

Atrial fibrillation ablation in practice: assessing CABANA generalizability

Peter A. Noseworthy^{1,2*}, Bernard J. Gersh², David M. Kent^{3,4}, Jonathan P. Piccini⁵, Douglas L. Packer², Nilay D. Shah^{1,6,7}, and Xiaoxi Yao^{1,6}

¹Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, 200 1st St SW, Rochester, MN, USA; ²Department of Cardiovascular Medicine, Mayo Clinic, 200 1st St SW, Rochester, MN, USA; ³Predictive Analytics and Comparative Effectiveness (PACE) Center, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center/Tufts University School of Medicine, 800 Washington St, Boston, MA 02111, USA; ⁴Department of Neurology, Tufts Medical Center/Tufts University School of Medicine, 800 Washington St, Boston, MA 02111, USA; ⁵Duke Center for Atrial Fibrillation, Duke Clinical Research Institute, Duke University Medical Center, 2400 Pratt St, Durham, NC 27705, USA; ⁶Division of Health Care Policy and Research, Department of Health Sciences Research, Mayo Clinic, 205 3rd Ave SW, Rochester, MN 55905, USA; and ⁷OptumLabs, One Main Street, 10th Floor, Cambridge, MA 02142, USA

Received 17 April 2018; revised 16 May 2018; editorial decision 9 February 2019; accepted 11 February 2019; online publish-ahead-of-print 15 March 2019

See page 1265 for the editorial comment on this article (doi: 10.1093/eurheartj/ehz168)

Aims

The Catheter Ablation vs. Antiarrhythmic Drug Therapy for Atrial Fibrillation (CABANA) trial aimed to assess the impact of ablation on morbidity and mortality. This observational study was conducted in parallel to CABANA to assess trial generalizability.

Methods and results

Using a large US administrative database, we identified 183 760 patients with atrial fibrillation (AF) treated with ablation or medical therapy (antiarrhythmic or rate control drugs) between 1 August 2009 and 30 April 2016 (CABANA enrolment period). Propensity score weighting was used to balance patients treated with ablation ($N=12\,032$) or medical therapy alone ($N=171\,728$) on 90 dimensions. Ablation was associated with a reduction in the composite endpoint of all-cause mortality, stroke, major bleeding, and cardiac arrest [hazard ratio (HR) 0.75, 95% confidence interval (CI) 0.70–0.81; $P<0.001$]. The majority of patients (73.8%) were potentially trial eligible; among whom the risk reduction associated with ablation was greatest (HR 0.70, 95% CI 0.63–0.77; $P<0.001$). Among the 3.8% of patients who failed to meet the inclusion criterion, i.e. patients under 65 years without stroke risk factors, the event rates were low and there was no significant relationship with ablation (HR 0.67, 95% CI 0.29–1.56; $P=0.35$). Among the 22.4% patients who met at least one of the trial exclusion criteria, there was a lesser but statistically significant reduction associated with ablation (HR 0.85, 95% CI 0.75–0.95; $P=0.01$).

Conclusion

In routine clinical care, ablation was associated with a reduction in the primary CABANA composite endpoint of all-cause mortality, stroke, major bleeding, and cardiac arrest, particularly in patients who were eligible for the trial.

Keywords

Atrial fibrillation • Ablation • Stroke • Mortality

Introduction

Among the 33 million people with atrial fibrillation (AF) worldwide, many suffer from not only bothersome symptoms and reduced quality of life, but also a five-fold increased risk of stroke and two-fold increased risk of death.^{1–4} Catheter ablation is a well-established therapy to relieve AF-associated symptoms,^{5,6} but emerging evidence suggests it may also improve cardiovascular outcomes.^{7–10} Currently, fewer than 5% of AF patients undergo ablation, but definitive

evidence that catheter ablation improves major adverse cardiovascular outcomes and survival could justify extending the indication for ablation beyond symptom control.^{11,12}

The Catheter Ablation vs. Antiarrhythmic Drug Therapy for Atrial Fibrillation (CABANA) trial randomized patients to ablation or medical therapy to determine whether ablation decreases morbidity and mortality in comparison to rhythm or rate control drugs.¹³ In order to be powered for hard clinical endpoints, the study included patients with at least one risk factor for stroke (<65 years, hypertension,

* Corresponding author. Tel: +1 507 255 2446, Fax: +1 507 255 2550, Email: noseworthy.peter@mayo.edu

Published on behalf of the European Society of Cardiology. All rights reserved. © The Author(s) 2019. For permissions, please email: journals.permissions@oup.com.

diabetes, congestive heart failure, prior stroke, transient ischaemic attack or systemic emboli, or vascular disease, left atrial size >5.0 cm, or ejection fraction $\leq 35\%$). Given the potential benefits, cost, and periprocedural risks of ablation,^{14,15} translation of this pivotal trial to practice will demand evidence of the trial's generalizability.

This study, designed to assess the CABANA trial generalizability, was conducted prior to the release of the trial results and was agnostic to its findings. We assessed the proportion of patients in routine practice who would have met trial eligibility and examined the association between ablation and clinical outcomes, stratified by potential trial eligibility.

Methods

Study population

This study was a retrospective cohort analysis using OptumLabs Data Warehouse, which contains patients with private insurance or Medicare Advantage of all ages and races throughout the US.^{16,17} The study population included adult patients (≥ 18 years) with AF, who were treated with ablation or medical therapy [antiarrhythmic drugs (AADs) or rate control drugs] between 1 August 2009 and 30 April 2016, the enrolment period of the CABANA trial (see patients selection flow diagram in [Supplementary material online, Figure S1](#)). To identify the ablated patients, we captured the ablation procedures with a primary diagnosis of AF. If a patient received multiple ablations, we anchored on the date of the first ablation, defined as the index date. Patients who received ablation before 1 August 2009 or after 30 April 2016 were excluded from the study. For the drug-treated cohort, we identified patients who did not undergo ablation at any time but received medical therapy. If patients received multiple study drugs, we randomly selected one drug, so called 'index medication', and defined the first fill date of the index medication as the index date. Using this method, the drug-treated group contained some patients who had received other study drugs before the index date, just as the ablation group contained patients with prior drug exposure. Patients were required to have at least 12 months of continuous enrolment in health insurance plans, defined as the baseline period, in order to capture an adequate medical history. The mean baseline period was 4.4 ± 3.4 years. The Mayo Clinic Institutional Review Board (IRB) exempted this study from review, because the study used pre-existing, de-identified data.

Outcomes

The primary outcome was a composite endpoint of all-cause mortality, stroke, major bleeding, and cardiac arrest, i.e. the same primary endpoint as the CABANA trial. The secondary outcomes were each of these outcomes considered separately. Patients were followed until the end of the study period (31 July 2017), the end of enrolment in health insurance plans, or death, whichever happened first. Stroke, major bleeding, and cardiac arrest were defined as a primary diagnosis during an emergency room visit or an inpatient stay. The codes used to define each endpoint are detailed in [Supplementary material online, Table S1](#). Mortality was identified based on the Social Security Death Master File and discharge status.

Statistical analysis

We calculated the proportion of patients who would be excluded from the trial based on the operational definition illustrated in [Supplementary material online, Table S2](#). We divided patients to three subgroups: (i) patients who would be eligible for CABANA; (ii) patients who failed to meet the inclusion criterion, i.e. those under 65 years without any stroke

risk factors; (iii) patients who met at least one of the exclusion criteria. A small proportion of patients both failed to meet the inclusion criterion and met the exclusion criteria. In the stratified analyses for clinical outcomes, such patients were classified as those who met the exclusion criteria.

Propensity score overlap weighting was used to account for the differences in baseline characteristics between patients who underwent ablation and those who were treated with medical therapy alone. A propensity score, the probability of undergoing ablation, was estimated using logistic regression based on socio-demographics, medical history, concurrent medication use, previous treatment with AADs or rate control drugs, the year of the index date, and the length of baseline period (variables in [Supplementary material online, Table S3](#)). The distribution of propensity scores is shown in [Supplementary material online, Figure S2](#). The overlap weight was calculated as $1 - \text{propensity score for the ablated patients, and the propensity score for the drug-treated patients}$. This weight is used to calculate the average treatment effect for the overlap population. This approach minimizes the asymptotic variance of the treatment effect, while also possessing a desirable exact balance property.¹⁸ The propensity score and weight were calculated separately in each of the three subgroups (patients who were eligible for CABANA, patients who failed to meet the inclusion criterion, and patients who met one of the exclusion criteria) in order to ensure optimal balance in each of the subgroups. Cox proportional hazards regression was used to compare patients treated with ablation and medical therapy in the propensity score-weighted cohort, with a robust sandwich estimator for variance estimation. The Fine and Gray method was used to consider death as a competing risk when assessing non-fatal outcomes (i.e. stroke, bleeding, or cardiac arrest when considered separately).¹⁹ The proportional hazards assumption was tested on the basis of Schoenfeld residuals.²⁰

Sensitivity analyses

First, we performed subgroup analyses for the primary outcome stratified by age, sex, race, CHA₂DS₂-VASc, hypertension with left ventricular hypertrophy, heart failure, cardiomyopathy, sleep apnoea, and prior thromboembolism. Second, one-to-one propensity score matching was used instead of propensity score weighting. Third, we conducted a stratified analysis based on whether the drug-treated patients were treated with AADs or with rate control drugs only. Fourth, we conducted a stratified analysis based on the medication adherence in the drug-treated patients. Fifth, we explored the impact of potential protocol deviation in the trial on the treatment effect, assuming 30% of the medical therapy cohort crossed over to the ablation cohort and 10% of the ablation cohort did not receive the procedure.¹³

We assessed residual confounding using two methods. First, we tested three falsification endpoints that are unlikely to be a result of ablation but might be related to unmeasured confounders such as frailty: chronic obstructive pulmonary disease, pneumonia, and fracture.²¹ Second, we used the method outlined by Lin *et al.*²² to assess whether the observed difference could be fully explained by an unmeasured confounder.

A two-sided *P*-value of <0.05 was considered to indicate statistical significance. No adjustment for multiple testing was performed. All the analyses except those related to the primary outcome were considered to be exploratory.

Results

Patient characteristics

This study identified 183 760 patients with AF treated with ablation or medical therapy between 1 August 2009 and 30 April 2016, among

Table 1 Baseline characteristics before and after propensity score weighting

	Before PS weighting				After PS Weighting		
	Drug treated (N = 171 728)	Ablation (N = 12 032)	Total (N = 183 760)	Standard mean difference	Drug treated (N = 171 728)	Ablation (N = 12 032)	Standard mean difference
Age (years), mean ± SD	70.7 ± 11.7	62.0 ± 10.9	70.1 ± 11.9	0.77	63.7 ± 11.9	63.7 ± 11.1	0.00
18–64	27.6%	58.4%	29.7%	0.65	50.7%	50.7%	0.00
65–74	28.1%	29.6%	28.2%	0.03	32.4%	32.4%	0.00
≥75	44.2%	12.0%	42.1%	0.77	16.9%	16.9%	0.00
Female	45.0%	31.0%	44.1%	0.29	34.5%	34.5%	0.00
Medical history							
Prior AF hospitalization	55.4%	53.3%	55.3%	0.04	52.2%	52.2%	0.00
Heart failure	45.6%	32.3%	44.7%	0.27	34.5%	34.5%	0.00
Systolic heart failure	17.9%	12.4%	17.6%	0.16	13.2%	13.2%	0.00
Hypertension	89.4%	83.1%	89.0%	0.18	85.2%	85.2%	0.00
Diabetes mellitus	37.6%	24.8%	36.7%	0.28	27.5%	27.5%	0.00
Ischaemic stroke	15.8%	7.8%	15.2%	0.25	9.3%	9.3%	0.00
Myocardial infarction	23.0%	14.2%	22.4%	0.23	15.9%	15.9%	0.00
CABG	14.6%	6.9%	14.1%	0.25	8.3%	8.3%	0.00
PCI	13.8%	11.1%	13.6%	0.08	12.1%	12.1%	0.00
Left ventricular hypertrophy	34.1%	35.5%	34.2%	0.03	34.8%	34.8%	0.00
Cardiomyopathy							
None	79.2%	76.3%	79.0%	0.07	76.6%	76.6%	0.00
Hypertrophic	1.3%	1.9%	1.3%	0.05	1.9%	1.9%	0.00
Ischaemic	3.3%	3.7%	3.3%	0.02	3.6%	3.6%	0.00
Dilated	16.3%	18.1%	16.4%	0.05	17.8%	17.8%	0.00
Implanted device							
None	84.3%	85.5%	84.4%	0.03	83.1%	83.1%	0.00
Biventricular or CRT pacemaker	1.0%	0.9%	1.0%	0.00	1.2%	1.2%	0.00
ICD	5.4%	5.4%	5.4%	0.00	6.5%	6.5%	0.00
Dual chamber pacemaker	6.6%	5.2%	6.5%	0.06	6.4%	6.4%	0.00
Single chamber pacemaker	2.5%	1.8%	2.5%	0.05	2.1%	2.1%	0.00
ILR	0.2%	1.1%	0.3%	0.11	0.8%	0.8%	0.00
Prior valve procedure	4.7%	1.7%	4.5%	0.17	2.1%	2.1%	0.00
Mitral stenosis	3.1%	2.2%	3.1%	0.06	2.6%	2.6%	0.00
Mitral regurgitation	41.4%	50.3%	42.0%	0.18	48.4%	48.4%	0.00
Major bleeding	28.7%	23.5%	28.3%	0.12	24.4%	24.4%	0.00
Intracranial bleeding	3.1%	1.3%	3.0%	0.12	1.7%	1.7%	0.00
Stage 3–5 CKD	17.0%	7.5%	16.3%	0.29	9.1%	9.1%	0.00
Cardioversion	11.9%	46.8%	14.2%	0.83	37.9%	37.9%	0.00
Previous drug treatment							
Number of previous AADs							
0	84.7%	23.6%	80.7%	1.55	37.6%	37.6%	0.00
1	13.4%	49.1%	15.7%	0.84	45.8%	45.8%	0.00
2	1.7%	20.8%	2.9%	0.63	13.4%	13.4%	0.00
≥3	0.3%	6.5%	0.7%	0.35	3.2%	3.2%	0.00
Concurrent medication							
Oral anticoagulant	28.2%	70.8%	31.0%	0.94	60.6%	60.6%	0.00
ACE inhibitors	26.5%	24.7%	26.4%	0.04	25.0%	25.0%	0.00
ARB	14.3%	14.7%	14.3%	0.01	14.9%	14.9%	0.00
Other calcium channel blockers	15.5%	9.7%	15.1%	0.18	10.8%	10.8%	0.00
Insulin	6.8%	3.9%	6.6%	0.13	4.5%	4.5%	0.00
CHA ₂ DS ₂ -VASc	4.4 ± 2.1	3.1 ± 1.9	4.3 ± 2.1	0.68	3.4 ± 1.9	3.4 ± 2.0	0.00

The CHA₂DS₂-VASc score is a 0- to 9-point stroke risk score where a higher point score indicates higher risk of stroke. The point score is calculated as follows: 1 point each for heart failure, hypertension, diabetes, vascular disease, age 65–74 years, and female sex and 2 points for age 75 years or older and prior thromboembolism (including ischaemic stroke, transient ischaemic attack, or systemic embolism). Concurrent medication was defined as a prescription fill within 3 months prior to the index date.

AAD, antiarrhythmic drug; ACE, angiotensin-converting enzyme; AF, atrial fibrillation; ARB, angiotensin II receptor blockers; CABG, coronary artery bypass grafting; CKD, chronic kidney disease; CRT, cardiac resynchronization therapy; ICD, implantable cardioverter-defibrillators; ILR, implantable loop recorder; PCI, percutaneous coronary intervention; PS, propensity score.

Table 2 Outcomes in overall propensity score-weighted patients (N = 183 760)

	Drug treated (N = 171 728)			Ablated (N = 12 032)			Absolute reduction in event rate (95% CI)	Hazard ratio (95% CI)	P-value
	Number of events	Person years	Event rate	Number of events	Person years	Event rate			
Composite	848	13 972	6.07	672	14 912	4.51	1.56 (1.19–1.92)	0.75 (0.70–0.81)	<0.001
All-cause mortality	520	14 522	3.58	369	15 513	2.38	1.20 (0.94–1.46)	0.67 (0.61–0.74)	<0.001
Ischaemic stroke	135	14 347	0.94	83	15 414	0.54	0.40 (0.27–0.53)	0.59 (0.48–0.73)	<0.001
Major bleeding	286	14 133	2.02	310	15 010	2.07	-0.04 (-0.28 to 0.19)	1.05 (0.94–1.18)	0.39
Cardiac arrest	48	14 519	0.33	27	15 508	0.17	0.16 (0.08–0.23)	0.54 (0.37–0.78)	0.001

Event rate was calculated as number of events per 100 person-years.

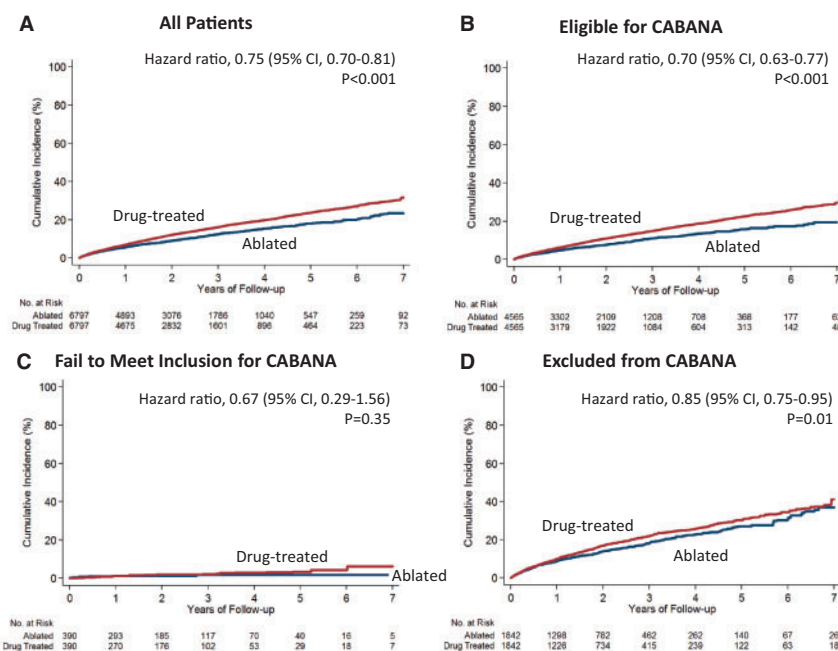


Figure 1 Primary endpoint (composite of mortality, stroke, major bleeding, or cardiac arrest) in ablated or drug-treated patients, stratified by CABANA trial eligibility criteria. The cumulative incidence in the overall cohort (A), in patients who would be potentially eligible for CABANA (B), in patients who failed to meet the inclusion criterion (C), and in patients who met at least one of the CABANA exclusion criteria (D). The drug-treated cohort was the reference group in the Cox proportional hazards regression analyses. All the curves and numbers were calculated using propensity score weighting.

whom 12 032 (6.5%) received ablation (Table 1). The majority of patients (73.8%) would have been eligible for enrolment in CABANA. Only 4.4% of patients did not meet the inclusion criteria (3.8% if patients who met the exclusion criteria were excluded from this group); and 22.4% of patients met at least one of the exclusion criteria. However, in patients who underwent ablation, 42.6% of patients would not have qualified for the trial (e.g. 8.9% of patients failed to meet the inclusion criterion, 6.5% of patients failed more than two AADs, 19.7% of patients failed amiodarone, and 7.3% of patients had other arrhythmias mandating AADs; [Supplementary material online, Table S2](#)). After propensity score weighting, patients treated with ablation and those treated with medical therapy alone

were identical on 90 dimensions ([Supplementary material online, Tables S3–S6](#)).

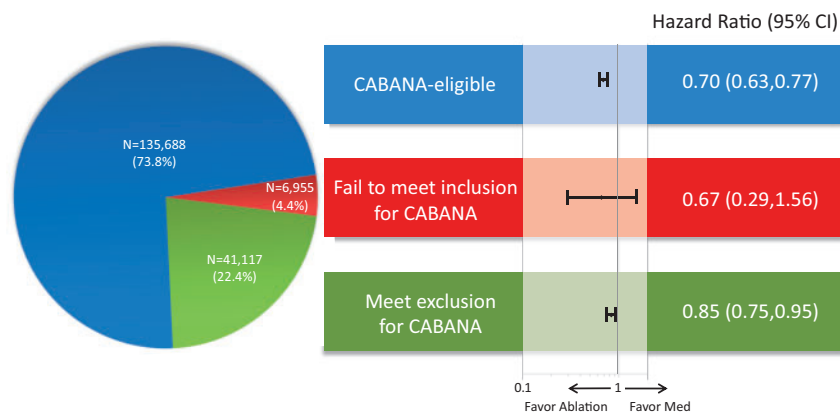
Outcomes

Patients were followed for a mean of 2.3 ± 1.7 years in the ablated patients (median 1.9 years) and 2.1 ± 1.7 years in the drug-treated patients (median 1.7 years). Ablation was associated with a reduction in the composite outcome of all-cause mortality, stroke, major bleeding, and cardiac arrest compared with medical therapy alone [4.51 vs. 6.07 events per 100 person years; hazard ratio (HR) 0.75, 95% confidence interval (CI) 0.70–0.81; $P < 0.001$] (Table 2 and Figure 1). The risk reduction associated with ablation was greatest in patients who

Table 3 Outcomes in propensity score-weighted patients stratified by trial eligibility

	Number of events	Person years	Event rate	Number of events	Person years	Event rate	Absolute reduction in event rate (95% CI)	Hazard ratio (95% CI)	P-value
Trial eligible	Drug treated (N = 128 781)			Ablated (N = 6907)					
Composite	527	9454	5.57	388	10 105	3.84	1.73 (1.32–2.14)	0.70 (0.63–0.77)	<0.001
All-cause mortality	312	9811	3.18	200	10 499	1.90	1.27 (0.98–1.56)	0.60 (0.53–0.69)	<0.001
Ischaemic stroke	86	9698	0.88	50	10 436	0.48	0.40 (0.26–0.55)	0.56 (0.43–0.73)	<0.001
Major bleeding	185	9558	1.94	192	10 168	1.89	0.05 (-0.22 to 0.32)	1.00 (0.87–1.16)	0.99
Cardiac arrest	30	9809	0.31	13	10 497	0.12	0.19 (0.11–0.26)	0.41 (0.24–0.69)	0.001
Failed to meet inclusion	N = 6130			N = 825					
Composite	—	—	0.89	—	—	0.58	0.31 (-0.22 to 0.84)	0.67 (0.29–1.56)	0.35
All-cause mortality	—	—	0.37	—	—	0.03	0.34 (0.19–0.50)	0.08 (0.01–0.58)	0.01
Ischaemic stroke	—	—	0.14	—	—	0.21	-0.07 (-0.39 to 0.26)	1.53 (0.27–8.69)	0.63
Major bleeding	—	—	0.39	—	—	0.34	0.06 (-0.33 to 0.45)	0.90 (0.30–2.73)	0.85
Cardiac arrest	—	—	0.04	—	—	0.00	0.04 (-0.01 to 0.09)	—	—
Met exclusion criterion	N=36 817			N=4300					
Composite	314	3678	8.53	279	3896	7.16	1.37 (0.46–2.27)	0.85 (0.75–0.95)	0.01
All-cause mortality	205	3867	5.30	169	4095	4.13	1.18 (0.50–1.85)	0.78 (0.67–0.91)	0.001
Ischaemic stroke	48	3806	1.27	31	4064	0.77	0.50 (0.19–0.80)	0.62 (0.45–0.87)	0.01
Major bleeding	97	3734	2.61	115	3926	2.94	-0.33 (-0.87 to 0.22)	1.16 (0.96–1.40)	0.13
Cardiac arrest	18	3865	0.47	14	4092	0.35	0.12 (-0.08 to 0.32)	0.76 (0.45–1.30)	0.32

Event rate was calculated as number of events per 100 person-years. To maintain de-identification, OptumLabs does not allow researchers to disclose the number of events when the number is 10 or fewer. Medicare data have similar requirements.



Take home figure Hazard ratio for primary outcome stratified by potential trial-eligibility in the overall population (N = 183 760).

met CABANA eligibility criteria (3.84 vs. 5.57 events per 100 person years; HR 0.70, 95% CI 0.63–0.77; $P < 0.001$; *Table 3* and *Take home figure*). In patients who failed to meet inclusion criteria, the overall event rates risks were low, and there was no significant relationship with ablation (0.58 vs. 0.89 events per 100 person years; HR 0.67, 95% CI 0.29–1.56; $P = 0.35$). In patients who met one of the trial criteria for exclusion, the risk reduction was more modest (7.16 vs. 8.53 events per 100 person years; HR 0.85, 95% CI 0.75–0.95; $P = 0.01$).

The treatment effect was not statistically different between patients who were eligible for the trial and those who failed to meet the inclusion criterion but was statistically different between patients who were eligible for the trial and those who met one of the exclusion criteria ($P = 0.01$ for interaction).

In all the cohorts except those who met at least one of the exclusion criteria, the curves continue to diverge over time so the proportional hazard assumption does not hold (*Figure 1*), so the HRs need

to be interpreted as an average effect over the observed times.^{23,24} Because of this, we also provided the cumulative risks and HRs at different time points to facilitate the interpretation of the effects over time (Supplementary material online, Table S7). The results for secondary outcomes were largely consistent with the primary outcome except that ablation was not associated with the risk of major bleeding in any of the cohorts (Tables 2 and 3).

Sensitivity analyses

All the results from the sensitivity analyses were consistent with the main findings (Supplementary material online, Tables S8–S10). Among trial-eligible patients, ablation was associated with a lower risk of the primary outcome in all subgroups, but the reduction was greater in patients under 65 years, men, patients with a low CHA₂DS₂-VASc score, and patients without heart failure. Among patients with trial exclusions, the reduction was greater in patients under 75 years, patients with a low CHA₂DS₂-VASc score, and patients without prior thromboembolism (Supplementary material online, Figures S3–S5). The sensitivity analysis to explore the impact of protocol deviation found that when there was crossover between treatment arms, the risk reduction associated with ablation would be smaller (HR 0.85, 95% CI 0.79–0.92; $P < 0.001$; Supplementary material online, Table S11).

The results using one-to-one propensity score matching (instead of overlap weights) were very similar to the primary results. For the composite endpoint, the HR was 0.70 (95% CI 0.64–0.77, $P < 0.001$) for ablation vs. medical therapy in the overall cohort, 0.62 (95% CI 0.55–0.70, $P < 0.001$) in patients who would be eligible for CABANA, 0.55 (95% CI 0.22–1.41, $P = 0.21$) in patients who failed to meet inclusion for CABANA, and 0.85 (95% CI 0.74–0.97, $P = 0.02$) in patients who would be excluded from CABANA. There were no significant relationships between ablation and any of the falsification endpoints (Supplementary material online, Table S12). In a test to assess the potential impact of an unmeasured confounder, we found that an unmeasured confounder could explain the association between ablation and the primary outcome only if the confounder were related to a substantially increased risk of the outcome (e.g. a HR > 2) and if there were substantial imbalance in its prevalence (e.g. 5% of patients in the treatment group had this characteristic vs. 50% of patients in the control group had; Supplementary material online, Figure S6).

Discussion

In this large cohort of 183 760 patients with AF encountered in routine practice, three in four would have met the CABANA trial enrolment criteria. Ablation was associated with a reduction in the primary CABANA endpoint, particularly in patients who met trial eligibility criteria. Our primary findings are consistent with the 'treatment-received' and 'per protocol' analyses of CABANA recently presented at Heart Rhythm Society, suggesting that patients who undergo ablation in practice have similar outcomes, relative to medical therapy, as those who underwent ablation within the trial. However, in patients who failed to meet the inclusion criterion, i.e. young patients without stroke risk factors, the risk was low and there was no significant relationship with ablation. Among patients who

met at least one of the trial exclusion criteria, there was a lesser but statistically significant reduction associated with ablation.

The 2016 ESC guidelines recommend AF ablation to improve symptoms and quality of life, and the class of recommendations ranges from I to IIa depending on the AF pattern and prior failure of AADs.²⁵ There are no indications for ablation to reduce major adverse cardiovascular events or mortality. However, AF is associated with considerable morbidity and mortality—nearly half of patients in one Medicare study died within five years of diagnosis.²⁶ In light of the evidence from this and other observational studies as well as RCTs including CASTLE-AF^{7–10} (and perhaps secondary analyses of CABANA), the role of ablation in reducing morbidity and mortality may warrant consideration for selected patients.

A major contribution of the current study is that it provides evidence to guide practice regarding how to select patients most likely to benefit from ablation in terms of cardiovascular risk reduction. The demonstration of stronger treatment effect in trial-eligible patients in our study suggests that the CABANA eligibility criteria could be used as a guide to patient selection. Although most patients with AF in practice would have qualified for CABANA, nearly half of the patients who actually undergo ablation would not. Such patients appear to have risk profiles towards the extremes of the disease continuum—many are young patients without stroke risk factors, while others have previously failed amiodarone or more than two AADs, or have other concomitant arrhythmias that would have excluded them from CABANA. As recommended in the guidelines, ablation may still be a reasonable strategy to relieve symptoms for many of these patients; however, our data suggest that the expected cardiovascular risk reduction may be more modest. Furthermore, subgroup analyses of the current study could also inform physicians as well as future guideline writing committees when making recommendations for populations not well represented in clinical trials.

In addition to the observational nature of our study, there are other key differences with the CABANA trial. First, while all patients in the ablation arm of our study underwent ablation (and no patients in the drug arm did), protocol deviations are common in clinical trials, particularly those like CABANA which are conducted over many years and involve treatments that are already in widespread clinical use. As such, the current study is more comparable to an 'as-treated' rather than the primary 'intention-to-treat' analysis of the trial. In a sensitivity analysis in our study, the HR for the primary endpoint increased from 0.70 to 0.85 when accounting for potential crossover between treatment arms (numerically similar to the primary CABANA ITT analysis). Second, trial participants tend to have better adherence to other medical therapies (e.g. oral anticoagulants) than do patients in routine practice and event rates are often lower in trials than in practice. Therefore, it may be more difficult to detect further risk reduction with ablation on top of optimal medical therapy within a trial, compared with an observational dataset in which guideline-directed medical therapy is under-used. Third, the distribution of patient characteristics (e.g. age and comorbidities) differ in our cohort and the trial. Fourth, the CABANA examined ablation outcomes in a global trial, whereas our cohort is limited to the US. We await the full publication of the trial before assessing whether there are important differences in outcome by region and whether this may affect the generalizability of our study to countries outside the US.

This study has several limitations. First, as an observational study, it is subject to residual confounding even after careful adjustment. However, the groups were identical on 90 dimensions were unlikely to substantially differ in other aspects, since many of the measured characteristics are highly correlated with unmeasured ones. For example, age, valvular heart disease, hypertension, previous AADs, and cardioversion are associated with unmeasured characteristics, such as left atrial diameter, AF pattern, or burden. Furthermore, the test of falsification endpoints provided some reassurance that there was no evidence for substantial residual confounding.

Second, this study relied on administrative data to ascertain baseline characteristics and clinical outcomes, which can be subject to misclassification. However, it is unlikely there are any systematic ascertainment differences between the two treatment groups, and thus, the misclassification should be non-differential. The diagnosis and procedure codes used in this study have demonstrated good performance in validation studies with positive predictive values around 90%.^{27–34}

Third, the non-significant relationship with ablation observed in patients who failed to meet the CABANA inclusion criteria could be due to the small number of patients and low event rates. In fact, the relatively risk reduction in this group (HR 0.67) was not significantly different from the relatively risk reduction in the trial-eligible group (HR 0.70). However, since the absolute risk is low, the magnitude of cardiovascular reduction with ablation is likely small in this group of patients.

In conclusion, three in four patients with AF in routine clinical practice would qualify for the CABANA trial. Similarly to the ‘treatment-received’ and ‘per protocol’ analyses of the CABANA trial itself, our study demonstrated an association between ablation and a reduction in the composite endpoint of all-cause mortality, stroke, major bleeding, or cardiac arrest, compared with medical therapy alone. The risk reduction was greatest among patients who met CABANA eligibility criteria but was absent in lower-risk patients who failed to meet the inclusion criterion and was more modest among those with trial exclusions.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Acknowledgements

X.Y. analysed the data and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Funding

National Institutes of Health, National Heart, Lung, and Blood Institute (NIH R21 HL140205 ‘Pairing Observational and Patient-level Clinical Trial Data to Assess Cardiovascular Risk Reduction with Catheter Ablation for Atrial Fibrillation’) to P.A.N. and X.Y. and Mayo Clinic Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery to P.A.N., N.D.S., and X.Y.

Conflict of interest: B.J.G. reports receiving consulting fees from Boston Scientific and Medtronic; J.P.P. reports receiving funding for clinical research from Abbott Medical and Boston Scientific and consulting fees

from Johnson & Johnson, Medtronic, and Sanofi; D.L.P. reports grants from NIH, St. Jude Medical Foundation and Corporation, Biosense Webster, Inc., Medtronic Corp., and Boston Scientific Corp., during the conduct of the study; other from Abbott, Aperture Diagnostics, Biosense Webster, Inc., Boston Scientific Corp., CardioFocus, Medtronic, Inc., St. Jude Medical, Spectrum Dynamics, and MediaSphere Medical; grants from Abbott, Biosense Webster, Boston Scientific, and Cardiolsight; grants and other from Medtronic, Inc., grants from NIH, grants and other from St. Jude Medical, grants and other from Siemens, grants and other from Thermedical, other from Other, other from Wiley & Sons, other from Oxford, other from St. Jude Medical, outside the submitted work. No other potential conflict of interest relevant to this article was reported.

References

- Chugh SS, Havmoeller R, Narayanan K, Singh D, Rienstra M, Benjamin EJ, Gillum RF, Kim YH, McAnulty JH Jr, Zheng ZJ, Forouzanfar MH, Naghavi M, Mensah GA, Ezzati M, Murray CJ. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation* 2014;**129**:837–847.
- Benjamin EJ, Wolf PA, D’Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of atrial fibrillation on the risk of death the Framingham Heart Study. *Circulation* 1998;**98**:946–952.
- Kannel WB, Wolf PA, Benjamin EJ, Levy D. Prevalence, incidence, prognosis, and predisposing conditions for atrial fibrillation: population-based estimates. *Am J Cardiol* 1998;**82**:2N–9N.
- Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991;**22**:983–988.
- Reynolds MR, Walczak J, White SA, Cohen DJ, Wilber DJ. Improvements in symptoms and quality of life in patients with paroxysmal atrial fibrillation treated with radiofrequency catheter ablation versus antiarrhythmic drugs. *Circ Cardiovasc Qual Outcomes* 2010;**3**:615–623.
- Ganesan AN, Shipp NJ, Brooks AG, Kuklik P, Lau DH, Lim HS, Sullivan T, Roberts-Thomson KC, Sanders P. Long-term outcomes of catheter ablation of atrial fibrillation: a systematic review and meta-analysis. *J Am Heart Assoc* 2013;**2**: e004549.
- Bunch TJ, May HT, Bair TL, Weiss JP, Crandall BG, Osborn JS, Mallender C, Anderson JL, Muhlestein BJ, Lappe DL, Day JD. Atrial fibrillation ablation patients have long-term stroke rates similar to patients without atrial fibrillation regardless of CHADS₂ score. *Heart Rhythm* 2013;**10**:1272–1277.
- Noseworthy PA, Kapa S, Deshmukh AJ, Madhavan M, Van Houten H, Haas LR, Mulpuru SK, McLeod CJ, Asirvatham SJ, Friedman PA, Shah ND, Packer DL. Risk of stroke after catheter ablation versus cardioversion for atrial fibrillation: a propensity-matched study of 24,244 patients. *Heart Rhythm* 2015;**12**:1154–1161.
- Friberg L, Tabrizi F, Englund A. Catheter ablation for atrial fibrillation is associated with lower incidence of stroke and death: data from Swedish health registries. *Eur Heart J* 2016;**37**:2478–2487.
- Marrouche NF, Brachmann J, Andresen D, Siebels J, Boersma L, Jordaens L, Merkely B, Pokushalov E, Sanders P, Proff J, Schunkert H, Christ H, Vogt J, Bänsch D. Catheter ablation for atrial fibrillation with heart failure. *N Engl J Med* 2018;**378**:417–427.
- Steinberg BA, Holmes DN, Ezekowitz MD, Fonarow GC, Kowey PR, Mahaffey KW, Naccarelli G, Reiffel J, Chang P, Peterson ED, Piccini JP. Rate versus rhythm control for management of atrial fibrillation in clinical practice: results from the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF) registry. *Am Heart J* 2013;**165**:622–629.
- Kirchhof P, Ammentorp B, Darius H, De Caterina R, Le Heuzey JY, Schilling RJ, Schmitt J, Zamorano JL. Management of atrial fibrillation in seven European countries after the publication of the 2010 ESC Guidelines on atrial fibrillation: primary results of the PREvention of thromboembolic events–European Registry in Atrial Fibrillation (PREFER in AF). *Europace* 2014;**16**:6–14.
- Packer DL, Mark DB, Robb RA, Monahan K, Bahnson T, Moretz K, Poole J, Mascette A, Rosenberg Y, Jeffries N. Catheter ablation versus antiarrhythmic drug therapy for atrial fibrillation (CABANA) Trial: study rationale and design. *Am Heart J* 2018;**199**:192–199.
- Shah RU, Freeman JV, Shilane D, Wang PJ, Go AS, Hlatky MA. Procedural complications, rehospitalizations, and repeat procedures after catheter ablation for atrial fibrillation. *J Am Coll Cardiol* 2012;**59**:143–149.
- Reynolds MR, Zimetbaum P, Josephson ME, Ellis E, Danilov T, Cohen DJ. Cost-effectiveness of radiofrequency catheter ablation compared with antiarrhythmic drug therapy for paroxysmal atrial fibrillation. *Circ Arrhythm Electrophysiol* 2009;**2**: 362–369.

16. Wallace PJ, Shah ND, Dennen T, Bleicher PA, Bleicher PD, Crown WH. Optum Labs: building a novel node in the learning health care system. *Health Aff (Millwood)* 2014;**33**:1187–1194.
17. Optum. *Optum Research Data Assets* 2015.
18. Li F, Morgan KL, Zaslavsky AM. Balancing covariates via propensity score weighting. *J Am Stat Assoc* 2018;**113**:390–400.
19. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc* 1999;**94**:496–509.
20. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994;**81**:515–526.
21. Prasad V, Jena AB. Prespecified falsification end points: can they validate true observational associations? *JAMA* 2013;**309**:241–242.
22. Lin DY, Psaty BM, Kronmal RA. Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics* 1998;**54**:948–963.
23. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag New York: Springer Science & Business Media; 2013.
24. Weintraub WS, Grau-Sepulveda MV, Weiss JM, O'Brien SM, Peterson ED, Kolm P, Zhang Z, Klein LW, Shaw RE, McKay C, Ritzenthaler LL, Popma JJ, Messenger JC, Shahian DM, Grover FL, Mayer JE, Shewan CM, Garratt KN, Moussa ID, Dangas GD, Edwards FH. Comparative effectiveness of revascularization strategies. *N Engl J Med* 2012;**366**:1467–1476.
25. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, Castella M, Diener HC, Heidbuchel H, Hendriks J, Hindricks G, Manolis AS, Oldgren J, Popescu BA, Schotten U, Van Putte B, Vardas P, Agewall S, Camm J, Baron Esquivias G, Budts W, Carerj S, Casselman F, Coca A, De Caterina R, Deftereos S, Dobrev D, Ferro JM, Filippatos G, Fitzsimons D, Gorennek B, Guenoun M, Hohnloser SH, Kolh P, Lip GY, Manolis A, McMurray J, Ponikowski P, Rosenhek R, Ruschitzka F, Savelieva I, Sharma S, Suwalski P, Tamargo JL, Taylor CJ, Van Gelder IC, Voors AA, Windecker S, Zamorano JL, Zeppenfeld K. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Europace* 2016;**18**:1609–1678.
26. Piccini JP, Hammill BG, Sinner MF, Hernandez AF, Walkey AJ, Benjamin EJ, Curtis LH, Heckbert SR. Clinical course of atrial fibrillation in older adults: the importance of cardiovascular events beyond stroke. *Eur Heart J* 2014;**35**:250–256.
27. Kumamaru H, Judd SE, Curtis JR, Ramachandran R, Hardy NC, Rhodes JD, Safford MM, Kissela BM, Howard G, Jalbert JJ. Validity of claims-based stroke algorithms in contemporary Medicare data reasons for geographic and racial differences in stroke (REGARDS) study linked with Medicare claims. *Circ Cardiovasc Qual Outcomes* 2014;**7**:611–619.
28. Tirschwell DL, Longstreth W. Validating administrative data in stroke research. *Stroke* 2002;**33**:2465–2470.
29. Kokotailo RA, Hill MD. Coding of stroke and stroke risk factors using international classification of diseases, revisions 9 and 10. *Stroke* 2005;**36**:1776–1781.
30. Jensen PN, Johnson K, Floyd J, Heckbert SR, Carnahan R, Dublin S. Identifying atrial fibrillation from electronic medical data: a systematic review. *Pharmacoepidemiol Drug Saf* 2012;**21**:141–147.
31. Noseworthy PA, Yao X, Abraham NS, Sangaralingham LR, McBane RD, Shah ND. Direct comparison of dabigatran, rivaroxaban, and apixaban for effectiveness and safety in nonvalvular atrial fibrillation. *Chest* 2016;**150**:1302–1312.
32. Yao X, Shah ND, Sangaralingham LR, Gersh BJ, Noseworthy PA. Non-vitamin K antagonist oral anticoagulant dosing in patients with atrial fibrillation and renal dysfunction. *J Am Coll Cardiol* 2017;**69**:2779–2790.
33. Fan J, Arruda-Olson AM, Leibson CL, Smith C, Liu G, Bailey KR, Kullo IJ. Billing code algorithms to identify cases of peripheral artery disease from administrative data. *J Am Med Inform Assoc* 2013;**20**:e349–e354.
34. Yao X, Tangri N, Gersh BJ, Sangaralingham LR, Shah ND, Nath KA, Noseworthy PA. Renal outcomes in anticoagulated patients with atrial fibrillation. *J Am Coll Cardiol* 2017;**70**:2621–2632.