Application of artificial intelligence to analyze data from randomized controlled trials: an example from DECAAF II

Mekhael, M.; Feng, H.; Akoum, N.; Sohns, C.; Sommer, P.; Mahnkopf, C.; ... ; Marrouche, N.

HeartRhythm

The Official Journal of the Heart Rhythm Society, The Cardiac Electrophysiology Society, and The Pediatric & Congenital Electrophysiology Society

Heart Rhythm Society SM

Check for updates

Application of artificial intelligence to analyze data from randomized controlled trials: An example from DECAAF II

Mario Mekhael, MD,[1,*] Han Feng, PhD,[1,*] Nazem Akoum, MD, FHRS,[2] Christian Sohns, MD,[3] Philipp Sommer, MD, FHRS,[3] Christian Mahnkopf, MD,[4] Eugene Kholmovski, PhD,[5] Jeroen J. Bax, MD,[6] Prashanthan Sanders, MD, FHRS,[7] Christopher McGann, MD,[8] Francis Marchlinski, MD, FHRS,[9] Moussa Mansour, MD, FHRS,[10] Gerhard Hindricks, MD,[11] David Wilber, MD, FHRS,[12] Hugh Calkins, MD, FHRS,[5] Pierre Jais, MD,[13] Hadi Younes, MD,[1] Ala Assaf, MD,[1] Charbel Noujaim, MD,[14] Chanho Lim, MS,[1] Chao Huang, PhD,[1] Amitabh Pandey, MD,[1] Oussama Wazni, MD,[14] Nassir Marrouche, MD, FHRS[1]

ABSTRACT

BACKGROUND Causal machine learning (ML) provides an efficient way of identifying heterogeneous treatment effect groups from hundreds of possible combinations, especially for randomized trial data.

OBJECTIVE The aim of this paper is to illustrate the potential of applying causal ML on the DECAAF II trial data. We proposed a causal ML model to predict the treatment response heterogeneity.

METHODS We applied causal tree learning to the DECAAF II trial data as an example of real applications, identifying subgroups that may be superior when subject to one of the treatments over the other through an easily interpretable process. For each subgroup identified, the characteristics were summarized, and the relationship between treatment arms and risk for recurrence of atrial tachyarrhythmia (aTA) among subjects was assessed.

RESULTS Causal tree learning demonstrated that, among all the preablation predictors, dividing subgroups according to age, with a cutoff of 58 years, provides the most heterogeneous subgroups in response to fibrosis-guided ablation in addition to pulmonary vein isolation (PVI) compared with PVI alone. The difference in the risk of recurrence of aTA between 2 treatments was nonsignificant in older patients (hazard ratio [HR] 1.06; 95% confidence interval [CI] 0.77–1.47; $P = .72$). However, among the younger patients, the risk of aTA recurrence was significantly lower in the fibrosis-guided ablation group compared with PVI-only (HR 0.50; 95% CI 0.28–0.90); $P = .02$).

CONCLUSION Applying causal ML on random controlled trial datasets helped us identify groups of patients that profited from the treatment of interest in an efficient and unbiased manner.

KEYWORDS Atrial cardiomyopathy; Atrial fibrillation; Atrial remodeling; Catheter ablation; Causal machine learning

(Heart Rhythm 2025;22:3035–3041)

Introduction

Randomized clinical trials (RCTs) are one of the major sources of knowledge accumulation and guidelines in medicine. RCTs have been intensively conducted and have changed the guidelines in the past century. Besides, one of the major uses of RCT data is to generate further hypotheses from

From the [1]Cardiology Department, Tulane University School of Medicine, New Orleans, Louisiana, [2]Cardiology Department, University of Washington Medical Center, Seattle, Washington, [3]Heart and Diabetes Center NRW, Bad Oeynhausen, Germany, [4]Klinikum Coburg, Coburg, Germany, [5]Cardiology Department, Johns Hopkins Medicine, Baltimore, Maryland, [6]Cardiology Department, Leiden University Medical Center, Leiden, The Netherlands, [7]Cardiology Department, Adelaide Medical School, Adelaide, Australia, [8]Swedish Center, Seattle, Washington, [9]Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania, [10]Cardiology Department, Massachusetts General Hospital, Boston, Massachusetts, [11]Cardiology Department, Leipzig University, Leipzig, Germany, [12]Cardiology Department, Loyola University Chicago, Chicago, Illinois, [13]Cardiology Department, Segalen University, Bordeaux, France, and [14]Cleveland Clinic, Cleveland, Ohio.

**These authors contributed equally to this work.*

https://doi.org/10.1016/j.hrthm.2025.01.008

secondary analyses. RCT data are usually very clean, although sometimes not generalizable, and researchers usually appreciate conducting secondary analyses of such data. Nevertheless, the potential of RCT data has still not been fully explored. One good example is identifying subgroups that respond to treatment differently in an unbiased, nonhypothesis-driven manner. Identifying such subpopulations can provide several benefits in personalizing clinical management and prioritizing a treatment approach over the other. Traditional statistical methods have long been used to analyze RCT data, focusing on associations between treatment and outcomes through techniques such as generalized linear models and Cox models. These methods are well suited for straightforward, hypothesis-driven analyses but can fall short when dealing with complex data formats or when the hypothesis is not predetermined.

Causal machine learning (ML), as a group of methods that combine causal inference from statistics with ML models, has been developed recently and has great potential to be applied in many scenarios to achieve the goal we discussed earlier. It employs techniques through estimating the conditional average treatment effect (CATE) while integrating traditional ML models to predict the treatment response heterogeneity, such as causal tree learning,[1] causal forests,[2] and double ML[3] to handle complex real-world clinical data. It can consider heterogeneous treatment effects (HTEs) among a wide range of variables and is of particular interest in fields such as electrophysiology, in which predictors of outcomes, such as recurrence of atrial fibrillation (AF), are still not well elucidated, especially in advanced disease such as persistent AF.

Catheter ablation (CA) has emerged as an effective treatment for AF, but freedom from arrhythmia recurrence remains low.[3,4] Several ablation strategies have been employed to individualize management of AF and to increase success of ablation, especially in patients with persistent AF. Despite improvements in ablation technology and techniques as well as imaging modalities, recurrence of AF following CA remains an issue, and pulmonary vein isolation (PVI) is still considered the cornerstone of AF ablation. Delayed enhancement magnetic resonance imaging (DE-MRI) has been used to evaluate atrial fibrosis and cardiomyopathy, which is known to be associated with treatment failure and postablation recurrence of arrhythmia.[5] The level of atrial fibrosis has also been shown to correlate with severity and progression of AF disease.[6] The DECAAF II (Efficacy of LGE-MRI–Guided Fibrosis Ablation vs Conventional Catheter Ablation of Atrial Fibrillation) trial studied the hypothesis that targeting areas of fibrosis based on DE-MRI, in addition to PVI might improve arrhythmia outcomes compared with PVI alone.[7] However, the intention-to-treat (ITT) analysis did not find a significant difference between the 2 treatment arms.[8] Other studies showed a significant improvement in freedom from recurrence of AF in patients who received substrate modification in addition to PVI.[9] Taking these observations into consideration, targeting extra-PV targets in persistent patients with AF remains a topic of interest and controversy.

In this study, we use the DECAAF II trial data as an example to illustrate how causal ML models can be applied to identify subpopulations that benefit more from either fibrosis-guided ablation in addition to PVI or PVI alone. To our knowledge, this is the first paper to apply causal ML to a persistent AF RCT in order to personalize management in this advanced disease.

Abbreviations

AF: atrial fibrillation

CA: catheter ablation

CATE: conditional average treatment effect

HTE: heterogenous treatment effect

LA: left atrium

ML: machine learning

PVI: pulmonary vein isolation

RCT: Randommized Clinical Trial

Methods

Causal machine learning

The development of a causal ML model fundamentally aims to identify and quantify the causal relationships among variables, primarily leveraging the concept of the CATE from causal inference from statistical theory. This methodology diverges from traditional ML models that aim to predict the outcomes accurately. Instead, it focuses on estimating the CATE, which measures the expected effect of a treatment on an outcome variable conditional on a set of observed covariates. Formally, the CATE for a given set of covariates $X=x$ is defined as follows:

$$\tau(x) = \mathbb{E}[Y(1) - Y(0)|X = x],$$

in which $Y(1)$ and $Y(0)$ represent the potential outcomes under treatment and control conditions, respectively. The underlying assumptions for this kind of model include the ignorability assumption that treatment assignment is conditionally independent of the potential outcomes given the covariates: that is, $(Y(1), Y(0)) \perp T|X$, in which T represents the treatment and the overlap assumption that each unit has a positive probability of receiving both treatment and control: that is, $0 < P(T = 1|X = x) < 1$.

These 2 assumptions guarantee the appropriateness of the results by natural hold under RCT settings but do not necessarily hold for real-world data: for example, electronic health records. Thus, RCT data are a great fit for applying causal ML methods.

Common causal ML techniques include double ML, causal forests, causal tree learning, and so forth. They are developed under different ML models, targeted for different purposes of analysis. We chose causal tree learning and the DECAAF II trial data in this study as an example to show the potential and the power of the causal ML method on RCT trial data.

Study design

The DECAAF II trial (clinicatrials.gov NCT02529319) design and methods have been previously described.[7] This substudy aimed to examine preablation predictors that can favor fibrosis-guided ablation over PVI-only for persistent AF treatment, based on an improved primary outcome in terms of

freedom from recurrence of arrhythmia. To achieve that, we used an ML model—causal tree learning—to identify heterogeneous subgroups that respond differently to the 2 types of ablation procedures.[10,11] The whole population was randomly split into 2 equal groups for training and validation purposes. Details regarding the ML model and statistical analysis are provided in the statistical analyses section.

The endpoint of the study was the first confirmed recurrence of any atrial tachyarrhythmia (aTA: including AF, or atrial tachycardia) lasting for > 30 seconds after a 90-day blanking period, demonstrated by at least 2 consecutive 1-lead smartphone electrocardiographic (ECG) device tracings, or by one positive reading on a clinical 12-lead ECG tracing, ambulatory monitor, or if the patient underwent repeat ablation. The daily smartphone ECGs were intended as the primary method for assessing recurrence of arrhythmia, but clinical and ambulatory ECGs served as back-up methods. A core laboratory at the University of Washington adjudicated the ECG findings.

Causal tree learning

Causal tree learning partitions the data into distinct subgroups to estimate the CATE and finds the optimal division of the subgroups by minimizing the expected mean squared error (EMSE). Thus, we provided the causal tree learning model with a group of variables being considered for dividing subgroups and let the model choose how to separate the population to obtain the most heterogeneous groups that respond differently to the treatment. It provides insights into HTEs while preserving high interpretability, allowing for more personalized interventions and broader practice. Briefly, we constructed 1 causal tree based on a randomly sampled half of the overall trial population, whereas the other half, as validation set, further pruned the tree obtained to preserve the generalizability of the findings in our study. In total 29 variables of different types were taken into consideration for the subgroup identification, including baseline demographics, risk factors, baseline imaging parameters, ablation parameters, and medications: namely, age, body mass index, sex, congestive heart failure, hypertension, diabetes mellitus, hyperlipidemia, history of stroke, vascular disease, tobacco use, coronary artery disease, history of coronary artery bypass graft, history of cardioversion, history of failed antiarrhythmic therapy; mitral valve disease; rheumatic fever; left atrial (LA) volume; LA fibrosis, LA appendage volume; average impedance drop; average ablation contact force; mean ablation lesion time; max ablation temperature and max ablation power; and use of antiarrhythmics, angiotensin receptor blockers, statins, aspirin, and anticoagulants. The size of the leaves was set to be at least 15% of the total population to avoid overfitting.

Statistical analyses

Eight hundred and fifteen patients were included in this analysis. We applied causal tree learning to identify subgroups with substantially different HTEs.[10,11] Subgroups were obtained based on the results of causal tree learning. For each subgroup, baseline demographics were presented as mean ± standard deviation. Normality was assessed using the Shapiro-Wilk test. Continuous variables were compared among study groups using the Kruskal-Wallis test or analysis of variance (ANOVA), according to whether the normality assumption was violated or not. The Fisher or χ^2 tests were used to compare categorical variables among study groups. A uni- and multivariable Cox regression model was performed to assess the relationship between the treatment arm and risk of aTA recurrence among the subjects for each subgroup. Covariates clinically relevant to aTA recurrence were adjusted in the multivariable Cox model, including treatment arm, age, sex, body mass index (BMI), LA volume, baseline fibrosis, congestive heart failure, hypertension, history of stroke, tobacco use, mean ablation lesion time, and max ablation power. Proportional hazard assumptions of variables are checked by introducing interactions with time, and the violation was only found in gender. The study flow is summarized in Figure 1. All the analyses were conducted under R version 4.2.0 (R Foundation for Statistical Computing, Vienna, Austria). Data will be made available upon reasonable request.

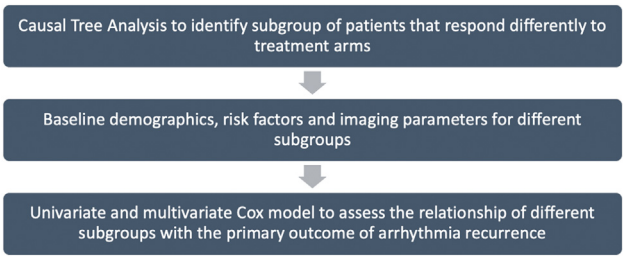

Figure 1
Summary of study analysis flow.

Results

Study subgroups and baseline characteristics

Evaluation of 815 participants from DECAAF II according to the causal tree learning demonstrated that age was the most important characteristic to separate the subpopulations, and there were 2 major subpopulations with different responses to treatments. The subgroups could be obtained by dividing the whole population according to age, with the cutoff being 58 years. Out of 815 patients, 575 participants were ≥58 years old, and 240 patients were younger (<58 years old). The median age of our population is 62.8 years. In the older cohort (≥58 years), 284 patients (49%) received fibrosis-guided ablation plus PVI vs 291 (51%) who received PVI only. In the younger cohort, 123 (51%) underwent fibrosis guided ablation plus PVI (Figure 2), whereas 117 (49%) subjects underwent PVI only. In patients <58 years old, the subgroup of fibrosis-guided ablation had more congestive heart failure (CHF) than the PVI-only subgroup (28.5% vs 15.4%, $P = .02$). All other differences were not significant (Table 1).

Primary endpoint

The event rate for the primary endpoint focusing on recurrence of aTA after ablation did not differ significantly between the

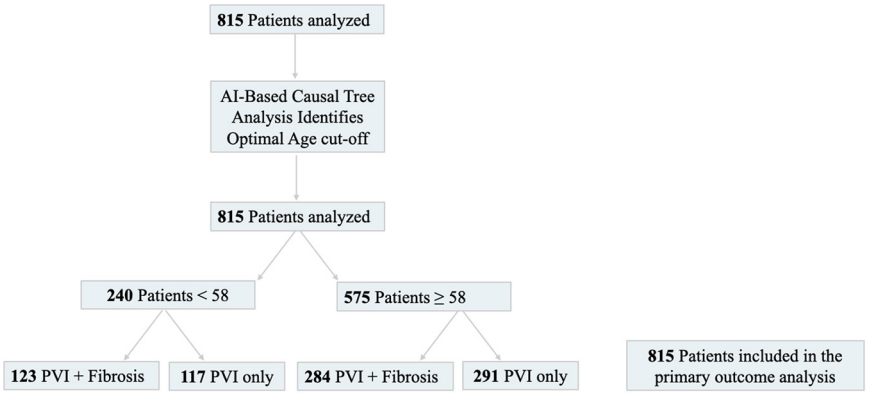

Figure 2
Study groups and flow chart.

fibrosis-guided group and the PVI-only group in both the younger and older cohort. In younger patients, the primary efficacy endpoint occurred in 36 of 123 patients (29.3%) in the fibrosis-guided group and in 45 of 117 (38.5%) in the PVI-only group (hazard ratio [HR] 0.73, 95% confidence interval [CI] 0.47–1.13; $P = .15$) (Supplemental Figure 1). In older patients, the primary efficacy endpoint occurred in 139 of 284 patients (48.9%) in the fibrosis-guided arm and in 143/291 (49.1%) in the PVI-only group (HR 1.03; 95% CI 0.81–1.30; $P = .84$).

After adjusting for covariates, the recurrence of aTA was similar for treatment arms in older patients (HR 1.06; 95% CI 0.77–1.47); $P = .72$) (Figure 3A). In contrast, recurrence of aTA was significantly lower in the fibrosis-guided group compared with the PVI-only group (HR 0.50; 95% CI 0.28–0.90); $P = .02$) in younger patients (Table 2, Figure 3A).

Moreover, when looking at different treatment arms, the adjusted Cox model in the fibrosis-guided ablation, younger age (HR 0.55; 95% CI 0.33–0.91); $P = .02$), female gender, and increased LA volume were all significant predictors of increased aTA after CA, whereas in the PVI-only arm, only LA volume was a significant predictor of aTA recurrence (HR 1.07, 95% CI 1.02–1.12; for every 10-mL increase; $P = .01$) (Supplemental Table S1). A schematic representation of all four Cox-analysis curves is shown in Supplemental Figure S2.

Discussion

The results of this subanalysis of DECAAF II trial demonstrate that applying an ML-based algorithm, particularly causal tree learning, allows us to identify a group of patients that might benefit from a therapy in an unbiased fashion. In this DECAAF II subanalysis, patients younger than 58 years derived more benefit from the addition of fibrosis-guided ablation than conventional PVI, whereas the same did not hold true for older patients. Younger patients who received fibrosis-guided ablation were 50% less likely to have recurrence of aTA compared

Table 1 Baseline characteristics of different study groups

	Age < 58 years			Age ≥ 58 years		
Characteristic number (%)	MRI-guided (n = 123)	PVI (n = 117)	*P value*	MRI-guided (n = 284)	PVI (n = 291)	*P* value
Baseline fibrosis levels	17.1% (7.1)	18.3% (7.7)	.26	19.5% (7.3)	19.2% (7.5)	.64
Age (years), mean (SD)	51.2 (6.2)	50.8 (5.7)	.40	66.3 (5.7)	66.9 (5.4)	.14
Female sex	13 (10.6%)	6 (5.1%)	.12	73 (25.7%)	77 (26.5%)	.84
Coronary artery disease	7 (5.7%)	7 (6.0%)	.92	45 (15.8%)	43 (14.8%)	.72
History of tobacco use	50 (40.7%)	50 (42.7%)	.74	95 (33.5%)	110 (37.8%)	.28
Mitral valve disease	5 (4.1%)	7 (6.0%)	.50	17 (6.0%)	18 (6.2%)	.92
Hyperlipidemia	26 (21.1%)	23 (19.7%)	.78	115 (40.5%)	115 (39.5%)	.81
Congestive heart failure	**35 (28.5%)**	**18 (15.4%)**	**.02**	52 (18.3%)	49 (16.8%)	.64
Hypertension	59 (48.0%)	50 (42.7%)	.42	183 (64.4%)	188 (64.6%)	.97
Diabetes mellitus	8 (6.5%)	7 (6.0%)	.87	30 (10.6%)	37 (12.7%)	.42
Stroke/transient ischemic attack/ thromboembolism	3 (2.4%)	6 (5.1%)	.27	32 (11.3%)	28 (9.6%)	.52
Antiarrhythmic medications	60 (48.8%)	50 (42.7%)	.35	134 (47.2%)	141 (48.5%)	.76
Vascular disease	10 (8.1%)	10 (8.5%)	.91	32 (11.3%)	29 (10.0%)	.61
Coronary artery bypass graft	0 (0.0%)	1 (0.9%)	.30	4 (1.4%)	7 (2.4%)	.38
Rheumatic fever	3 (2.4%)	1 (0.9%)	.34	4 (1.4%)	2 (0.7%)	.39
Cardioverted	101 (82.1%)	101 (86.3%)	.37	241 (84.9%)	240 (82.5%)	.44

Bold indicates statistically significant.
MRI = magnetic resonance imaging; PVI = pulmonary vein isolation; SD = standard deviation.

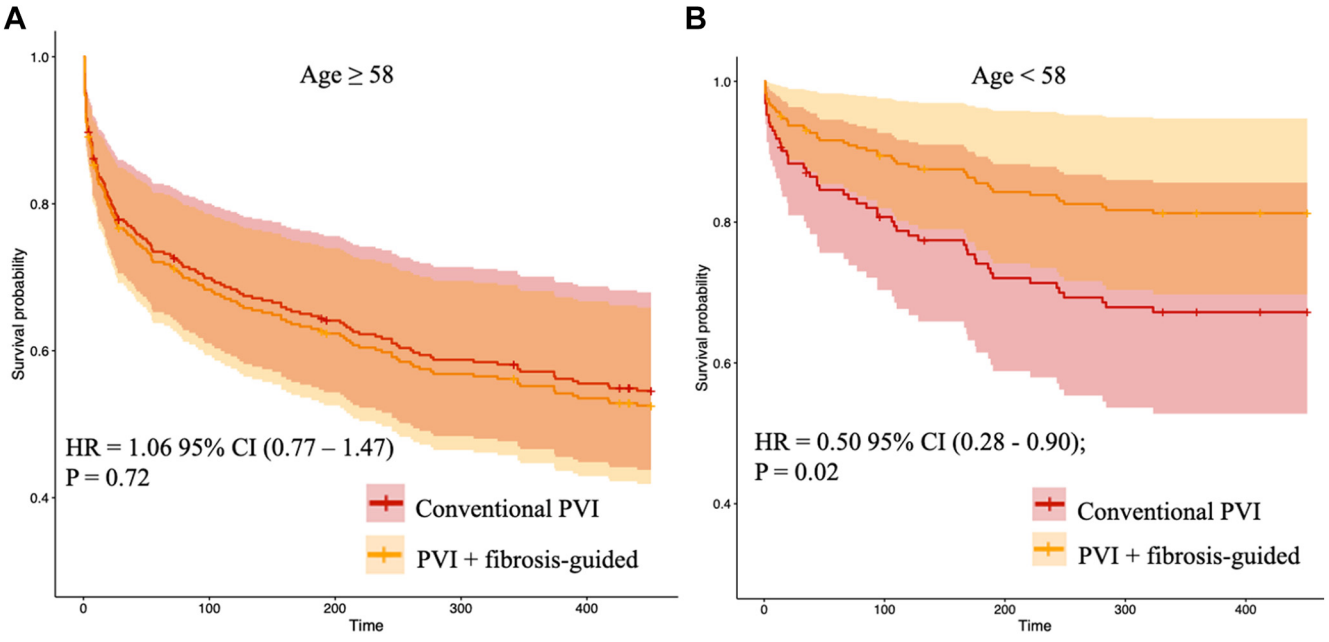

Figure 3
Cox hazard model survival curve analysis demonstrating differences between treatment arms in patients ≥58 years (**A**) and patients <58 years (**B**).

with those who received conventional PVI. In contrast, older patients had similar rates of recurrence regardless of treatment strategy.

Application of causal machine learning in medicine

The current gold standard for understanding the effect of interventions on clinical outcomes are RCTs. However, performing an RCT can be lengthy and difficult and requires preliminary evidence supporting the hypothesis. Previously, researchers resorted to secondary subanalyses for generation of hypotheses. Traditional subanalyses are usually underpowered and susceptible to many biases, such as estimation bias. Most importantly, these analyses test only 1 hypothesis at a time and fail to combine all factors and confounders that may affect treatment outcomes in a comprehensive way.[12]

On the other hand, causal ML methods compare multiple treatment effects in a population comprehensively through a data-driven fashion.[10,13] Causal ML allows for a hypothesis-free approach to study the data and variables.[14] These techniques also have the advantage of avoiding overfitting and maintaining valid hypothesis testing.[15] Specifically, causal ML models are especially useful when applied to RCT data, as their purpose is to evaluate outcome effect after administration of a certain treatment with minimum baseline characteristics differences.[16] One previous work in the field of electrophysiology has applied causal ML on a retrospective claims dataset and was able to identify patients with AF who would benefit from certain types of anticoagulation.[17] This helped generate a hypothesis to test prospectively.[17] Of note applying causal ML methods on a non-RCT dataset

Table 2 Predictors of AF recurrence using multivariable Cox regression model in the 2 study groups

	Age < 58 years		Age ≥ 58 years	
Variable (Ref)	HR [CI 95%]	*P* value	HR [CI 95%]	*P* value
PVI + fibrosis-guided ablation	**0.52 [0.28–0.90]**	**.02**	1.06 [0.77–1.47]	.72
Sex (male)	0.63 [0.26–1.55]	.32	**0.64 [0.44–0.93]**	**.02**
Left atrial volume[10]	**1.09 [1.001–1.18]**	**.047**	**1.06 [1.02–1.10]**	**.01**
Baseline fibrosis[10]	0.68 [0.43–1.07]	.10	1.12 [0.89–1.42]	.33
Congestive heart failure (yes)	**2.05 [1.10–3.82]**	**.02**	0.88 [0.57–1.38]	.58
Hypertension (yes)	0.64 [0.33–1.22]	.18	0.98 [0.68–1.40]	.90
Stroke/TIA (yes)	5.71 [0.63–51.9]	.12	0.75 [0.43–1.31]	.31
Tobacco (yes)	1.59 [0.86–2.96]	.14	1.01 [0.72–1.44]	.93
BMI[10]	1.22 [0.76–1.94]	.41	**0.74 [0.56–0.995]**	**.046**
Mean max power[10]	0.77 [0.51–1.16]	.22	**1.02 [1.01–1.03]**	**.005**
Mean lesion duration[10]	1.05 [0.69–1.61]	.81	1.00 [0.79–1.27]	.98

Bold indicates statistically significant.
BMI = body mass index; CI = confidence interval; HR = hazard ratio; TIA = transient ischemic attack.

may need more careful preprocessing and in some situations may not be valid.

In this work, we used causal tree learning on the DECAAF II trial data to evaluate subpopulations of persistent AF patients that benefit from PVI + fibrosis CA to personalize management. Advanced disease processes such as persistent AF are very heterogeneous, and outcomes differ on an individual level. It has been difficult to predict patients at high risk for recurrence of AF or other clinical outcomes in the population with persistent AF. Whereas previous studies focused on traditional statistics, and testing a single hypothesis related to demographics, risk factors, or imaging parameters, in this study, we examined treatment and outcome effects of a wide range of variables consisting of demographics, risk factors, imaging, and ablation parameters. Patients aged younger than 58 were found to have the most benefit of undergoing PVI + fibrosis ablation compared with PVI alone.

Fibrosis-guided ablation improves outcome in younger patients

Aging leads to increased atrial fibrosis and myopathy, which can lead to the emergence of more triggers outside the pulmonary veins. This, in turn, reduces the success rates of PVI in older patients.[18–21] This age-related effect appears particularly pronounced in patients with persistent or long-standing permanent AF, as they seem to benefit less from CA than patients with paroxysmal AF and less atrial fibrosis.[6,22] Several prospective trials have been designed to test this hypothesis with mixed results. Although most of these trials yielded negative results,[7,23–25] others showed significant improvement with substrate modification.[9] The ALICIA (Guided Fibrosis Ablation for the Treatment of Atrial Fibrillation) trial examined MRI-guided fibrosis ablation in addition to PVI alone and showed no significant differences between the 2 strategies.[23] However, patients in the ALICIA trial had very low fibrosis levels, with 98% of patients in Utah stages 1 and 2 of fibrosis (only 1 patient in Utah 4 and no patients in Utah 3).[23] The authors[23] concluded that the lack of atrial disease might have contributed to the lack of superiority. In the DECAAF II population, patients had a relatively higher baseline fibrosis (19%), making it a more suitable population to test the hypothesis with more ablation targets.[7] The main study showed no difference between both treatment arms.[8] In this subanalysis, younger patients seemed to benefit more from fibrosis-guided ablation, possibly because of improved remodeling after the procedure. Even though baseline fibrosis was significantly higher in older patients, recurrence in the fibrosis-guided ablation group seemed to be driven by age, female sex, and LA volume rather than underlying substrate (Supplemental Table S1). This further emphasizes age as an independent factor in determining treatment success after fibrosis-guided ablation adjunctive to PVI. A common assumption across the negative studies was the view of atrial myopathy as a homogenous process, suggesting a uniform response to ablation. The findings from this study challenge this assumption and shed light on the variability in treatment outcomes.

Limitations

This study is an ML-based retrospective analysis of DECAAF II and is not part of the intention-to-treat analysis. The trial was not powered for this analysis. The findings of this study should be interpreted as hypothesis generating and need to be further confirmed in future prospective trials. In addition, the difference was noted in adjusted analysis, and crude comparison of outcomes was not statistically different between patients <58 years of age and patients ≥58 years of age. Measured confounders include CHF, which was significantly differentially distributed in patients younger than 58 years. The low sample size in each subgroup raises the concern for unmeasured bias and should be kept in mind when interpreting the results. Moreover, other factors, such as the underrepresentation of women in the younger group, might have induced some bias. Also, because of the limitation in the number of subjects, we are not able to address the clustering effect originating from sites in our multivariable analysis. There were some parameters that were studied in other DECAAF II substudies such as fibrosis regionality or natriuretic peptides that were found to determine recurrence of arrhythmia as well but were not included in this study.[26,27] Fibrosis regionality is an inherent part of total fibrosis burden and is therefore very dependent on the latter. Including regionality along with total fibrosis would cause bias in the analysis. Moreover, fibrosis regionality is a very novel tool that may not be applicable in many centers. For natriuretic peptides, the analysis was done on patients without heart failure, and so a small substudy of DECAAF II. Analyzing the effect of natriuretic peptides on recurrence in both patients with and without heart failure may not be ideal and potentially biased.

Conclusion

ML causal tree analysis can help identify patients that may benefit from a certain therapy using RCT databases. Using a hypothesis-free approach, causal tree learning allows us to identify variables with the highest effects on the outcome in an unbiased fashion. In this example, on the DECAAF II trial, patients with persistent AF younger than 58 years of age were shown to have a lower risk of recurrence of aTA with fibrosis-guided ablation than with conventional PVI alone. Further prospective evidence is still needed to confirm these findings. The use of ML, particularly causal ML on RCT datasets, should be encouraged to generate potential guideline-changing hypotheses.

Appendix

Supplementary data

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.hrthm.2025.01.008.

Funding Sources: None.

Disclosures: Dr Marrouche has received consulting fees from Biosense Webster, Boston Scientific, and AtriCure; he has

served as speaker for Abbott, Biosense Webster, AtriCure, and Sanofi and has received research support Abbott, Medtronic, Biosense Webster, Siemens, GE, Boston Scientific, Sanofi, and Samsung. Dr Marrouche has a family member as the CEO of Cardiac Designs. He is the founder of Marrek, being named in a patent issued for MRI fibrosis imaging and being a previous shareholder of Cardiac Designs. Dr Wazni has received personal fees (for consulting services) from Biosense Webster and from Boston Scientific during the conduct of the study. Dr Kholmovski has received personal fees and other (share ownership) from Marrek Inc during the conduct of the study and outside the submitted work; grants from Medtronic outside the submitted work; and a patent issued for US 9713436 licensed to Marrek Inc, US 10004425 licensed to Marrek Inc, and US 10726545. Dr Mansour has received personal fees (for consulting services) from Biosense Webster, Boston Scientific, and Medtronic and personal fees (holding equity) from EPD Solutions and New Pace Ltd outside the submitted work. Dr Marchlinski has received personal fees (scientific advisory board) from Abbott Medical, Biosense Webster, and Medtronic outside the submitted work. Dr Wilber has served on executive committees for clinical trials for Abbott and Boston Scientific; he has served as coprimary investigator of a clinical trial for Atricure; and has received grants from Abbott, Atricure, and Biosense; and personal fees from the American College of Cardiology Foundation (editor-in-chief, Journal of the American College of Cardiology editor's page), Biosense (consulting), and Medtronic (lectures for fellows) outside the submitted work. Dr Jais has received grants from Biosense Webster and Boston Scientific during the conduct of the study and grants from Acutus and Medtronic outside the submitted work. Dr Sanders has received advisory board and research grants to his institution from Medtronic, Abbott Medical, and Boston Scientific; has served on advisory boards for CathRx and Pacemate outside the submitted work; and has received support from a practitioner fellowship from the National Health and Medical Research Council of Australia. Dr Bax has received departmental unrestricted research grants from Abbott, Edwards Lifesciences, Medtronic, Biotronik, Boston Scientific, GE Healthcare, Novartis, and Bayer outside the submitted work. Dr Calkins has received personal fees from Biosense Webster, Abbott, and Boston Scientific outside the submitted work. The remaining authors have no conflicts of interest to disclose.

Data Availability: The data that support the findings of this study are available from the corresponding author, Dr Marrouche, upon reasonable request.

Address reprint requests and correspondence: Dr Nassir Marrouche, Tulane University School of Medicine, 1430 Tulane Ave, New Orleans, LA 70001. E-mail address: nmarrouche@tulane.edu

References

1. Li J, Ma S, Le T, Liu L, Liu J. Causal decision trees. IEEE Trans Knowl Data Eng 2016;29:257–271.
2. Wager S, Athey S. Estimation and inference of heterogeneous treatment effects using random forests. J Am Stat Assoc 2018;113:1228–1242.
3. Calkins H, Hindricks G, Cappato R, et al. 2017 HRS/EHRA/ECAS/APHRS/SOLAECE expert consensus statement on catheter and surgical ablation of atrial fibrillation. Heart Rhythm 2017;14:e275–e444.
4. Verma A, Jiang CY, Betts TR, et al. Approaches to catheter ablation for persistent atrial fibrillation. N Engl J Med 2015;372:1812–1822.
5. Marrouche NF, Wilber D, Hindricks G, et al. Association of atrial tissue fibrosis identified by delayed enhancement MRI and atrial fibrillation catheter ablation: the DECAAF study. JAMA 2014;311:498–506.
6. Mekhael M, Noujaim C, Hajjar AHE, et al. Persistent atrial fibrillation patients have higher atrial fibrosis and arrhythmia recurrence. J Am Coll Cardiol 2022;79(9_Suppl):104.
7. Marrouche NF, Greene T, Dean JM, et al. Efficacy of LGE-MRI-guided fibrosis ablation versus conventional catheter ablation of atrial fibrillation: the DECAAF II trial: study design. J Cardiovasc Electrophysiol 2021;32:916–924.
8. Marrouche NF, Wazni O, McGann C, et al. Effect of MRI-guided fibrosis ablation vs conventional catheter ablation on atrial arrhythmia recurrence in patients with persistent atrial fibrillation: the DECAAF II randomized clinical trial. JAMA 2022;327:2296–2305.
9. Huo Y, Gaspar T, Schönbauer R, et al. Low-voltage myocardium-guided ablation trial of persistent atrial fibrillation. N Engl J Med Evid 2022;1:EVIDoa2200141.
10. Athey S, Imbens G. Recursive partitioning for heterogeneous causal effects. Proc Nat Acad Sci 2016;113:7353–7360.
11. Rubin DB. Bayesian inference for causal effects: the role of randomization. Ann Stat 1978;6:34–58.
12. Taguchi A, Kato K, Furusawa A, et al. Heterogeneous treatment effect of dose-dense paclitaxel plus carboplatin therapy for advanced ovarian cancer. Int J Cancer 2024;155:1068–1077.
13. Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine: reporting of subgroup analyses in clinical trials. N Engl J Med 2007;357:2189–2194.
14. Basu S, Raghavan S, Wexler DJ, Berkowitz SA. Characteristics associated with decreased or increased mortality risk from glycemic therapy among patients with type 2 diabetes and high cardiovascular risk: machine learning analysis of the ACCORD trial. Diabetes Care 2018;41:604–612.
15. Watson JA, Holmes CC. Machine learning analysis plans for randomised controlled trials: detecting treatment effect heterogeneity with strict control of type I error. Trials 2020;21:1–10.
16. Petito LC, García-Albéniz X, Logan RW, et al. Estimates of overall survival in patients with cancer receiving different treatment regimens: emulating hypothetical target trials in the Surveillance, Epidemiology, and End Results (SEER)–Medicare Linked Database. JAMA Netw Open 2020;3:e200452–e.
17. Ngufor C, Yao X, Inselman JW, et al. Identifying treatment heterogeneity in atrial fibrillation using a novel causal machine learning method. Am Heart J 2023;260:124–140.
18. Iwasaki Y-k, Nishida K, Kato T, Nattel S. Atrial fibrillation pathophysiology. Circulation 2011;124:2264–2274.
19. Huo Y, Gaspar T, Pohl M, et al. Prevalence and predictors of low voltage zones in the left atrium in patients with atrial fibrillation. Europace 2017;20:956–962.
20. Goette A, Kalman JM, Aguinaga L, et al. EHRA/HRS/APHRS/SOLAECE expert consensus on atrial cardiomyopathies: definition, characterization, and clinical implication. Heart Rhythm 2017;14:e3–e40.
21. Goette A, Corradi D, Dobrev D, et al. Atrial cardiomyopathy revisited–evolution of a concept: a clinical consensus statement of the European Heart Rhythm Association (EHRA) of the ESC, the Heart Rhythm Society (HRS), the Asian Pacific Heart Rhythm Society (APHRS), and the Latin American Heart Rhythm Society (LAHRS). Europace 2024;26:euae204.
22. Metzner I, Wissner E, Tilz RR, et al. Ablation of atrial fibrillation in patients ≥75 years: long-term clinical outcome and safety. Europace 2016;18:543–549.
23. Bisbal F, Benito E, Teis A, et al. Magnetic resonance imaging-guided fibrosis ablation for the treatment of atrial fibrillation: the ALICIA trial. Circ Arrhythm Electrophysiol 2020;13:e008707.
24. Yang B, Jiang C, Lin Y, et al. STABLE-SR (Electrophysiological Substrate Ablation in the Left Atrium During Sinus Rhythm) for the treatment of nonparoxysmal atrial fibrillation. Circ Arrhythm Electrophysiol 2017;10:e005405.
25. Yang G, Zheng L, Jiang C, et al. Circumferential pulmonary vein isolation plus low-voltage area modification in persistent atrial fibrillation. JACC Clin Electrophysiol 2022;8:882–891.
26. Assaf A, Mekhael M, Noujaim C, et al. Effect of fibrosis regionality on atrial fibrillation recurrence: insights from DECAAF II. Europace 2023;25:euad199.
27. Younes H, Mekhael M, Feng H, et al. Baseline natriuretic peptides as a predictor of atrial fibrillation recurrence after radiofrequency-based pulmonary vein isolation in a non-heart failure population: a subanalysis from DECAAF II. Pacing Clin Electrophysiol 2023;46:848–854.