA_2DS_2-VASc generally had a similar C statistic to $CHADS_2$ but was better at identifying the patients at truly low risk and categorised only a small proportion into the intermediate risk category.[13] In a further study in a small elderly "real world" cohort with anticoagulated atrial fibrillation, the $CHADS_2$ and CHA_2DS_2-VASc had similar strength (C statistics) for predicting thromboembolism.[14]

An ideal validation cohort for a thromboembolic risk scheme would be a large real world cohort of patients with atrial fibrillation, without any use of anticoagulation treatment. In Denmark, the national patient registry allows such an analysis in a large cohort of real world patients, and the first objective of the analysis reported here was to assess the effects of the individual factors of $CHADS_2$ and CHA_2DS_2-VASc on the risk of thromboembolism. Secondly, we evaluated the predictive capability of $CHADS_2$ and CHA_2DS_2-VASc for thromboembolism.

METHODS

Registry data sources

In Denmark, all citizens have a permanent and personal registration number that enables linkage of the nationwide registries at an individual level. Since 1978, all admissions from Danish hospitals have been registered in the Danish national patient registry with one primary discharge diagnosis and, if appropriate, one or more secondary discharge diagnoses according to ICD-8 (international classification of diseases, 8th revision) up to 1993 and the ICD-10 from 1994 onwards.[15] From 1996, invasive therapeutic procedures (such as surgery and percutaneous interventions) have been coded according to the Nordic Medical Statistics Committees Classification of Surgical Procedures. Since 1995, all prescriptions dispensed from Danish pharmacies have been accurately registered in the Danish registry of medicinal product statistics (prescription registry) according to the international anatomical therapeutic chemical classification system.[16] The civil registration system holds information on vital status for all citizens, and the national causes of death registry holds information on primary and contributing causes of death.

Study population

From the national patient registry, we identified all patients with non-valvular atrial fibrillation or atrial flutter in the period 1997-2006. We defined non-valvular atrial fibrillation by a discharge diagnosis of atrial fibrillation or atrial flutter (diagnosis code I48), no previous diagnoses of mitral or aortic valve disease (394-396, 4240, 4241, I05, I06, I34, I35), and no mitral or aortic valve surgery (surgical procedure codes KFK, KFM, KFP), as done previously.[17] Because drug treatment may be changed or intensified in relation to hospital admission, we started follow-up seven days after discharge. We excluded patients if they died or had a

Table 2 | Event rate (95% CI) of hospital admission and death due to thromboembolism* per 100 person years

Score/risk category	1 year's follow-up	5 years' follow-up	10 years' follow-up
$CHADS_2$:			
0	1.67 (1.47 to 1.89)	1.28 (1.19 to 1.38)	1.24 (1.16 to 1.33)
1	4.75 (4.45 to 5.07)	3.70 (3.55 to 3.86)	3.56 (3.42 to 3.70)
2	7.34 (6.88 to 7.82)	5.58 (5.35 to 5.83)	5.40 (5.18 to 5.63)
3	15.47 (14.62 to 16.36)	10.29 (9.87 to 10.73)	9.89 (9.50 to 10.31)
4	21.55 (20.03 to 23.18)	14.00 (13.22 to 14.82)	13.70 (12.95 to 14.48)
5	19.71 (16.93 to 22.93)	12.98 (11.52 to 14.63)	12.57 (11.18 to 14.14)
6	22.36 (14.58 to 34.30)	16.75 (11.91 to 23.56)	17.17 (12.33 to 23.92)
$CHADS_2$:			
Low risk (0)	1.67 (1.47 to 1.89)	1.28 (1.19 to 1.38)	1.24 (1.16 to 1.33)
Intermediate risk (1)	4.75 (4.45 to 5.07)	3.70 (3.55 to 3.86)	3.56 (3.42 to 3.70)
High risk (2-6)	12.27 (11.84 to 12.71)	8.30 (8.08 to 8.51)	7.97 (7.77 to 8.17)
CHA_2DS_2-VASc:			
0	0.78 (0.58 to 1.04)	0.69 (0.59 to 0.81)	0.66 (0.57 to 0.76)
1	2.01 (1.70 to 2.36)	1.51 (1.37 to 1.67)	1.45 (1.32 to 1.58)
2	3.71 (3.36 to 4.09)	3.01 (2.83 to 3.20)	2.92 (2.76 to 3.09)
3	5.92 (5.53 to 6.34)	4.41 (4.21 to 4.61)	4.28 (4.10 to 4.47)
4	9.27 (8.71 to 9.86)	6.69 (6.41 to 6.99)	6.46 (6.20 to 6.74)
5	15.26 (14.35 to 16.24)	10.42 (9.95 to 10.91)	9.97 (9.53 to 10.43)
6	19.74 (18.21 to 21.41)	12.85 (12.07 to 13.69)	12.52 (11.78 to 13.31)
7	21.50 (18.75 to 24.64)	13.92 (12.49 to 15.51)	13.96 (12.57 to 15.51)
8	22.38 (16.29 to 30.76)	14.07 (10.80 to 18.33)	14.10 (10.90 to 18.23)
9	23.64 (10.62 to 52.61)	16.08 (8.04 to 32.15)	15.89 (7.95 to 31.78)
CHA_2DS_2-VASc:			
Low risk (0)	0.78 (0.58 to 1.04)	0.69 (0.59 to 0.81)	0.66 (0.57 to 0.76)
Intermediate risk (1)	2.01 (1.70 to 2.36)	1.51 (1.37 to 1.67)	1.45 (1.32 to 1.58)
High risk (2-9)	8.82 (8.55 to 9.09)	6.01 (5.88 to 6.14)	5.72 (5.60 to 5.84)

*Includes peripheral artery embolism, ischaemic stroke, and pulmonary embolism.

thromboembolism in this seven day quarantine period. We identified drug treatment status from prescription claims from 180 days before discharge to seven days after discharge, and we excluded patients if they had received vitamin K antagonists (medicine code B01AA) or heparins (B01AB) (fig 1). We censored patients at time of death or at the end of the follow-up periods—that is, at one, five, and 10 years.

Covariates of $CHADS_2$ and CHA_2DS_2-VASc

We identified patients with congestive heart failure from the combination of a previous diagnosis of heart failure (425, 4270, 4271, I110, I42, I50, J819) in the national patient registry and treatment with loop diuretics (C03C).[18] We identified patients with hypertension from combination treatment with at least two of the following classes of antihypertensive drugs: α adrenergic blockers (C02A, C02B, C02C), non-loop diuretics (C02DA, C02L, C03A, C03B, C03D, C03E, C03X, C07C, C07D, C08G, C09BA, C09DA, C09XA52), vasodilators (C02DB, C02DD, C02DG, C04, C05), β blockers (C07), calcium channel blockers (C07F, C08, C09BB, C09DB), and renin-angiotensin system inhibitors (C09). This definition of hypertension was validated in a previously described randomly selected cohort of people from the Danish population aged 16 years and older.[19] Of the 14 994 people in this cohort, 2028 reported having taken drugs for hypertension within a two week period before the interview. The positive predictive value of treatment with two classes of antihypertensive drugs to predict hypertension was 80.0%, and the specificity was 94.7%. We defined diabetes mellitus as a claimed prescription for a glucose lowering drug (A10). Information on previous thromboembolism—that is, peripheral artery embolism, stroke, transient ischaemic attack, and pulmonary embolism (433-438, 444, 450, G458, G459, I26, I63, I64, I74)—came from the national patient registry (from 1978), as did information on previous vascular disease—that is, myocardial infarction, peripheral artery disease, and aortic plaque (410, 440, I21, I22, I700, I702-I709), as defined by Lip and colleagues.[13 20-23]

The $CHADS_2$ score was the sum of points obtained after addition of one point each for heart failure, hypertension, age≥75, and diabetes and two points for previous thromboembolism. This score thus ranged from 0 to 6.[11] The CHA_2DS_2-VASc score was the sum of points after addition of one point each for heart failure, hypertension, diabetes, vascular disease, age 65-74 years, and female sex and two points each for previous thromboembolism and age≥75 years. This score thus ranged from 0 to 9.[13] In both risk schemes, we considered a score of 0 to represent low risk, 1 to represent intermediate risk, and ≥2 to represent high risk of thromboembolism.

Outcomes

The primary study end point was admission to hospital with or death from thromboembolism—that is, peripheral artery embolism, ischaemic stroke, or pulmonary embolism (I26, I63, I64, I74), as defined by Lip and colleagues.[13] We also did a sensitivity analysis confining the primary study end point to peripheral artery

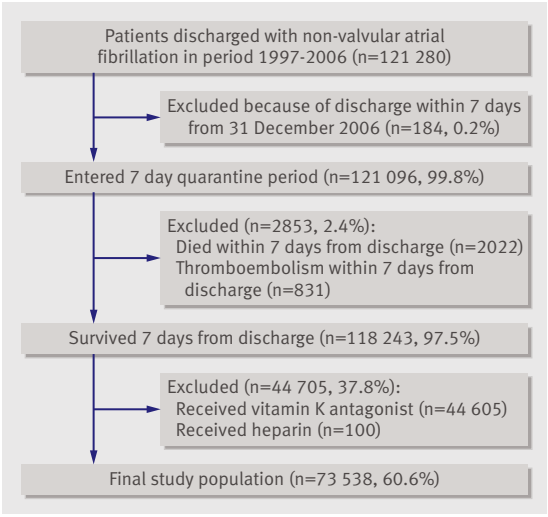

Fig 1 | Flow chart of study population

Table 3 | All cause mortality rate (95% CI) per 100 person years

Score/risk category	1 year's follow-up	5 years' follow-up	10 years' follow-up
$CHADS_2$:			
0	9.33 (8.85 to 9.84)	5.14 (4.95 to 5.33)	4.70 (4.54 to 4.86)
1	24.50 (23.82 to 25.20)	16.61 (16.29 to 16.93)	15.93 (15.65 to 16.22)
2	29.21 (28.30 to 30.14)	21.10 (20.65 to 21.56)	20.57 (20.16 to 20.99)
3	39.41 (38.08 to 40.78)	28.42 (27.74 to 29.11)	27.90 (27.27 to 28.55)
4	43.61 (41.50 to 45.84)	32.69 (31.56 to 33.86)	32.30 (31.22 to 33.41)
5	53.68 (49.10 to 58.70)	38.92 (36.44 to 41.57)	38.90 (36.50 to 41.45)
6	83.30 (67.17 to 103.29)	53.45 (44.56 to 64.12)	51.47 (43.04 to 61.55)
$CHADS_2$:			
Low risk (0)	9.33 (8.85 to 9.84)	5.14 (4.95 to 5.33)	4.70 (4.54 to 4.86)
Intermediate risk (1)	24.50 (23.82 to 25.20)	16.61 (16.29 to 16.93)	15.93 (15.65 to 16.22)
High risk (2-6)	35.47 (34.75 to 36.20)	25.46 (25.11 to 25.83)	24.87 (24.53 to 25.21)
CHA_2DS_2-VASc:			
0	4.85 (4.31 to 5.45)	2.56 (2.36 to 2.78)	2.29 (2.12 to 2.47)
1	10.32 (9.61 to 11.08)	5.81 (5.52 to 6.10)	5.33 (5.09 to 5.58)
2	21.17 (20.31 to 22.05)	13.65 (13.27 to 14.04)	12.93 (12.59 to 13.27)
3	27.06 (26.22 to 27.93)	19.11 (18.71 to 19.52)	18.52 (18.16 to 18.89)
4	31.29 (30.27 to 32.35)	22.67 (22.17 to 23.20)	22.23 (21.76 to 22.71)
5	39.45 (37.99 to 40.97)	28.50 (27.75 to 29.27)	28.16 (27.45 to 28.88)
6	44.96 (42.67 to 47.36)	33.00 (31.79 to 34.26)	32.52 (31.37 to 33.71)
7	51.12 (46.91 to 55.70)	37.71 (35.43 to 40.14)	37.98 (35.76 to 40.33)
8	77.74 (65.91 to 91.69)	50.31 (44.07 to 57.43)	48.98 (43.02 to 55.77)
9	105.51 (72.85 to 152.81)	65.77 (47.00 to 92.05)	62.40 (44.59 to 87.33)
CHA_2DS_2-VASc:			
Low risk (0)	4.85 (4.31 to 5.45)	2.56 (2.36 to 2.78)	2.29 (2.12 to 2.47)
Intermediate risk (1)	10.32 (9.61 to 11.08)	5.81 (5.52 to 6.10)	5.33 (5.09 to 5.58)
High risk(2-9)	30.46 (29.97 to 30.96)	21.07 (20.83 to 21.30)	20.32 (20.10 to 20.54)

embolism and ischaemic stroke (that is, excluding pulmonary embolism). The secondary outcome was death from any cause.

Statistical analysis

In patients discharged with non-valvular atrial fibrillation who were not receiving treatment with vitamin K antagonists or heparins, we estimated event rates for thromboembolism and death for the various $CHADS_2$ and CHA_2DS_2-VASc scores and for the specific covariate combinations forming the scores of 1 or 2. We estimated the risk of thromboembolism by using Cox proportional hazard regression models. In the Cox models, we analysed the risk associated with all possible risk factor combinations for $CHADS_2$ score=1 (four combinations) and $CHADS_2$ score=2 (seven combinations); we used $CHADS_2$ score=0 as the reference. In the same manner, other Cox models analysed the risk associated with all possible risk factor combinations for CHA_2DS_2-VASc score=1 (six combinations) and CHA_2DS_2-VASc score=2 (17 combinations), with CHA_2DS_2-VASc score=0 used as the reference. We did all analyses for one, five, and 10 years of follow-up. In additional Cox regression models, we included concomitant treatment with antiplatelet drugs (that is, primary acetylsalicylic acid, clopidogrel, and dipyridamole), to adjust for this potential confounder. We also did sensitivity analyses by not including pulmonary embolism as an outcome.

We used C statistics estimated from Cox regression models to assess the predictive capability of $CHADS_2$ and CHA_2DS_2-VASc for thromboembolism, using the method described by Liu and colleagues.[24] C statistics give a measure of how well the risk prediction scheme identifies patients who will have a future event. For estimating C statistics, we analysed $CHADS_2$ and CHA_2DS_2-VASc as risk scores (0-6 and 0-9) and as risk groups (low, intermediate, and high). We also evaluated the scores both as categorical and as continuous covariates. We constructed survival curves, based on Kaplan-Meier estimates of the probability of remaining free of thromboembolism with a score of 0 and 1, for the two risk stratification schemes. We considered a two sided P value <0.05 to be statistically significant. In all Cox models, the model assumptions (that is, proportional hazards, linearity of continuous covariates, and lack of interactions) were found to be valid. We used SAS statistical software version 9.1 and Stata statistical software version 11.0 for the analyses.

RESULTS

Of 121 280 patients with non-valvular atrial fibrillation, 73 538 (60.6%) fulfilled the study inclusion criteria (fig 1). Table 1 shows the baseline characteristics for all patients discharged with non-valvular atrial fibrillation and for the study population. During the one, five, and 10 years of follow-up, 9097 (12.4%), 13 966 (19.0%), and 15 344 (20.9%) of the non-anticoagulated patients claimed at least one prescription for a vitamin K antagonist, and exclusion of these patients from the time of starting vitamin K antagonist treatment (censoring) did not alter the results of our analyses (data not shown). Of the 16 406 patients categorised by $CHADS_2$ as being at low risk, 6472 (39.5%) were at intermediate risk and 3565 (21.7%) were at high risk when categorised by CHA_2DS_2-VASc. Of

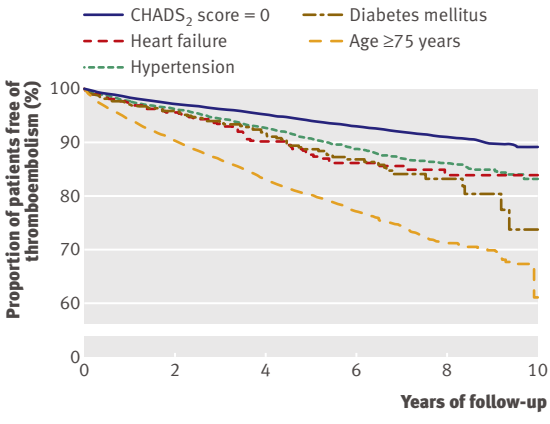

Fig 2 | Kaplan-Meier estimate of probability of remaining free of thromboembolism with $CHADS_2$ score 0 and 1. Only patients with $CHADS_2$ scores 0 and 1 were included, and patients were censored at death for causes other than thromboembolism

Table 4 | Event rates (95% CI) for hospital admission and death due to thromboembolism* per 100 person years by $CHADS_2$ score and by covariates

Score and covariates	1 year's follow-up	5 years' follow-up	10 years' follow-up
$CHADS_2$ score=0	1.67 (1.47 to 1.89)	1.28 (1.19 to 1.38)	1.24 (1.16 to 1.33)
$CHADS_2$ score=1:			
Heart failure	2.80 (1.81 to 4.34)	2.57 (2.00 to 3.30)	2.31 (1.82 to 2.93)
Hypertension	2.42 (2.04 to 2.87)	1.95 (1.77 to 2.16)	1.94 (1.77 to 2.13)
Age≥75	5.97 (5.55 to 6.41)	4.77 (4.55 to 5.00)	4.64 (4.44 to 4.85)
Diabetes mellitus	3.00 (1.97 to 4.55)	2.37 (1.84 to 3.05)	2.42 (1.93 to 3.04)
$CHADS_2$ score=2:			
Diabetes + heart failure	6.36 (3.31 to 12.23)	6.36 (4.33 to 9.34)	5.96 (4.12 to 8.63)
Diabetes + hypertension	2.81 (1.80 to 4.41)	2.75 (2.14 to 3.53)	2.78 (2.21 to 3.50)
Diabetes + age≥75	7.83 (6.13 to 10.01)	5.66 (4.75 to 6.74)	5.37 (4.52 to 6.36)
Heart failure + hypertension	4.44 (3.23 to 6.10)	3.44 (2.80 to 4.22)	3.28 (2.71 to 3.97)
Heart failure + age≥75	6.63 (5.77 to 7.62)	5.56 (5.06 to 6.10)	5.50 (5.03 to 6.02)
Hypertension + age≥75	6.93 (6.30 to 7.62)	5.65 (5.31 to 6.01)	5.47 (5.15 to 5.80)
Previous thromboembolism	15.46 (13.41 to 17.83)	8.25 (7.40 to 9.20)	7.74 (6.98 to 8.57)

*Includes peripheral artery embolism, ischaemic stroke, and pulmonary embolism.

the 23 730 patients categorised by $CHADS_2$ as being at intermediate risk, 21 999 (92.7%) were at high risk when categorised by CHA_2DS_2-VASc.

Table 2 shows rates of thromboembolism per 100 person years according to $CHADS_2$ and CHA_2DS_2-VASc risk scores at one, five, and 10 years of follow-up. The thromboembolic rates after one year of follow-up in the low risk category (score=0) were 1.67 (95% confidence interval 1.47 to 1.89) for $CHADS_2$ and 0.78 (0.58 to 1.04) for CHA_2DS_2-VASc. In the intermediate risk category (score=1), the rate of thromboembolism per 100 person years was 4.75 (4.45 to 5.07) for $CHADS_2$ and 2.01 (1.70 to 2.36) for CHA_2DS_2-VASc. This risk pattern was generally sustained at five and 10 years of follow-up; patients classified as being at intermediate risk by $CHADS_2$ had a higher rate of thromboembolism (approximately 3.6) than did those classified as being at intermediate risk by CHA_2DS_2-VASc (approximately 1.5). The high risk categories (score≥2) as determined by either $CHADS_2$ or CHA_2DS_2-VASc had markedly increased rates of thromboembolism compared with the low or intermediate risk categories.

We also estimated rates of thromboembolism in the vitamin K antagonist treated patients. In all risk categories, except for patients classified with CHA_2DS_2-VASc score=0, the thromboembolic rate was lower in the vitamin K antagonist treated patients. In these patients, thromboembolic rates after one year of follow-up in the low risk category were 1.27 (1.06 to 1.53) per 100 person years for $CHADS_2$ and 0.81 (0.56 to 1.17) for CHA_2DS_2-VASc. In the intermediate risk category, the rates were 2.27 (2.02 to 2.56) for $CHADS_2$ and 1.23 (0.98 to 1.56) for CHA_2DS_2-VASc.

Table 3 shows mortality rates according to $CHADS_2$ and CHA_2DS_2-VASc scores. We found a clear relation between increasing $CHADS_2$ and CHA_2DS_2-VASc score and increasing mortality rates. The low risk and intermediate risk categories as determined by CHA_2DS_2-VASc had much lower mortality rates than did patients categorised in these two risk groups by $CHADS_2$.

Table 4 shows rates of thromboembolism according to the risk factors composing $CHADS_2$ scores 0, 1, and 2; table 5 shows hazard ratios from Cox proportional hazard analysis. The risk associated with $CHADS_2$ score=1 depended on the specific conditions (risk factors) composing the score. The risk factor associated with the highest risk was age≥75 (hazard ratio 3.52, 95% confidence interval 3.05 to 4.07, at one year's

Table 5 | Hazard ratios for hospital admission and death due to thromboembolism* by combinations of covariates composing $CHADS_2$ scores 1 and 2

	1 year's follow-up		5 years' follow-up		10 years' follow-up	
Score and covariates	Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value
$CHADS_2$ score=0	1.00		1.00		1.00	
$CHADS_2$ score=1:						
Heart failure	1.67 (1.06 to 2.63)	0.03	1.99 (1.53 to 2.58)	<0.0001	1.84 (1.43 to 2.35)	<0.0001
Hypertension	1.45 (1.17 to 1.79)	0.0006	1.52 (1.34 to 1.73)	<0.0001	1.56 (1.39 to 1.74)	<0.0001
Age≥75	3.52 (3.05 to 4.07)	<0.0001	3.62 (3.31 to 3.96)	<0.0001	3.59 (3.31 to 3.90)	<0.0001
Diabetes mellitus	1.79 (1.16 to 2.77)	0.009	1.84 (1.41 to 2.39)	<0.0001	1.93 (1.53 to 2.45)	<0.0001
$CHADS_2$ score=2:						
Diabetes + heart failure	3.74 (1.93 to 7.28)	0.0001	4.75 (3.21 to 7.03)	<0.0001	4.52 (3.10 to 6.59)	<0.0001
Diabetes + hypertension	1.67 (1.05 to 2.67)	0.03	2.11 (1.62 to 2.74)	<0.0001	2.17 (1.71 to 2.76)	<0.0001
Diabetes + age≥75	4.57 (3.47 to 6.02)	<0.0001	4.16 (3.44 to 5.04)	<0.0001	3.98 (3.31 to 4.79)	<0.0001
Heart failure + hypertension	2.63 (1.87 to 3.70)	<0.0001	2.61 (2.09 to 3.24)	<0.0001	2.55 (2.08 to 3.12)	<0.0001
Heart failure + age≥75	3.84 (3.19 to 4.64)	<0.0001	4.04 (3.58 to 4.56)	<0.0001	4.04 (3.61 to 4.53)	<0.0001
Hypertension + age≥75	4.08 (3.48 to 4.77)	<0.0001	4.22 (3.83 to 4.66)	<0.0001	4.14 (3.78 to 4.53)	<0.0001
Previous thromboembolism	9.13 (7.55 to 11.04)	<0.0001	6.30 (5.52 to 7.19)	<0.0001	6.05 (5.35 to 6.83)	<0.0001

Results from Cox proportional hazard analyses; $CHADS_2$ score=0 was reference.
*Includes peripheral artery embolism, ischaemic stroke, and pulmonary embolism.

Table 6 | Event rates (95% CI) for hospital admission and death due to thromboembolism* per 100 person years by CHA_2DS_2-VASc score and by covariates

Score and covariates	1 year's follow-up	5 years' follow-up	10 years' follow-up
CHA_2DS_2-VASc score= 0	0.78 (0.58 to 1.04)	0.69 (0.59 to 0.81)	0.66 (0.57 to 0.76)
CHA_2DS_2-VASc score=1:			
Heart failure	1.50 (0.37 to 5.98)	2.35 (1.30 to 4.24)	1.78 (0.99 to 3.21)
Hypertension	2.14 (1.46 to 3.15)	1.60 (1.26 to 2.01)	1.49 (1.21 to 1.84)
Diabetes mellitus	3.47 (1.65 to 7.27)	2.28 (1.42 to 3.66)	2.02 (1.29 to 3.16)
Vascular disease	0.75 (0.24 to 2.33)	1.40 (0.91 to 2.15)	1.47 (1.01 to 2.12)
Age 65-74	2.88 (2.29 to 3.62)	2.13 (1.85 to 2.46)	2.09 (1.83 to 2.38)
Female sex	1.24 (0.89 to 1.73)	0.86 (0.70 to 1.06)	0.82 (0.68 to 1.00)
CHA_2DS_2-VASc score=2:			
Diabetes + heart failure	4.53 (0.64 to 32.17)	3.52 (1.13 to 10.91)	3.83 (1.44 to 10.21)
Diabetes + hypertension	3.29 (1.37 to 7.91)	1.79 (0.93 to 3.44)	1.94 (1.10 to 3.42)
Diabetes + age 65-74	1.49 (0.48 to 4.61)	1.92 (1.11 to 3.30)	1.98 (1.21 to 3.22)
Diabetes + vascular disease	0	1.06 (0.15 to 7.55)	1.80 (0.45 to 7.19)
Diabetes + female sex	1.11 (0.16 to 7.85)	0.62 (0.16 to 2.49)	1.23 (0.51 to 2.96)
Heart failure + hypertension	4.11 (1.96 to 8.62)	3.19 (1.98 to 5.14)	2.81 (1.79 to 4.41)
Heart failure + age 65-74	1.84 (0.69 to 4.90)	2.49 (1.55 to 4.01)	2.46 (1.59 to 3.82)
Heart failure + vascular disease	3.55 (0.50 to 25.17)	1.91 (0.48 to 7.66)	1.49 (0.37 to 5.97)
Heart failure + female sex	0	0.55 (0.08 to 3.87)	0.87 (0.22 to 3.49)
Hypertension + age 65-74	2.54 (1.74 to 3.70)	2.22 (1.79 to 2.76)	2.30 (1.89 to 2.78)
Hypertension + vascular disease	1.56 (0.70 to 3.48)	1.48 (0.96 to 2.30)	1.52 (1.02 to 2.24)
Hypertension + female sex	1.84 (1.09 to 3.11)	1.48 (1.09 to 2.02)	1.43 (1.08 to 1.89)
Age 65-74 + vascular disease	2.90 (1.72 to 4.89)	2.47 (1.82 to 3.35)	2.54 (1.93 to 3.35)
Age 65-74 + female sex	2.82 (2.21 to 3.60)	2.10 (1.81 to 2.45)	2.06 (1.80 to 2.36)
Vascular disease + female sex	2.87 (0.93 to 8.91)	1.95 (0.93 to 4.08)	2.26 (1.21 to 4.19)
Age≥75	4.75 (4.14 to 5.44)	4.37 (4.02 to 4.75)	4.27 (3.94 to 4.62)
Previous thromboembolism	16.07 (11.64 to 22.18)	7.87 (6.12 to 10.11)	6.98 (5.50 to 8.85)

*Includes peripheral artery embolism, ischaemic stroke, and pulmonary embolism.

follow-up), whereas hypertension was associated with the lowest thromboembolic risk (1.45, 1.17 to 1.79, at one year's follow-up). In patients with $CHADS_2$ score=2, previous thromboembolism as a single risk factor clearly carried the highest risk. Multivariable analyses including treatment with antiplatelet drugs yielded very similar results (web extra table A). Figure 2 shows thromboembolic-free survival curves for $CHADS_2$ scores 0 and 1.

Table 6 shows rates of thromboembolism according to risk factors composing CHA_2DS_2-VASc scores 0, 1, and 2; table 7 shows relative risks (hazard ratios). Again, the risk associated with a particular CHA_2DS_2-VASc score was strongly dependent on the specific conditions (risk factors) composing the score. For patients with CHA_2DS_2-VASc score=1, diabetes was associated with the highest thromboembolic rate (3.47, 1.65 to 7.27) and age 65-74 had the second highest rate (2.88, 2.29 to 3.62) at one year of follow-up. In patients with CHA_2DS_2-VASc score=2, previous thromboembolism as a single risk factor was clearly associated with the highest risk, followed by age≥75. Again, multivariable analyses including treatment with antiplatelet drugs yielded very similar results (web extra table B). Figure 3 shows thromboembolic-free survival curves for CHA_2DS_2-VASc scores 0 and 1. The thromboembolic rate with CHA_2DS_2-VASc score=1 was lower than that with $CHADS_2$ score=1, so the thromboembolic rate associated with one specific risk factor (for example, diabetes) was lower in the intermediate risk category (score=1) determined by CHA_2DS_2-VASc than by $CHADS_2$. Likewise, the thromboembolic rate with CHA_2DS_2-VASc score=0 was lower than with $CHADS_2$ score=0; using score=0 as the reference, the hazard ratio with a specific risk factor was higher in the intermediate risk category determined by CHA_2DS_2-VASc than by $CHADS_2$.

Table 8 shows how accurately $CHADS_2$ and CHA_2DS_2-VASc identified patients who had a thromboembolism during follow-up (C statistics based on Cox regression models); scores were entered in the analysis as categorical or continuous variables. The predictive abilities with $CHADS_2$ and CHA_2DS_2-VASc analysed as scores (0-6 and 0-9) were very similar, whereas the predictive ability of CHA_2DS_2-VASc was clearly superior to $CHADS_2$ for categorisation of patients into risk groups (low, intermediate, and high risk). At one, five, and 10 years of follow-up, C statistics with $CHADS_2$ were 0.722, 0.796, and 0.812; the corresponding C statistics with CHA_2DS_2-VASc were 0.850, 0.880, and 0.888. The 95% confidence intervals for $CHADS_2$ and CHA_2DS_2-VASc did not overlap.

Of the thromboembolic events representing the primary study outcome, pulmonary embolism comprised 7.7%. Sensitivity analyses excluding pulmonary embolism from the study outcome yield results similar to the main findings (web extra tables C and D). The predictive abilities (C statistics) for categorising patients into risk groups at one, five, and 10 years of follow-up were 0.711, 0.789, and 0.806 for $CHADS_2$ and 0.845, 0.877, and 0.885 for CHA_2DS_2-VASc. Again, the 95% confidence intervals did not overlap.

DISCUSSION

This nationwide study used the largest real world cohort of non-anticoagulated patients with non-

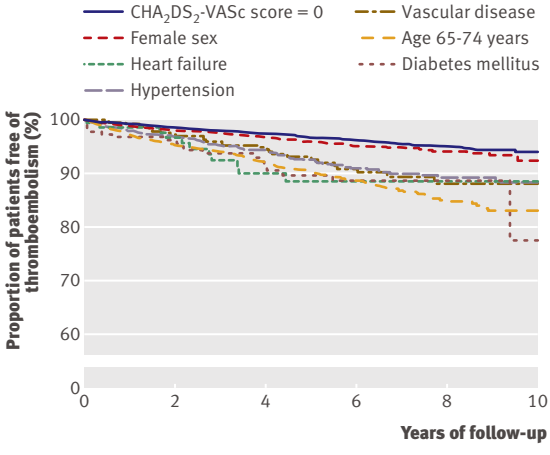

Fig 3 | Kaplan-Meier estimate of probability of remaining free of thromboembolism with CHA_2DS_2-VASc score 0 and 1. Only patients with CHA_2DS_2-VASc scores 0 and 1 were included, and patients were censored at death for causes other than thromboembolism

Table 7 | Hazard ratios for hospital admission and death due to thromboembolism* by combinations of covariates composing CHA_2DS_2-VASc scores 1 and 2

	1 year's follow-up		5 years' follow-up		10 years' follow-up	
Score and covariates	Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value
CHA_2DS_2-VASc score=0	1.00		1.00		1.00	
CHA_2DS_2-VASc score=1:						
Heart failure	1.92 (0.47 to 7.91)	0.37	3.39 (1.84 to 6.26)	<0.0001	2.69 (1.47 to 4.95)	0.001
Hypertension	2.76 (1.70 to 4.48)	<0.0001	2.32 (1.75 to 3.07)	<0.0001	2.26 (1.75 to 2.92)	<0.0001
Diabetes mellitus	4.46 (2.01 to 9.89)	0.0002	3.31 (2.00 to 5.46)	<0.0001	3.03 (1.89 to 4.86)	<0.0001
Vascular disease	0.97 (0.30 to 3.11)	0.96	2.04 (1.29 to 3.22)	0.002	2.22 (1.49 to 3.30)	<0.0001
Age 65-74	3.68 (2.54 to 5.34)	<0.0001	3.07 (2.48 to 3.80)	<0.0001	3.12 (2.57 to 3.78)	<0.0001
Female sex	1.60 (1.02 to 2.49)	0.04	1.25 (0.96 to 1.63)	0.10	1.24 (0.98 to 1.57)	0.08
CHA_2DS_2-VASc score=2:						
Diabetes + heart failure	5.80 (0.80 to 42.09)	0.08	5.13 (1.64 to 16.07)	0.005	5.79 (2.15 to 15.59)	0.0005
Diabetes + hypertension	4.23 (1.68 to 10.65)	0.002	2.58 (1.32 to 5.05)	0.006	2.90 (1.62 to 5.21)	0.0003
Diabetes + age 65-74	1.89 (0.59 to 6.09)	0.29	2.75 (1.56 to 4.85)	0.0005	2.94 (1.77 to 4.90)	<0.0001
Diabetes + vascular disease	–	0.98	1.54 (0.22 to 10.98)	0.67	2.69 (0.67 to 10.82)	0.16
Diabetes + female sex	1.42 (0.20 to 10.30)	0.73	0.90 (0.22 to 3.64)	0.88	1.86 (0.76 to 4.51)	0.17
Heart failure + hypertension	5.26 (2.37 to 11.66)	<0.0001	4.56 (2.76 to 7.53)	<0.0001	4.19 (2.62 to 6.72)	<0.0001
Heart failure + age 65-74	2.33 (0.84 to 6.47)	0.11	3.55 (2.15 to 5.86)	<0.0001	3.65 (2.30 to 5.78)	<0.0001
Heart failure + vascular disease	4.55 (0.63 to 33.02)	0.13	2.78 (0.69 to 11.22)	0.15	2.26 (0.56 to 9.09)	0.25
Heart failure + female sex	–	0.97	0.79 (0.11 to 5.67)	0.82	1.32 (0.33 to 5.32)	0.70
Hypertension + age 65-74	3.26 (2.02 to 5.25)	<0.0001	3.21 (2.45 to 4.20)	<0.0001	3.44 (2.71 to 4.37)	<0.0001
Hypertension + vascular disease	2.02 (0.86 to 4.73)	0.11	2.14 (1.35 to 3.42)	0.001	2.28 (1.50 to 3.46)	0.0001
Hypertension + female sex	2.37 (1.30 to 4.32)	0.005	2.15 (1.52 to 3.04)	<0.0001	2.16 (1.58 to 2.95)	<0.0001
Age 65-74 + vascular disease	3.70 (2.03 to 6.74)	<0.0001	3.56 (2.52 to 5.02)	<0.0001	3.80 (2.79 to 5.18)	<0.0001
Age 65-74 + female sex	3.61 (2.46 to 5.28)	<0.0001	3.04 (2.44 to 3.79)	<0.0001	3.11 (2.55 to 3.78)	<0.0001
Vascular disease + female sex	3.69 (1.15 to 11.88)	0.03	2.81 (1.32 to 5.99)	0.008	3.38 (1.79 to 6.38)	0.0002
Age≥75	5.96 (4.32 to 8.23)	<0.0001	6.16 (5.14 to 7.38)	<0.0001	6.21 (5.27 to 7.33)	<0.0001
Previous thromboembolism	20.44 (13.23 to 31.57)	<0.0001	11.27 (8.37 to 15.18)	<0.0001	10.44 (7.91 to 13.78)	<0.0001

Results from Cox proportional-hazard analyses; CHA_2DS_2-VASc score=0 was reference.
*Includes peripheral artery embolism, ischaemic stroke, and pulmonary embolism.

valvular atrial fibrillation ever investigated. We found that CHA_2DS_2-VASc performed better than $CHADS_2$ for categorisation of patients into risk groups (low, intermediate, and high risk) for stroke and for identification of patients at "truly low risk" (score=0). With a $CHADS_2$ score of 0 or 1, the thromboembolic risk was still appreciable. Also, not all risk factors composing $CHADS_2$ score=1 were associated with an equal risk; a particularly high risk was associated with age≥75. Patients with CHA_2DS_2-VASc score=1 had a lower rate of thromboembolism, which would seem more "truly moderate." Again, not all risk factors in CHA_2DS_2-VASc score=1 were associated with an equal risk, and diabetes and age 65-74 were associated with the highest thromboembolic rates. With $CHADS_2$ or CHA_2DS_2-VASc score=2, again the thromboembolic rates were strongly dependent on the specific covariates composing the score, and the risk associated with previous thromboembolism was markedly increased compared with all other combinations of risk factors. We also found that with CHA_2DS_2-VASc score=0 the risk was "truly low" and no reduction in thromboembolic rate occurred with vitamin K antagonist treatment, whereas the thromboembolic rate was reduced in vitamin K antagonist treated patients with $CHADS_2$ scores 0-1 and CHA_2DS_2-VASc score=1.

Implications of findings

The advantage of $CHADS_2$ is its simplicity, although its limitations are well recognised.[25] $CHADS_2$ was developed by amalgamation of risk schemes derived from clinical trial cohorts; it was initially validated in a cohort of patients with atrial fibrillation admitted to hospital, with event rates reflecting clinical practice more than a decade ago.[11] Even the "C" in $CHADS_2$ has been debated, as a history of congestive heart failure does not seem to be an independent risk factor for thromboembolism.[2 11 26] The risk of thromboembolism increases with increasing age above 65 years,[27] rising approximately 1.5-fold per decade.[2] The increased effect of age≥75 as a single high risk factor was suggested by cohort analyses and a recent semi-systematic review,[6 28] and this finding was confirmed in our study. Other risk factors such as female sex and previous vascular disease have been recognised.[12 29-32] Our study

Table 8 | C statistics (95% CI) based on Cox regression models with covariates analysed as categorical or continuous variables

	1 year's follow-up	5 years' follow-up	10 years' follow-up
Covariates analysed as categorical variables			
$CHADS_2$; score 0-6	0.663 (0.634 to 0.691)	0.762 (0.744 to 0.780)	0.781 (0.764 to 0.797)
$CHADS_2$; 3 groups	0.722 (0.694 to 0.748)	0.796 (0.778 to 0.812)	0.812 (0.796 to 0.827)
CHA_2DS_2-VASc; score 0-9	0.661 (0.633 to 0.690)	0.758 (0.740 to 0.776)	0.777 (0.760 to 0.793)
CHA_2DS_2-VASc; 3 groups	0.850 (0.829 to 0.871)	0.880 (0.866 to 0.893)	0.888 (0.875 to 0.900)
Covariates analysed as continuous variables			
$CHADS_2$; score 0-6	0.691 (0.663 to 0.719)	0.787 (0.770 to 0.804)	0.804 (0.788 to 0.819)
$CHADS_2$; 3 groups	0.722 (0.694 to 0.748)	0.796 (0.778 to 0.812)	0.812 (0.796 to 0.827)
CHA_2DS_2-VASc; score 0-9	0.682 (0.653 to 0.709)	0.775 (0.757 to 0.793)	0.792 (0.776 to 0.808)
CHA_2DS_2-VASc; 3 groups	0.852 (0.830 to 0.873)	0.882 (0.868 to 0.895)	0.890 (0.877 to 0.902)

suggests the value of these conditions for prediction of thromboembolism; in patients with CHA_2DS_2-VASc score=1, female sex increased the risk of thromboembolism at one year of follow-up and vascular disease increased thromboembolic risk at five and 10 years of follow-up.

Because of the benefit of oral anticoagulation over aspirin, in patients with atrial fibrillation and $CHADS_2$ score=2 the clinical impetus would be to anti coagulate.[5-7] With $CHADS_2$ scores 0-1, or where a more comprehensive stroke risk and vitamin K antagonist risk/benefit assessment is necessary, a need clearly exists to consider other risk factors not included in the $CHADS_2$ score. This large study in a non-anticoagulated cohort with non-valvular atrial fibrillation clearly shows advantages of CHA_2DS_2-VASc for further refinement of thromboembolic risk stratification, with improvements in C statistics for identification of patients at low, intermediate, and high risk of thromboembolism and a convincing identification of those with a truly low risk of thromboembolism. Use of CHA_2DS_2-VASc could therefore simplify thromboprophylaxis, with CHA_2DS_2-VASc score=0 identifying patients at truly low risk for whom no antithrombotic treatment may be considered. With CHA_2DS_2-VASc score=1, oral anticoagulation can be used, given the limited evidence for the efficacy of aspirin (which also has a potential for bleeding[8]) and with consideration of the future availability of new oral anticoagulant drugs that can overcome the clinical disadvantages of vitamin K antagonists (for example, without the need for monitoring of anticoagulation and with less risk of bleeding). Also, when the intermediate risk category was defined as CHA_2DS_2-VASc score=1, only 11.2% of patients were categorised in this group, compared with 32.3% when the $CHADS_2$ score=1 definition was used. Based on the 2006 ACC/AHA/ESC guidelines, which recommend vitamin K antagonist or aspirin for this intermediate risk category, using the $CHADS_2$ score rather than the CHA_2DS_2-VASc score would open more patients to the uncertainty of vitamin K antagonist or aspirin and could even result in aspirin being used instead of vitamin K antagonist, as the guidelines do not provide definitive recommendations. Finally, the decision to anticoagulate is based not only on thromboembolic risk but also on the risk of bleeding, and the European guidelines on atrial fibrillation incorporate a new bleeding prediction scheme to help with this decision making.[33]

Limitations of study

The main limitation of this study was inherent to its observational nature. We had no information on the reason(s) for absence of vitamin K antagonist treatment in this specific cohort of patients with non-valvular atrial fibrillation, and we could not differentiate between paroxysmal, persistent, and permanent atrial fibrillation and atrial fibrillation that had been triggered by a single episode of acute illness. Even though the positive predictive value of the diagnosis of atrial fibrillation is very high in the registry (99%),[29] and data on prescription claims are accurate,[16] retrospective studies may be affected by misclassification and inclusion bias—for example, by including only patients admitted to hospital with atrial fibrillation we might have increased the proportion of patients who were at higher risk of thromboembolic events and death. Furthermore, to account for treatment started in relation to the admission for atrial fibrillation, we defined the study baseline at day seven from discharge, thereby excluding 2.5% of the population with atrial fibrillation from the study.

The frequencies of risk factors in the study population were also underestimated, as we identified patients with heart failure, hypertension, and diabetes from prescription claims and thus did not detect patients treated with diet control and lifestyle interventions alone. Therefore, the estimated thrombotic risk must be applied with caution in these populations. Furthermore, we were not able to account for effects of smoking, family history of thromboembolism, alcohol intake, or body mass index. The outcome diagnoses of stroke and pulmonary embolism have previously been validated in the registry; the positive predictive value of ischaemic stroke (I63) was 97%, haemorrhagic strokes only comprised 5.8% of the unspecified strokes (I64), and pulmonary embolism (I26) had a positive predictive value of 82.1%.[21 23] However, patients with previous strokes were excluded in the validation study and in our study they were not, so the risk remains that some of the stroke outcomes in this study may in fact have been recoding of old strokes, which again would lead to overestimation of the observed high risk associated with previous stroke.

Conclusions

The risk associated with a specific risk score in both $CHADS_2$ and CHA_2DS_2-VASc depends on the risk factors composing the score. CHA_2DS_2-VASc performed better than $CHADS_2$ in predicting patients at high risk and can also be used to identify patients with non-valvular atrial fibrillation with a truly low risk of thromboembolism.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Thromboembolic risk stratification of patients with non-valvular atrial fibrillation is essential for selection of optimal antithrombotic treatment

The most commonly used risk stratification scheme is $CHADS_2$; CHA_2DS_2-VASc was developed to complement $CHADS_2$ by considering additional thromboembolic risk modifiers

WHAT THIS STUDY ADDS

CHA_2DS_2-VASc is more valid for stroke prediction in patients categorised as being at low and intermediate risk by the $CHADS_2$ scheme

This is clinically important, as many of the patients at low risk according to $CHADS_2$ are not at "truly low risk" and treatment guidelines are not conclusive for those at intermediate risk

The importance of each component of the $CHADS_2$ and CHA_2DS_2-VASc scores for thromboembolism risk has been estimated

Contributors: JBO made primary contributions to data collection and analysis, interpretation of results, and writing of the manuscript. GYHL helped to write the first draft. GYHL, MLH, PRH, GHG, and CT-P contributed to the study conception and design. All authors contributed to interpretation of results, all revised the manuscript critically for important intellectual content, and all approved the final manuscript. JBO is the guarantor.
Funding: None.
Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.
Ethical approval: No ethics approval is needed for retrospective register studies in Denmark. The study was approved by the Danish Data Protection Agency (No 2007-41-1667).
Data sharing: No additional data available.

1 Hughes M, Lip GY. Stroke and thromboembolism in atrial fibrillation: a systematic review of stroke risk factors, risk stratification schema and cost effectiveness data. *Thromb Haemost* 2008;99:295-304.
2 Stroke Risk in Atrial Fibrillation Working Group. Independent predictors of stroke in patients with atrial fibrillation: a systematic review. *Neurology* 2007;69:546-54.
3 Lim HS, Lip GY. Thromboprophylaxis in acute ischaemic stroke: how can we PREVAIL? *Lancet Neurol* 2007;6:578-9.
4 Fang MC, Go AS, Chang Y, Borowsky L, Pomernacki NK, Singer DE. Comparison of risk stratification schemes to predict thromboembolism in people with nonvalvular atrial fibrillation. *J Am Coll Cardiol* 2008;51:810-5.
5 Lee BH, Park JS, Park JH, Kwak JJ, Hwang ES, Kim SK, et al. The effect and safety of the antithrombotic therapies in patients with atrial fibrillation and CHADS score 1. *J Cardiovasc Electrophysiol* 2010;21:501-7.
6 Gorin L, Fauchier L, Nonin E, de Labriolle A, Haguenoer K, Cosnay P, et al. Antithrombotic treatment and the risk of death and stroke in patients with atrial fibrillation and a CHADS2 score=1. *Thromb Haemost* 2010;103:833-40.
7 Healey JS, Hart RG, Pogue J, Pfeffer MA, Hohnloser SH, De Caterina R, et al. Risks and benefits of oral anticoagulation compared with clopidogrel plus aspirin in patients with atrial fibrillation according to stroke risk: the atrial fibrillation clopidogrel trial with irbesartan for prevention of vascular events (ACTIVE-W). *Stroke* 2008;39:1482-6.
8 Sato H, Ishikawa K, Kitabatake A, Ogawa S, Maruyama Y, Yokota Y, et al. Low-dose aspirin for prevention of stroke in low-risk patients with atrial fibrillation: Japan Atrial Fibrillation Stroke Trial. *Stroke* 2006;37:447-51.
9 Van Walraven C, Hart RG, Wells GA, Petersen P, Koudstaal PJ, Gullov AL, et al. A clinical prediction rule to identify patients with atrial fibrillation and a low risk for stroke while taking aspirin. *Arch Intern Med* 2003;163:936-43.
10 Lip GY, Halperin JL. Improving stroke risk stratification in atrial fibrillation. *Am J Med* 2010;123:484-8.
11 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.
12 Fuster V, Ryden LE, Cannom DS, Crijns HJ, Curtis AB, Ellenbogen KA, et al. ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: full text: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the European Society of Cardiology Committee for Practice Guidelines (writing committee to revise the 2001 guidelines for the management of patients with atrial fibrillation) developed in collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. *Europace* 2006;8:651-745.
13 Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. 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 Poli D, Lip GY, Antonucci E, Grifoni E, Lane D. Stroke risk stratification in a "real-world" elderly anticoagulated atrial fibrillation population. *J Cardiovasc Electrophysiol* 2010 July 23;Epub ahead of print.
15 Andersen TF, Madsen M, Jorgensen J, Mellemkjoer L, Olsen JH. The Danish national hospital register: a valuable source of data for modern health sciences. *Dan Med Bull* 1999;46:263-8.
16 Gaist D, Sorensen HT, Hallas J. The Danish prescription registries. *Dan Med Bull* 1997;44:445-8.
17 Go AS, Hylek EM, Borowsky LH, Phillips KA, Selby JV, Singer DE. Warfarin use among ambulatory patients with nonvalvular atrial fibrillation: the anticoagulation and risk factors in atrial fibrillation (ATRIA) study. *Ann Intern Med* 1999;131:927-34.
18 Kumler T, Gislason GH, Kirk V, Bay M, Nielsen OW, Kober L, et al. Accuracy of a heart failure diagnosis in administrative registers. *Eur J Heart Fail* 2008;10:658-60.
19 Ekholm O, Hesse U, Davidsen M, Kjoller M. The study design and characteristics of the Danish national health interview surveys. *Scand J Public Health* 2009;37:758-65.
20 Madsen M, Davidsen M, Rasmussen S, Abildstrom SZ, Osler M. The validity of the diagnosis of acute myocardial infarction in routine statistics: a comparison of mortality and hospital discharge data with the Danish MONICA registry. *J Clin Epidemiol* 2003;56:124-30.
21 Krarup LH, Boysen G, Janjua H, Prescott E, Truelsen T. Validity of stroke diagnoses in a national register of patients. *Neuroepidemiology* 2007;28:150-4.
22 Joensen AM, Jensen MK, Overvad K, Dethlefsen C, Schmidt E, Rasmussen L, et al. Predictive values of acute coronary syndrome discharge diagnoses differed in the Danish national patient registry. *J Clin Epidemiol* 2009;62:188-94.
23 Severinsen MT, Kristensen SR, Overvad K, Dethlefsen C, Tjonneland A, Johnsen SP. Venous thromboembolism discharge diagnoses in the Danish national patient registry should be used with caution. *J Clin Epidemiol* 2010;63:223-8.
24 Liu L, Forman S, Barton B. Fitting Cox model using PROC PHREG and beyond in SAS. 2009. http://support.sas.com/resources/papers/proceedings09/236-2009.pdf.
25 Karthikeyan G, Eikelboom JW. The CHADS2 score for stroke risk stratification in atrial fibrillation—friend or foe? *Thromb Haemost* 2010;104:45-8.
26 Stroke Risk in Atrial Fibrillation Working Group. Comparison of 12 risk stratification schemes to predict stroke in patients with nonvalvular atrial fibrillation. *Stroke* 2008;39:1901-10.
27 Van Walraven C, Hart RG, Connolly S, Austin PC, Mant J, Hobbs FD, et al. Effect of age on stroke prevention therapy in patients with atrial fibrillation: the atrial fibrillation investigators. *Stroke* 2009;40:1410-6.
28 Marinigh R, Lip GY, Fiotti N, Giansante C, Lane DA. Age as a risk factor for stroke in atrial fibrillation patients: implications for thromboprophylaxis. *J Am Coll Cardiol* 2010;56:827-37.
29 Frost L, Andersen LV, Vestergaard P, Husted S, Mortensen LS. Trend in mortality after stroke with atrial fibrillation. *Am J Med* 2007;120:47-53.
30 Lane DA, Lip GY. Female gender is a risk factor for stroke and thromboembolism in atrial fibrillation patients. *Thromb Haemost* 2009;101:802-5.
31 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.
32 Lip GY. Coronary artery disease and ischemic stroke in atrial fibrillation. *Chest* 2007;132:8-10.
33 Camm AJ, Kirchhof P, Lip GY, Schotten U, Savelieva I, Ernst S, 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.

Accepted: 28 December 2010