Pre-Procedure Application of Machine Learning and Mechanistic Simulations Predicts Likelihood of Paroxysmal Atrial Fibrillation Recurrence Following Pulmonary Vein Isolation

Running title: *Shade & Ali et al.; ML and Simulations predict AF recurrence after PVI*

Julie K. Shade, BS[1,2]*; Rheeda L. Ali, PhD[1]*; Dante Basile[1,2]; Dan Popescu, BS[1,3];

Tauseef Akhtar, MD[4]; Joseph E. Marine, MD[4]; David D. Spragg, MD[4]; Hugh Calkins, MD[4];

Natalia A. Trayanova, PhD[1,2,5]

[1]Alliance for Cardiovascular Diagnostic and Treatment Innovation, [2]Department of Biomedical Engineering, [3]Department of Applied Math and Statistics, Johns Hopkins University; [4]Division of Cardiology, Department of Medicine, [5]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD
*contributed equally

Correspondence:

Natalia Trayanova, PhD

Johns Hopkins University

Hackerman 216

3400 North Charles Street

Baltimore, MD 21218

Tel: +1 (410) 516-4375

Email: ntrayanova@jhu.edu

Journal Subject Terms: Atrial Fibrillation; Computational Biology; Magnetic Resonance Imaging (MRI); Fibrosis; Electrophysiology

Abstract:

Background - Pulmonary vein isolation (PVI) is an effective treatment strategy for patients with atrial fibrillation (AF), but many experience AF recurrence and require repeat ablation procedures. The goal of this study was to develop and evaluate a methodology which combines machine learning (ML) and personalized computational modeling to predict, prior to PVI, which patients are most likely to experience AF recurrence after PVI.

Methods - This single-center retrospective proof-of-concept study included 32 patients with documented paroxysmal AF who underwent PVI and had pre-procedural late gadolinium enhanced magnetic resonance imaging (LGE-MRI). For each patient, a personalized computational model of the left atrium simulated AF induction via rapid pacing. Features were derived from pre-PVI LGE-MRI images and from results of simulations (SimAF). The most predictive features were used as input to a quadratic discriminant analysis ML classifier, which was trained, optimized, and evaluated with 10-fold nested cross validation to predict the probability of AF recurrence post-PVI.

Results - In our cohort, the ML classifier predicted probability of AF recurrence with an average validation sensitivity and specificity of 82% and 89%, respectively, and a validation AUC of 0.82. Dissecting the relative contributions of SimAF and raw images to the predictive capability of the ML classifier, we found that when only features from SimAF were used to train the ML classifier, its performance remained similar (validation AUC=0.81). However, when only features extracted from raw images were used for training, the validation AUC significantly decreased (0.47).

Conclusions - ML and personalized computational modeling can be used together to accurately predict, using only pre-PVI LGE-MRI scans as input, whether a patient is likely to experience AF recurrence following PVI, even when the patient cohort is small.

Key words: magnetic resonance imaging; atrial fibrillation; pulmonary vein isolation; computer-based model; machine learning

Non-standard Abbreviations and Acronyms

PVI: pulmonary vein isolation

ML: machine learning

LGE-MRI: late gadolinium enhanced magnetic resonance imaging

SimAF: simulations of atrial fibrillation induction

AUC: area under the curve

LA: left atrium

RD: reentrant driver

MAT: macro-reentrant atrial tachycardia

MV: mitral valve

PV: pulmonary vein

QDA: quadratic discriminant analysis

LIPV: left inferior pulmonary vein

LSPV: left superior pulmonary vein

LAA: left atrial appendage

RIPV: right inferior pulmonary vein

n_{RD+MAT}: number of re-entrant drivers and macro-reentrant atrial tachycardias observed

n_{RD}: number of re-entrant drivers observed

n_{MAT}: number of macro-reentrant atrial tachycardias observed

P_{RD+MAT}: proportion of pacing locations from which AF (either re-entrant driver and macro-reentrant atrial tachycardia) was induced

P_{RD}: proportion of pacing locations from which re-entrant driver was induced

P_{MAT}: proportion of pacing locations from which macro-reentrant atrial tachycardia was induced

Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia,[1] with paroxysmal atrial fibrillation (PxAF) accounting for around 25% of AF cases.[2] Untreated AF leads to increased risk of stroke and heart failure.[2] For many PxAF patients, pulmonary vein isolation (PVI) ablation is a successful treatment strategy.[2] However, in a meta-analysis of PVI outcomes, only 78% of patients were free from AF at 12 months.[3] Patients who experience AF recurrence may require repeat PVI or additional substrate modification in the fibrotic left atrium (LA).[4,5] A methodology which identifies, prior to PVI, patients who are likely to experience post-PVI AF recurrence would allow development of targeted ablation strategies for these patients, reducing redo procedures and decreasing the risk of morbidity and mortality.[6]

Given the importance of early and effective intervention for atrial arrhythmias, many machine learning (ML)-based healthcare technologies have focused on AF detection and clinical outcome prediction.[7] There have been several attempts to use ML to predict AF recurrence after ablation, including a study which used deep learning-based LA shape analysis,[8] and another which used ML on imaging and clinical biomarkers to predict cryo-balloon PVI outcomes.[9] However, ample experimental and clinical evidence supports the primary role of fibrosis remodeling in the atria in the pathophysiology of AF,[10–13] which thus far has not been accounted for in ML approaches aimed at predicting AF recurrence after PVI.

Atrial fibrosis promotes the initiation and perpetuation of re-entrant activity underlying AF by disrupting conduction and establishing regions of pro-fibrillatory substrate.[11,13] However, it remains unknown to what degree the patient-specific fibrosis distribution prior to ablation is a contributing factor to AF recurrence after PVI. Personalized biophysically-detailed computational models of the atria based on the patient's late gadolinium enhanced magnetic

resonance imaging (LGE-MRI), which visualizes the personalized fibrosis distribution, allow for clinically-validated,[14] non-invasive investigation of the susceptibility of a patient's fibrotic substrate to sustaining reentrant activity.[15] Such atrial models of arrhythmogenic propensity may have potential to predict, pre-procedure, the probability of AF recurrence after PVI, but reducing hundreds of thousands of transmembrane voltage measurements over thousands of milliseconds to meaningful predictive features is a difficult task. In addition, mechanistic modeling does not explicitly consider clinical biomarkers or quantitative measures of the structural remodeling derived from raw imaging data. ML classifiers are ideal for identifying predictive patterns in high-dimensional data and combining predictive features derived from multiple sources, thus we hypothesize that a combination of ML and mechanistic modeling may provide accurate pre-procedure prediction of AF recurrence after PVI.

In this proof-of-concept study, we develop a novel approach that combines mechanistic computational modeling and ML to predict, before the ablation procedure, the individual patient's probability of AF recurrence post-PVI. We show that this approach results in an ML classifier that achieves high validation sensitivity and specificity even when even when the patient cohort available for training is small.

Methods

Data Materials, and Code Disclosure

The authors declare that all data supporting the findings of the study are available within the paper and its supplementary information. Source data for activation maps are available from the corresponding author upon request. The LGE-MRI images used to construct the personalized LA

computational models are available on request and on approval of the Johns Hopkins Institutional Review Board.

Overview of Methodology

In this institutional review board-approved retrospective study, we present a novel predictor of AF recurrence post-PVI using ML and personalized LA computational models in patients with PxAF and discernable fibrosis on LGE-MRI. The predictor is designed to be applied pre-procedure; an overview of it is presented in Figure 1. In a cohort of 32 patients, we evaluated the predictive capability of this classifier. In doing so, for each patient, a personalized computational model of the LA was constructed from the pre-procedure LGE-MRI images and used to simulate AF induction via rapid pacing. Features were derived from the results of these simulations (SimAF) and from raw LGE-MRI images, and served as input into a quadratic discriminant analysis (QDA) ML classifier. Features from SimAF were chosen in two ways: i) they were based on general knowledge of AF dynamics, and ii) they were left to be chosen by the ML training algorithm, unsupervised. The classifier was trained, optimized, and evaluated with 10-fold nested cross validation. Finally, we assessed the capabilities of ML risk predictors that used features from SimAF only and from LGE-MRI imaging only, thus distinguishing the relative contributions to the overall predictive capability.

Patient Cohort and Cardiac MRI Image Acquisition

This study included adult patients from a single center with documented PxAF who received pre-procedural LGE-MRI scans and underwent PVI ablation between December 2011 and December 2015. PxAF was defined as an episode of AF that terminated spontaneously or with intervention within 7 days.[2] Patients were excluded from the study if their LGE-MRI had motion or breathing artifact or if the myocardium was not correctly nulled, resulting in insufficient visualization of

the LA geometry and fibrosis distribution for model reconstruction and ML. 32 patients were included in the study.

All patients were observed overnight in the hospital for hemodynamic monitoring and resumption of anticoagulation. Routine follow-up with electrocardiograms (ECGs) and clinical assessment was performed at 3, 6, and 12 months. Additional follow-up for symptomatic patients was performed if necessary, including Holter monitoring. Any recurrence of AF/atrial tachycardia (AT) documented by ECG or a device-recording system lasting ≥30 seconds, outside of a 3-month post-procedure blanking period, was considered recurrence.

Twelve (38%) patients experienced AF recurrence in the follow-up period. All patients had PVI with either radiofrequency (RF) or cryo-balloon ablation. 28 (88%) patients underwent circumferential linear RF ablation around the left and right pulmonary veins. The other 4 (12%) patients had cryo-balloon ablation performed with a 23 or 28mm cryo-balloon. The median time between the PVI procedure and last date of follow-up was 366 days (IQR: 365-467 days). The median time to reported AF recurrence was 310 days (IQR: 204-381 days).

Pre-procedure LGE-MRI scans were acquired using a 1.5 T Avanto MR system for the purpose of visualizing and reconstructing the atrial geometry and fibrosis distribution. Scans were performed in the axial orientation 10–27min following 0.2mmol/kg of gadobenate dimeglumine contrast agent using a fat-saturated 3-dimensional (3D) IR-prepared fast spoiled gradient recalled echo sequence, with electrocardiogram-triggered and respiratory navigator gating. Image resolution was 1.25x1.25x2.5mm.

Methodology for Personalized Atrial Computational Modeling

A full description of the personalized atrial geometric model reconstruction workflow can be found in our previous publications.[5,16,17] Briefly, the LA epicardial and endocardial walls were

manually delineated on the pre-procedure LGE-MRI using ITK-snap.[18] Fibrotic voxels were classified with image intensity ratio greater than 1.22,[19] this threshold has been validated clinically by using thresholded fibrosis in atrial models to predict AF ablation targets.[16] High-resolution tetrahedral finite-element meshes were generated from the up-sampled segmented images.[20] Realistic myocardial fiber orientations were incorporated using a diffeomorphic mapping technique from an atlas geometry.[17,21,22]

Electrophysiological properties were assigned to non-fibrotic and fibrotic tissue in the geometric models as described previously.[14,23,24] Specifically, a human chronic AF action potential model[25] with modifications to fit clinical monophasic action potential recordings from patients with AF[26] was used to represent membrane kinetics in non-fibrotic myocardium. In fibrotic regions, further ionic modifications were implemented,[27–29]. At the tissue scale, fibrotic regions had reduced conductivities to represent impaired cell-to-cell coupling, as we have described previously.[29] The rapid pacing atrial arrhythmia induction protocol is described on our previous publications, and involved rapid-pacing from 30 uniformly distributed locations on the LA.[17,24] Simulations were performed in the CARP software package (https://carp.medunigraz.at/).[30,31] Persistent reentrant drivers (RDs) were identified using the wavefront tip analysis method.[32,33] We also identified macro-reentrant atrial tachycardias (MAT), wavefront propagation around a non-conductive obstacle such as the mitral valve (MV) or pulmonary vein (PV).[17]

Extraction of Features from SimAF

Once the personalized simulations of AF induction in all 32 models were completed, features from SimAF were selected for input into the ML classifier. These were chosen in two ways: i) by the authors based on general knowledge of AF dynamics (deductive features), and ii) by the ML

training algorithm, unsupervised (inductive features). Deductive features included presence of RDs, theorized to correlate with or predict likelihood of AF recurrence,[5] as well as other features we thought might be meaningful, such as the number of RDs inside the regions isolated by PVI. The latter could be predictive of AF recurrence in the case of PV reconnection, which has been shown to occur in as many as 85% of patients experiencing AF recurrence.[34] In contrast, inductive features of SimAF were "learned" in an unsupervised manner during classifier training by analyzing SimAF in models of patients who experienced AF recurrence. This category of features was included to allow learning of predictive features of SimAF not previously described in the literature and to reduce bias that may be introduced by hand-crafting the choice of features. Detail regarding the extraction of inductive features from SimAF is presented in Supplementary Materials. A complete list of features extracted from SimAF is presented in Table 1.

Extraction of Features from LGE-MRI Images

While LA models were based on pre-procedure LGE-MRI images, model reconstruction involved binarizing the fibrosis distribution via thresholding, as well as its interpolation and mapping to the 3D mesh. As unprocessed (raw) images of fibrosis distribution might contain additional prognostic information pertinent to AF recurrence, features from the pre-procedure LGE-MRI atrial scans were also made available to the classifier. A complete list of features extracted from LGE-MRI scans is presented in Table 1. They included, among others, the heterogeneity and quantity of the fibrosis distribution, both suggested to correlate with AF propensity,[5] and a fractal dimension-based feature[35] which quantifies how quickly the complexity of the 3D surface of the fibrosis volume decreases as resolution decreases; the latter

was calculated by analyzing the differences in the number of cubes of various sizes required to cover the entire surface of the fibrotic region.

Training, Optimization, and Evaluation of QDA Classifier

Ten-fold nested cross-validation was used to train, validate, and test the classifier.[36] Random forests were used for unbiased feature selection,[37] then a QDA classifier was trained[38,39] using the selected features to predict the probability of AF recurrence after PVI. The selected features are listed in Table 1; all of these were extracted, as described, either directly from raw pre-PVI LGE-MRI or from SimAF in computational LA models reconstructed from pre-PVI LGE-MRI. Features extracted from raw images and those derived from SimAF were not treated differently in any way in the fully automatic feature section process, thus preventing bias towards either subset of features. Optimized hyperparameters included the number of features selected for the classifier and various parameters used to calculate the inductive SimAF features. The inductive SimAF features learned during training were recorded for further analysis. Further details regarding ML are provided in the online Supplementary Methods.

Results

Clinical Characteristics of Study Cohort

There were no significant differences in clinical characteristics, including several known AF risk factors, between patients who did and did not experience AF recurrence (Table 2). Accordingly, we did not attempt to train classifiers with any of these clinical biomarkers as none of them were associated with AF recurrence in this retrospective cohort.

AF Inducibility in Patient-Specific LA Models

Figure 2 presents several examples of reconstructed LA models showing the pre-procedure patient-specific atrial geometries and fibrosis distributions, as well as examples of induced reentrant activity following the rapid pacing protocol in the models. A MAT was observed around the left inferior pulmonary vein (LIPV) in the LA model of Patient 1 (Fig. 2A). An RD was found on the posterior left atrium adjacent to the MV in the LA model of patient 2 (Fig. 2B). In the LA model of patient 3, RDs were found on the inferior posterior wall and LIPV, and a MAT perpetuated around the RIPV (Fig. 2C).

We first tested whether the results of SimAF themselves could be used, pre-procedure, to predict AF recurrence post-PVI. We found that reentry was induced from a larger number of pacing sites in the LA models of patients who experienced AF recurrence post-PVI (9.2±1.8) compared to those of patients who did not (5.7±1.7), but this did not reach statistical significance (p=0.19, Fig. 2D). We also examined whether the number of pacing sites from which reentry was induced in the models was associated with AF recurrence after PVI-- the resulting AUC was 0.72. More RDs and MATs were observed in the models of patients who experienced AF recurrence (2.6±0.4) compared to those of patients who did not (1.7±0.4), but this also did not reach statistical significance (p=0.18, Fig. 2E). Using the number of pre-ablation simulated RDs and MATs to predict have AF recurrence after PVI, the AUC was 0.69 and the sensitivity and specificity were 75% and 60%. The lack of statistical significance in the hand-picked features and relatively low training AUCs indicates that this approach may not perform well when applied to previously unseen patients.

Prediction of Post-PVI AF Recurrence Using ML on Features from Pre-Procedure LGE-MRI Imaging and SimAF

Our risk predictor, which used ML on features extracted from both pre-procedure LGE-MRI imaging and SimAF, predicted post-PVI AF recurrence with an average validation sensitivity and specificity of 82% and 89%, respectively, and a validation AUC of 0.82 (Fig. 3). The training AUC was similar: 0.90. This indicates that the classifier is generalizable, or likely to correctly predict whether a previously unseen patient will experience AF recurrence after PVI, despite the small data set available for training.

When only features derived from SimAF we used in an ML classifier, the predictive capability was similar: an average validation sensitivity and specificity of 79% and 89%, respectively, and a validation AUC of 0.81. This indicates that, in this small cohort, features extracted directly from raw images did not bestow additional predictive capabilities to the ML classifier, over those based solely on SimAF features, as pre-procedure LGE-MRIs are already included in the personalized models.

In comparison, an ML classifier using as input only features extracted from raw LGE-MRI (without imaging-based simulations) had validation sensitivity and specificity of 57% and 61% respectively. This classifier achieved a training AUC similar to that of the classifier using combined inputs (0.85 vs. 0.90, respectively), but a much lower validation AUC (0.47 vs. 0.82, respectively), the latter indicating that it was not generalizable. This suggests that data from a much larger patient cohort would be required to train an ML classifier to correctly predict AF recurrence risk of a previously unseen patient using only features extracted from raw LGE-MRI images, further supporting the inclusion of features extracted from imaging-based simulation results regarding AF propensity in ML-based AF recurrence risk prediction.

Analysis of Inductive Features Learned by the ML Algorithm from SimAF

To gain insight into how the ML classifier predicted AF recurrence after PVI, we analyzed the inductive SimAF features learned from the training data in each "outer loop" of cross-validation. Figure 4 presents the learned features of SimAF (reentry locations and pacing locations) that were most predictive of AF recurrence and were thus selected by the ML classifier. All of the most predictive reentry-inducing pacing locations that contributed to the selected inductive SimAF features were outside the PVs. Additionally, the most predictive inductive features frequently involved the numbers of RDs and MATs on the LIPV and mid anterior wall, as well as the MV for MAT.

Discussion

The goal of this study is to develop and evaluate a novel methodology for prediction of AF recurrence post-PVI using ML and personalized mechanistic modeling of AF induction in the LA of patients with PxAF and fibrotic remodeling on LGE-MRI. The ML classifier is designed to be applied before the ablation procedure, using only data available up to that time point. It uses as inputs features extracted from LGE-MRI-based simulation results for AF propensity in the fibrotic substrate, as well as those extracted from raw pre-procedure LGE-MRI images. We demonstrate that this approach results in a highly predictive and generalizable classifier, even when the patient cohort used for training is small. Our approach achieved an average validation sensitivity and specificity of 82% and 89%, respectively, and a validation AUC of 0.82, indicating that the classifier is generalizable and likely to accurately predict the AF risk of a previously unseen patient. To our knowledge, this is the first study to demonstrate that ML and

mechanistic cardiac modeling can be used together to develop an accurate and generalizable classifier that predicts the risk of adverse clinical events.

Our ML-based AF recurrence risk prediction methodology incorporated both inductive SimAF features learned during training of the ML classifier and deductive SimAF features chosen prior to training. The inclusion of inductive features is analogous to deep learning, a popular form of ML in which feature extraction is performed in the process of training a classifier, rather than prior to training.[40] Analysis of the inductive features extracted during training confirms that the classifier can learn patterns of SimAF which have not necessarily been previously evaluated for correlation with AF recurrence, but are clinically explainable in the context of published studies describing mechanisms of AF recurrence. For example, all pacing locations selected by the unsupervised algorithm for inductive SimAF feature extraction were outside the PVs, which is supported by research suggesting that triggers outside the PVs contribute to AF recurrence.[4]

Additionally, among the most predictive inductive SimAF features were the numbers of RDs and MATs on the LIPV and mid anterior wall, and the number of MATs around the MV. Since the mid anterior wall and the MV are outside the PV region and would not be electrically isolated by PVI, it is understandable that fibrosis distribution that can result in RDs and MATs forming there would be highly predictive of post-PVI AF recurrence. Furthermore, while it may seem counter-intuitive that reentry around the LIPV would predict AF recurrence, since PVI should electrically isolate this region, this feature might predict in which patients AF would recur should PVs reconnect after the ablation procedure. Jiang et al. found that re-connection of 1 or more PVs occurs in 85.5% of patients with AF recurrence, and 58.6% in patients without AF recurrence,[34] which supports our findings.

In designing this study, we hypothesized that an ML classifier which included features extracted from both image-based SimAF and the raw LGE-MRI images would provide the best prediction of post-PVI AF recurrence. However, we found that the features extracted from the raw LGE-MRI images did not contribute significantly to the predictive capability of the ML classifier. This indicates that while model generation from LGE-MRI images effectively reduced the dimensionality of the images, it nonetheless retained enough image information about the fibrotic substrate to be able to accurately predict whether a given patient will experience post-PVI AF recurrence.

When we constructed an ML classifier which only included features extracted from raw LGE-MRI images, in our small training cohort it achieved a much lower validation AUC than training AUC. This result indicates that an ML classifier based on raw images (that does not incorporate personalized simulations) would need to be trained on a much larger imaging data set to be able to correctly predict AF recurrence post-PVI of a previously unseen patient.

The validation AUC of the ML classifier which included only features derived from SimAF to predict AF recurrence (Fig. 3B) was higher than the AUCs achieved using measures of AF inducibility in the LA models without ML (Fig. 2F). While both risk prediction methods relied on the same underlying SimAF in the same LA models, the ML classifier training algorithm allowed selection of multiple features derived from SimAF and weighting of these features, so the resulting classifier was finely tuned. In contrast, the non-ML method predicted the risk of AF recurrence via thresholding of a single measure of AF propensity (number of reentry locations or number of reentry-inducing pacing sites); it did not consider multiple features, their relative importance, or inductive features learned from the training data.

In the presented approaches to predict post-PVI AF recurrence here, LGE-MRI is a central data input, as it has the ability to accurately visualize LA geometry and spatial distribution of fibrosis. This ability is paramount to our study and to many others that have utilized atrial LGE-MRI in clinical studies.[5,16] Fibrosis on LGE-MRI has been shown to correlate with that in histological studies of biopsied patients.[10] Furthermore, the measure used to define fibrosis in our study (image intensity ratio) has been validated by correlation with local intracardial LA bipolar voltage measurements.[19] Finally, Chubb et al. [41] found excellent reproducibly of LGE-MRI, demonstrating that LGE scar assessment is very specific, albeit not particularly sensitive. The latter is an advantage in our study as it prevents us from predicting false AF inducibility in a "false-positive" fibrotic substrate.

A limitation of this study is the small data set, often a serious concern in ML studies. However, by using 10-fold nested cross-validation and aggregating the results, we demonstrated nonetheless excellent predictive capability. Our combined approach achieved similarly high validation and training AUCs, indicating that the resulting classifier was generalizable despite the small data set. Prospective validation in a larger cohort would confirm the predictive capability of the proposed post-PVI AF risk prediction methodology. Further, although this study only included PxAF patients with adequate clinical follow-up at 3, 6, and 12 months, the methods for defining recurrence of AF are a limitation of any clinical study that uses such follow-up, especially the difficulty in identifying asymptomatic recurrent AF.

Another limitation is the lack of published detailed methods and quantitative results for previously proposed imaging-based ML AF recurrence prediction methodologies, which limited our ability to compare these with our methodology. A recent approach which used only imaging features[8] achieved a F1-score (harmonic mean of precision and recall) of 0.33, while our

approach achieved a testing F1-score of 0.70. However, the study did not state what part of the data (training, testing, or validation) this F1-score applied to, so this may be a poor comparison. Further, two clinical risk scores have been proposed to predict the risk of AF recurrence after PVI: the ATLAS score[42] and the CAAP-AF score[43], but our retrospective registry did not contain all the necessary variables to calculate either of these risk scores. However, we note that our ML classifier achieved a validation AUC of 0.82, which is greater than the censored C-statistic of 0.75 achieved by the ATLAS score and the development AUC of 0.69 achieved by the CAAP-AF score.

In this study, we developed an ML classifier able to accurately predict, pre-procedure, AF recurrence post-PVI. The classifier uses as inputs features extracted from LGE-MRI-based simulation results for AF propensity in the fibrotic substrate and those extracted from raw pre-procedure LGE-MRI images. The ML classifier is designed to be applied before the ablation procedure. Our vision for the use of this classifier is that should it predicts that AF will recur post-PVI, then the patient's LA model would be used to also predict the treatment strategy, i.e. the personalized ablation targets outside the wide-area PVI, as we have done prospectively in a recent study.[16] An important characteristic of our resulting classifier is that it considers the potential patient-specific mechanisms of arrhythmogenesis resulting from the fibrotic substrate in the LA, making it clinically explainable. To our knowledge, this is the first study to demonstrate the potential of combining computational cardiac modeling and ML to make clinical predictions.

Sources of Funding: This work was supported by National Institutes of Health grant (U01-HL141074) to N.T., a grant from the Lowenstein Foundation to N.T., a National Science Foundation Graduate Research Fellowship (DGE-1746891) to J.S, and a research fellowship from Johns Hopkins University to R.A.

Disclosures: None

References:

1. Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby J V., Singer DE. Prevalence of Diagnosed Atrial Fibrillation in Adults. *JAMA*. 2001;285:2370.

2. Calkins H, Hindricks G, Cappato R, Kim Y, Saad EB, Aguinaga L, Akar JG, Badhwar V, Brugada J, Camm J, et al. 2017 HRS / EHRA / ECAS / APHRS / SOLAECE expert consensus statement on catheter and surgical ablation of atrial fi brillation. *Heart Rhythm*. 2017;14:e275–e444.

3. Kis Z, Muka T, Franco OH, Bramer WM, De Vries LJ, Kardos A, Szili-Torok T. The Short and Long-Term Efficacy of Pulmonary Vein Isolation as a Sole Treatment Strategy for Paroxysmal Atrial Fibrillation: A Systematic Review and Meta-Analysis. *Curr Cardiol Rev*. 2017;13:199–208.

4. Lemola K, Hall B, Cheung P, Good E, Han J, Tamirisa K, Chugh A, Bogun F, Pelosi F, Morady F, et al. Mechanisms of recurrent atrial fibrillation after pulmonary vein isolation by segmental ostial ablation. *Heart Rhythm*. 2004;1:197–202.

5. Ali RL, Hakim JB, Boyle PM, Zahid S, Sivasambu B, Marine JE, Calkins H, Trayanova NA, Spragg DD. Arrhythmogenic propensity of the fibrotic substrate after atrial fibrillation ablation: a longitudinal study using magnetic resonance imaging-based atrial models. *Cardiovasc Res*. 2019;115:1757–1765.

6. Cheng EP, Liu CF, Yeo I, Markowitz SM, Thomas G, Ip JE, Kim LK, Lerman BB, Cheung JW. Risk of Mortality Following Catheter Ablation of Atrial Fibrillation. *J Am Coll Cardiol*. 2019;74:2254–2264.

7. Cantwell CD, Mohamied Y, Tzortzis KN, Garasto S, Houston C, Chowdhury RA, Ng FS, Bharath AA, Peters NS. Rethinking multiscale cardiac electrophysiology with machine learning and predictive modelling. *Comput Biol Med*. 2019;104:339–351.

8. Bhalodia R, Goparaju A, Sodergren T, Whitaker R, Morris A, Kholmovski E, Marrouche N, Cates J, Elhabian S. Deep Learning for End-to-End Atrial Fibrillation Recurrence Estimation. In: 2018 Computing in Cardiology Conference (CinC). Computing in Cardiology. 2018.

9. Budzianowski J, Hiczkiewicz J, Burchardt P, Pieszko K, Rzeźniczak J. Predictors of atrial fibrillation early recurrence following cryoballoon ablation of pulmonary veins using statistical assessment and machine learning algorithms. *Heart Vessels*. 2019;34:352–359.

10. McGann C, Akoum N, Patel A, Kholmovski E, Revelo P, Damal K, Wilson B, Cates J, Harrison A, Ranjan R, et al. Atrial fibrillation ablation outcome is predicted by left atrial remodeling on MRI. *Circ Arrhythmia Electrophysiol*. 2014;7:23–30.

11. Lau DH, Linz D, Schotten U, Mahajan R, Sanders P, Kalman JM. Pathophysiology of Paroxysmal and Persistent Atrial Fibrillation: Rotors, Foci and Fibrosis. *Hear Lung Circ*. 2017;26:887–893.

12. Saha M, Roney CH, Bayer JD, Meo M, Cochet H, Dubois R, Vigmond EJ. Wavelength and fibrosis affect phase singularity locations during atrial fibrillation. *Front Physiol*. 2018;9.

13. Platonov PG. Atrial fibrosis: An obligatory component of arrhythmia mechanisms in atrial fibrillation? *J Geriatr Cardiol*. 2017;14:233–237.

14. Boyle PM, Hakim JB, Zahid S, Franceschi WH, Murphy MJ, Vigmond EJ, Dubois R, Haïssaguerre M, Hocini M, Jaïs P, et al. Comparing reentrant drivers predicted by image-based computational modeling and mapped by electrocardiographic imaging in persistent atrial fibrillation. *Front Physiol*. 2018;9:1–12.

15. Aronis KN, Ali R, Trayanova NA. The role of personalized atrial modeling in understanding atrial fibrillation mechanisms and improving treatment. *Int J Cardiol*. 2019;287:139–147.

16. Boyle PM, Zghaib T, Zahid S, Ali RL, Deng D, Franceschi WH, Hakim JB, Murphy MJ, Prakosa A, Zimmerman SL, et al. Computationally guided personalized targeted ablation of persistent atrial fibrillation. *Nat Biomed Eng*. 2019;3:870-879.

17. Zahid S, Cochet H, Boyle PM, Schwarz EL, Whyte KN, Vigmond EJ, Dubois R, Hocini M, Haïssaguerre M, Jaïs P, et al. Patient-derived models link re-entrant driver localization in atrial fibrillation to fibrosis spatial pattern. *Cardiovasc Res*. 2016;110:443–454.

18. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, Gerig G. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*. 2006;31:1116–1128.

19. Khurram IM, Beinart R, Zipunnikov V, Dewire J, Yarmohammadi H, Sasaki T, Spragg DD, Marine JE, Berger RD, Halperin HR, et al. Magnetic resonance image intensity ratio, a normalized measure to enable interpatient comparability of left atrial fibrosis. *Heart Rhythm*. 2014;11:85–92.

20. Vadakkumpadan F, Rantner LJ, Tice B, Boyle P, Prassl AJ, Vigmond E, Plank G, Trayanova N. Image-based models of cardiac structure with applications in arrhythmia and defibrillation studies. *J Electrocardiol*. 2009;42:157.e1-157.e10.

21. Krueger MW, Schmidt V, Tobón C, Weber FM, Lorenz C, Keller DUJ, Barschdorf H, Burdumy M, Neher P, Plank G, et al. Modeling atrial fiber orientation in patient-specific geometries: A semi-automatic rule-based approach. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2011;6666 LNCS:223–232.

22. McDowell KS, Vadakkumpadan F, Blake R, Blauer J, Plank G, MacLeod RS, Trayanova

NA. Methodology for patient-specific modeling of atrial fibrosis as a substrate for atrial fibrillation. *J Electrocardiol*. 2012;45:640–645.

23. Boyle PM, Hakim JB, Zahid S, Franceschi WH, Murphy MJ, Prakosa A, Aronis KN, Zghaib T, Balouch M, Ipek EG, et al. The fibrotic substrate in persistent atrial fibrillation patients: Comparison between predictions from computational modeling and measurements from focal impulse and rotor mapping. *Front Physiol*. 2018;9:1–12.

24. Zahid S, Whyte KN, Schwarz EL, Blake RC, Boyle PM, Chrispin J, Prakosa A, Ipek EG, Pashakhanloo F, Halperin HR, et al. Feasibility of using patient-specific models and the "minimum cut" algorithm to predict optimal ablation targets for left atrial flutter. *Heart Rhythm*. 2016;13:1687–1698.

25. Courtemanche M, Ramirez RJ, Nattel S. Ionic mechanisms underlying human atrial action potential properties: Insights from a mathematical model. *Am J Physiol.* 1998;275.

26. Krummen DE, Bayer JD, Ho J, Ho G, Smetak MR, Clopton P, Trayanova NA, Narayan SM. Mechanisms of human atrial fibrillation initiation clinical and computational studies of repolarization restitution and activation latency. *Circ Arrhythmia Electrophysiol*. 2012;5:1149–1159.

27. Corradi D, Callegari S, Maestri R, Benussi S, Alfieri O. Structural remodeling in atrial fibrillation. *Nat Clin Pract Cardiovasc Med*. 2008;5:782–796.

28. Nattel S, Burstein B, Dobrev D. Atrial remodeling and atrial fibrillation: mechanisms and implications. *Circ Arrhythm Electrophysiol*. 2008;1:62–73.

29. Avila G, Medina IM, Jiménez E, Elizondo G, Aguilar CI. Transforming growth factor-β1 decreases cardiac muscle L-type Ca 2+ current and charge movement by acting on the Cav1.2 mRNA. *Am J Physiol.* 2007;292:622–631.

30. Plank G, Zhou L, Greenstein JL, Cortassa S, Winslow RL, O'Rourke B, Trayanova NA. From mitochondrial ion channels to arrhythmias in the heart: Computational techniques to bridge the spatio-temporal scales. *Philos Trans R Soc A Math Phys Eng Sci*. 2008;366:3381–3409.

31. Vigmond EJ, Weber dos Santos R, Prassl AJ, Deo M, Plank G. Solvers for the cardiac bidomain equations. *Prog Biophys Mol Biol*. 2008;96:3–18.

32. Deng D, Murphy MJ, Hakim JB, Franceschi WH, Zahid S, Pashakhanloo F, Trayanova NA, Boyle PM. Sensitivity of reentrant driver localization to electrophysiological parameter variability in image-based computational models of persistent atrial fibrillation sustained by a fibrotic substrate. *Chaos*. 2017;27:1–10.

33. Eason J, Trayanova N. Phase singularities and termination of spiral wave reentry. *J Cardiovasc Electrophysiol*. 2002;13:672–679.

34. Jiang R hong, Jiang C yang. Pulmonary Vein Reconnection in Patients With and Without Atrial Fibrillation Recurrence After Ablation. *JACC Clin Electrophysiol*. 2016;2:484–486.

35. Beheshti SMA, AhmadiNoubari H, Fatemizadeh E, Khalili M. An Efficient Fractal Method for Detection and Diagnosis of Breast Masses in Mammograms. *J Digit Imaging*. 2014;27:661–669.

36. Cawley GC, Talbot NLC. On over-fitting in model selection and subsequent selection bias in performance evaluation. *J Mach Learn Res*. 2010;11:2079–2107.

37. Yoh WY. Regression Trees with Unbiased Variable Selection. *Korean J Appl Stat*. 2004;17:459–473.

38. Srivastava S, Gupta MR, Frigyik BA. Bayesian quadratic discriminant analysis. *J Mach Learn Res*. 2007;8:1277–1305.

39. Guo Y, Hastie T, Tibshirani R. Regularized linear discriminant analysis and its application in microarrays. *Biostatistics*. 2007;8:86–100.

40. Dara S, Tumma P. Feature Extraction by Using Deep Learning: A Survey. *Proc 2nd Int Conf Electron Commun Aerosp Technol ICECA 2018*. 2018;1795–1801.

41. Chubb H, Karim R, Whitaker J, Harrison J, Wright M, O'Neill M, Razavi R. Reproducibility of late gadolidium enhancement of atrial ablation scar. *J Cardiovasc Magn Reson*. 2016;18:1–2.

42. Mesquita J, Ferreira AM, Cavaco D, Moscoso Costa F, Carmo P, Marques H, Morgado F, Mendes M, Adragão P. Development and validation of a risk score for predicting atrial fibrillation recurrence after a first catheter ablation procedure - ATLAS score. *Europace*. 2018;20:f428–f435.

43. Winkle RA, Jarman JWE, Mead RH, Engel G, Kong MH, Fleming W, Patrawala RA. Predicting atrial fibrillation ablation outcome: The CAAP-AF score. *Heart Rhythm*. 2016;13:2119–2125.

Table 1. List of all features extracted from raw LGE-MRI images and imaging-based simulation results.

Feature Category	Feature Description	Included in classifier (%)
Inductive SimAF	n_{RD+MAT} within N most predictive anatomical regions	100%
	n_{RD} within N most predictive anatomical regions	20%
	n_{MAT} within N most predictive anatomical regions	
	Mean p_{RD+MAT} per anatomical region which induced RD or MAT localized to single most predictive anatomical region	
	Mean p_{RD} per anatomical region which induced RD localized to single most predictive anatomical region	
	Mean p_{MAT} per anatomical region which induced which induce MAT localized to single most predictive anatomical region	
	Mean p_{RD+MAT} per anatomical region which induced RD or MAT localized to any of N most predictive anatomical regions	10%
	Mean p_{RD} per anatomical region which induced RD localized to any of N most predictive anatomical regions	10%
	Mean p_{MAT} per anatomical region which induced MAT localized to any of N most predictive anatomical regions	10%
	p_{RD+MAT} within P most predictive anatomical regions	20%
	p_{RD} within P most predictive anatomical regions	
	p_{MAT} within P most predictive anatomical regions	
	p_{RD+MAT} within P most predictive anatomical regions that induce RD or MAT localized to any of N most predictive reentry locations	40%
	p_{RD} within P most predictive anatomical regions that induce RD localized to any of N most predictive RD locations	
	p_{MAT} within P most predictive anatomical regions that induce MAT localized to any of N most predictive MAT locations	10%
Deductive SimAF	p_{RD+MAT} (all pacing locations)	
	p_{RD} (all pacing locations)	
	n_{RD} (all locations)	

	n_{MAT} (all locations)	
	n_{RD} localized to PV region	
	n_{RD} localized outside PV region	
	n_{MAT} localized outside PV region	
	p_{RD} that lead to RDs outside PV region (all pacing locations)	
	n_{RD+MAT} (all locations)	
	n_{RD} / p_{RD} (number of RDs per inducible pacing site)	
Imaging	Ratio of fibrotic tissue to entire atrial myocardium	
	Ratio of fibrotic tissue to non-fibrotic tissue	
	Fibrosis entropy	
	Difference in ratio of FD of fibrosis between 1^3 to 2^3 voxels and 2^3 to 4^3 voxels	10%
	Difference in ratio of FD of fibrosis between 2^3 to 4^3 voxels and 4^3 to 8^3 voxels	
	Difference in ratio of FD of fibrosis between 4^3 to 8^3 voxels and 8^3 to 16^3 voxels	
	Difference in ratio of FD of fibrosis between 8^3 to 16^3 voxels and 16^3 to 32^3 voxels	10%
	Difference in ratio of FD of fibrosis between 16^3 to 32^3 voxels and 32^3 to 64^3 voxels	
	Difference in ratio of FD of fibrosis between 32^3 to 64^3 voxels and 64^3 to 128^3 voxels	
	Difference in ratio of FD of fibrosis between 64^3 to 128^3 voxels and 128^3 to 256^3 voxels	

Features extracted for each patient from raw LGE-MRI images and imaging-based simulation results and the frequency with which each feature was selected for inclusion in the ML classifier. 10 outer loops of cross validation were performed; for example, 100% indicates that the feature was among the most predictive in all 10 outer loops. machine learning (ML), simulations of atrial fibrillation induction (SimAF), reentrant driver (RD), macro-reentrant tachycardia (MAT), number of RD and MAT observed (n_{RD+MAT}), number of RD observed (n_{RD}), number of MAT observed (n_{MAT}), proportion of pacing sites that led to RD or MAT (p_{RD+MAT}), proportion of pacing sites that led to RD (p_{RD}), proportion of pacing sites that led to MAT (p_{MAT}), pulmonary vein (PV), fractal dimension (FD)

Table 2. Baseline characteristics of the paroxysmal AF patient cohort.

Clinical Characteristics	Freedom from AF (n=20)	AF Recurrence (n=12)	p
Age (years)	63 ± 8	63 ± 9	0.87
Male sex	9	4	0.71
Body mass index (kg/m^3)	27 ± 6	27 ± 5	0.99
Hypertension	10 (50%)	7 (58%)	0.73
Impaired Glucose or Diabetes	5 (25%)	1 (8%)	0.63
Congestive heart failure	4 (20%)	2 (17%)	1
Ablation Procedure Details			
Cryo Ablation	3 (15%)	1 (8%)	1
Flutter Line	5 (20%)	3 (25%)	1

P-values were calculated with Student's t-test or Fisher's exact test as appropriate. atrial fibrillation (AF)

Figure Legends:

Figure 1. Overview of AF recurrence risk prediction methodology. The flowchart shows our novel machine learning (ML) methodology to predict, before the ablation procedure, atrial fibrillation (AF) recurrence following pulmonary vein isolation for patients with paroxysmal AF. For each patient, a personalized computational model of the left atrium was constructed from late gadolinium enhanced magnetic resonance imaging (LGE-MRI) to simulate AF induction via rapid pacing. Features were derived from the results of simulations (SimAF) as well as from the raw LGE-MRI images to serve as inputs into a quadratic discriminant analysis classifier. Features from SimAF were chosen in two ways: i) based on general knowledge of AF dynamics, and ii) left to be chosen by the ML training algorithm, unsupervised. The ML classifier was trained, optimized, and validated with 10-fold nested cross validation, resulting in the validation receiver operating characteristic curve shown (right).

Figure 2. Mechanistic modeling results of AF induction in patient-specific LA models. **(a)** Reconstructed 3-dimensional (3D) left atrial model for patient 2. Pacing locations which did (yellow) and did not (red) induce reentrant drivers (RD) and/or macro-reentrant tachycardias (MAT) (left). Activation map showing MAT around the left inferior pulmonary vein (right). **(b)** Reconstructed 3D left atrial model for patient 4 (left). Activation map showing RD adjacent to mitral valve on posterior left atrium adjacent to the mitral valve (right). **(c)** Reconstructed 3D atrial model for patient 7 (far left). Activation maps showing RD inferior to the pulmonary vein (middle left), RD at left inferior pulmonary vein (middle right), and MAT around right inferior pulmonary vein (far right). **(d)** Number of AF-inducing pacing sites for each patient vs.

pulmonary vein isolation (PVI) outcome. **(e)** Number of RDs and MATs for each patient vs. PVI outcome. **(f)** Receiver operating characteristic curves for prediction of AF recurrence after PVI using the number of AF-inducing pacing sites (solid red line) and number of RDs and MATs (dashed blue line). reentrant driver (RD), macro-reentrant tachycardia (MAT), area under the curve (AUC).

Figure 3. Prediction of AF recurrence post-PVI using ML on Features Derived from raw LGE-MRI and SimAF. Training **(a)**, validation **(b)**, and testing **(c)** receiver operating characteristic (ROC) curve for quadratic discriminant analysis classifier trained using features derived from imaging and SimAF (solid yellow line), imaging only (dashed blue line), and SimAF only (dashed red line). ROC curves were calculated by aggregating the results of 10-fold nested cross validation. simulations of atrial fibrillation induction (SimAF), imaging (I), area under the curve (AUC).

Figure 4. Analysis of inductive features learned by the ML algorithm from SimAF. **(a)** Reentry (reentrant driver, RD, and macro-reentrant tachycardia, MAT) locations most predictive of atrial fibrillation (AF) recurrence (red). **(b)** RD locations most predictive of AF recurrence (yellow). **(c)** MAT locations most predictive of AF recurrence (gray). (d) Locations (in purple) of the highest proportion of reentry-inducing pacing sites found to be predictive of AF recurrence. **(e)** The proportion of pacing sites in these locations (light blue) inducing MAT was predictive of AF recurrence. **(f)** Frequency with which each characteristic of SimAF was used to calculate an inductive feature selected by the random forest for inclusion in the ML classifier in the 10 outer

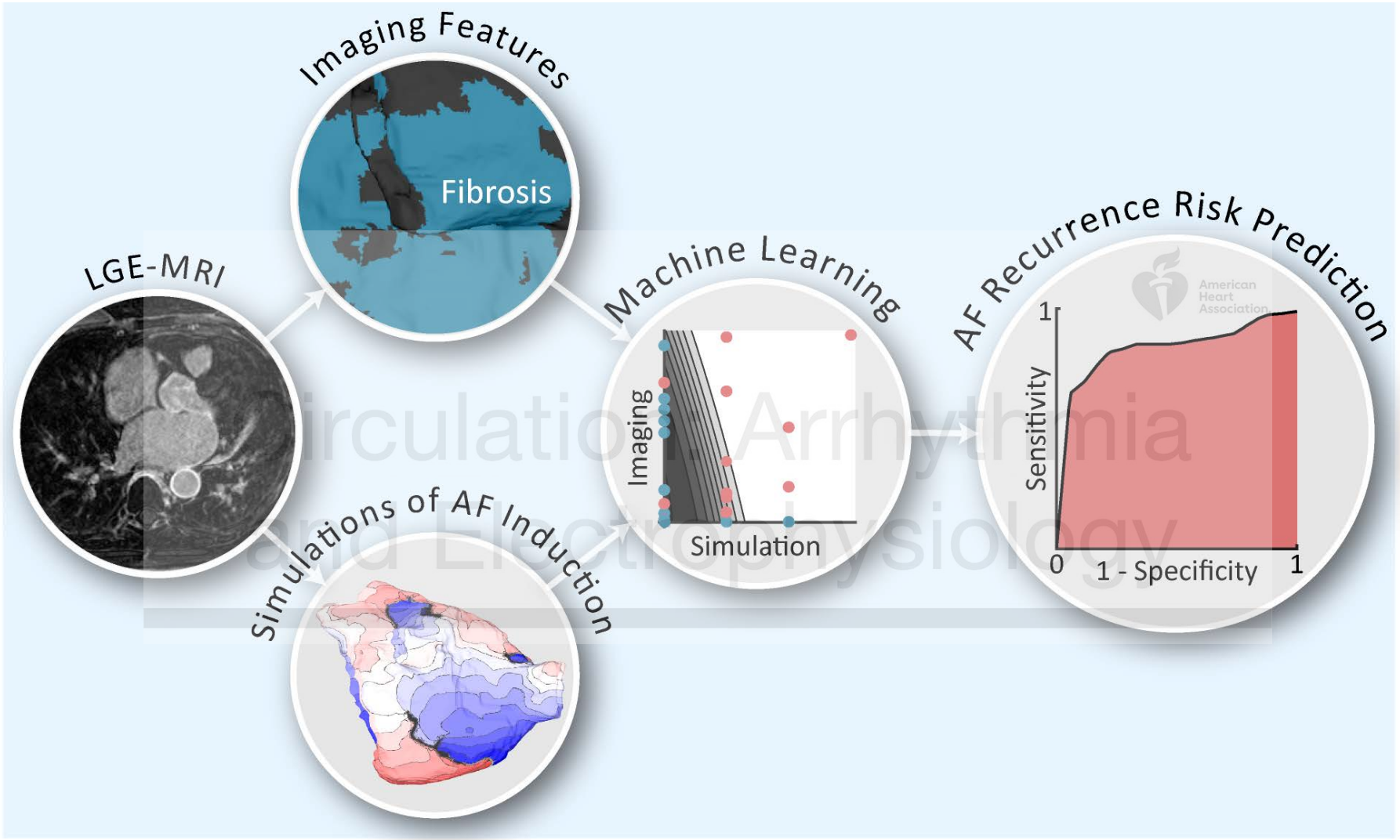
loops of cross validation. mitral valve (MV), left superior pulmonary vein (LSPV), left inferior pulmonary vein (LIPV), left atrial appendage (LAA), right inferior pulmonary vein (RIPV)

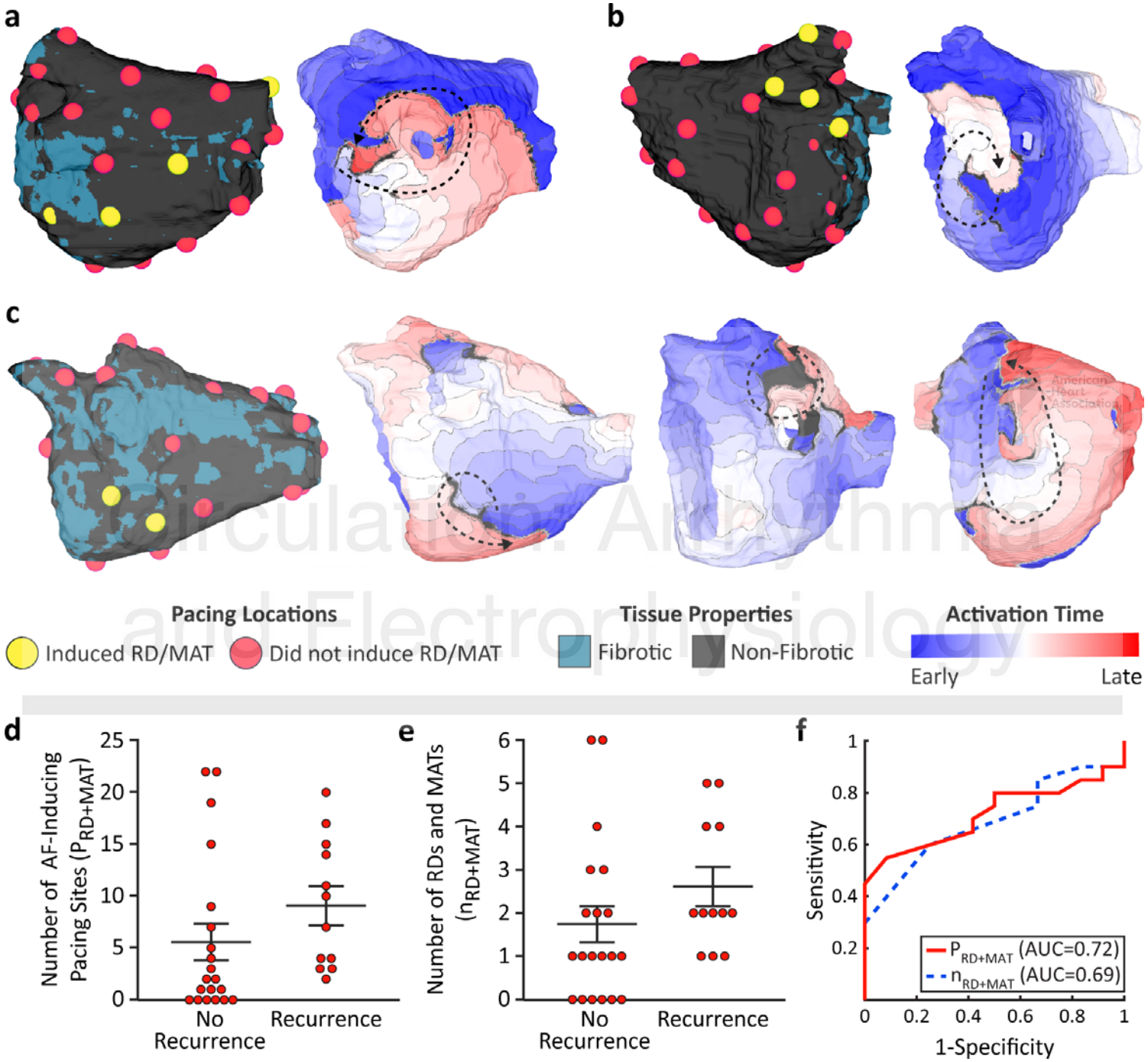
What is known?

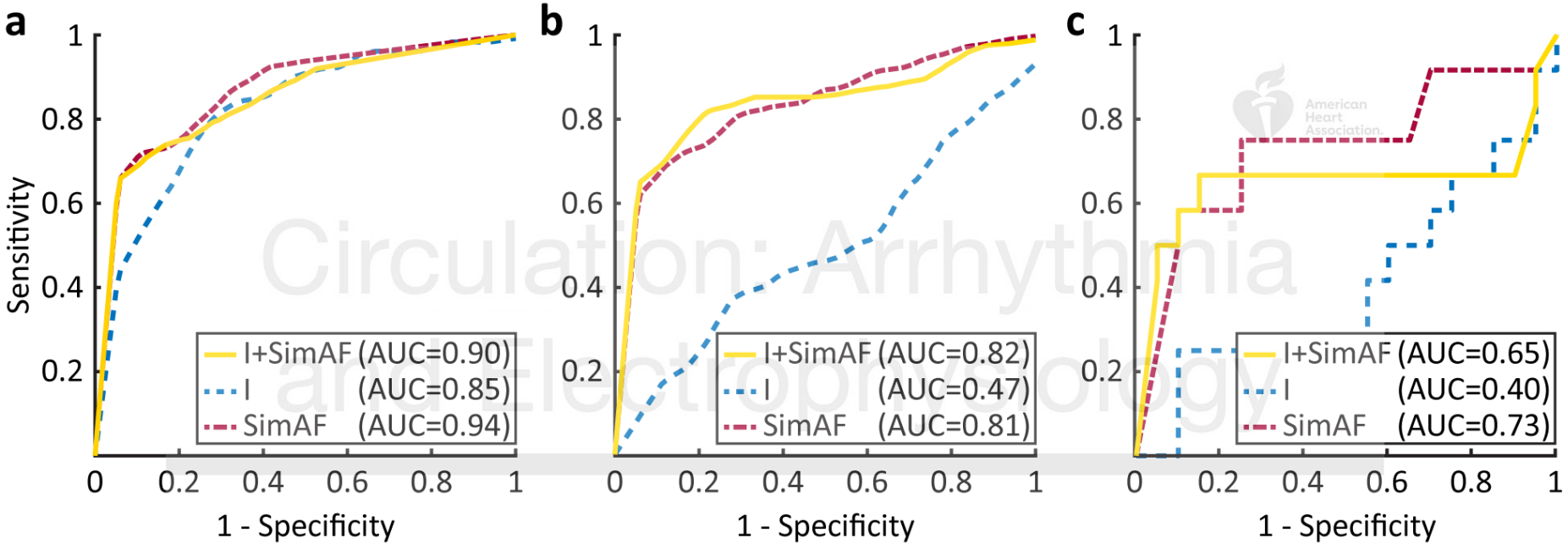
- In patients with paroxysmal atrial fibrillation (AF), recurrence of AF after ablation treatment via pulmonary vein isolation (PVI) is high.
- Atrial fibrosis promotes re-entrant activity underlying AF, thus presence of fibrosis in the atria could be a factor contributing to AF recurrence after PVI.
- Currently, there are no approaches to predict, before the ablation procedure, the individual patient's probability of AF recurrence post-PVI.

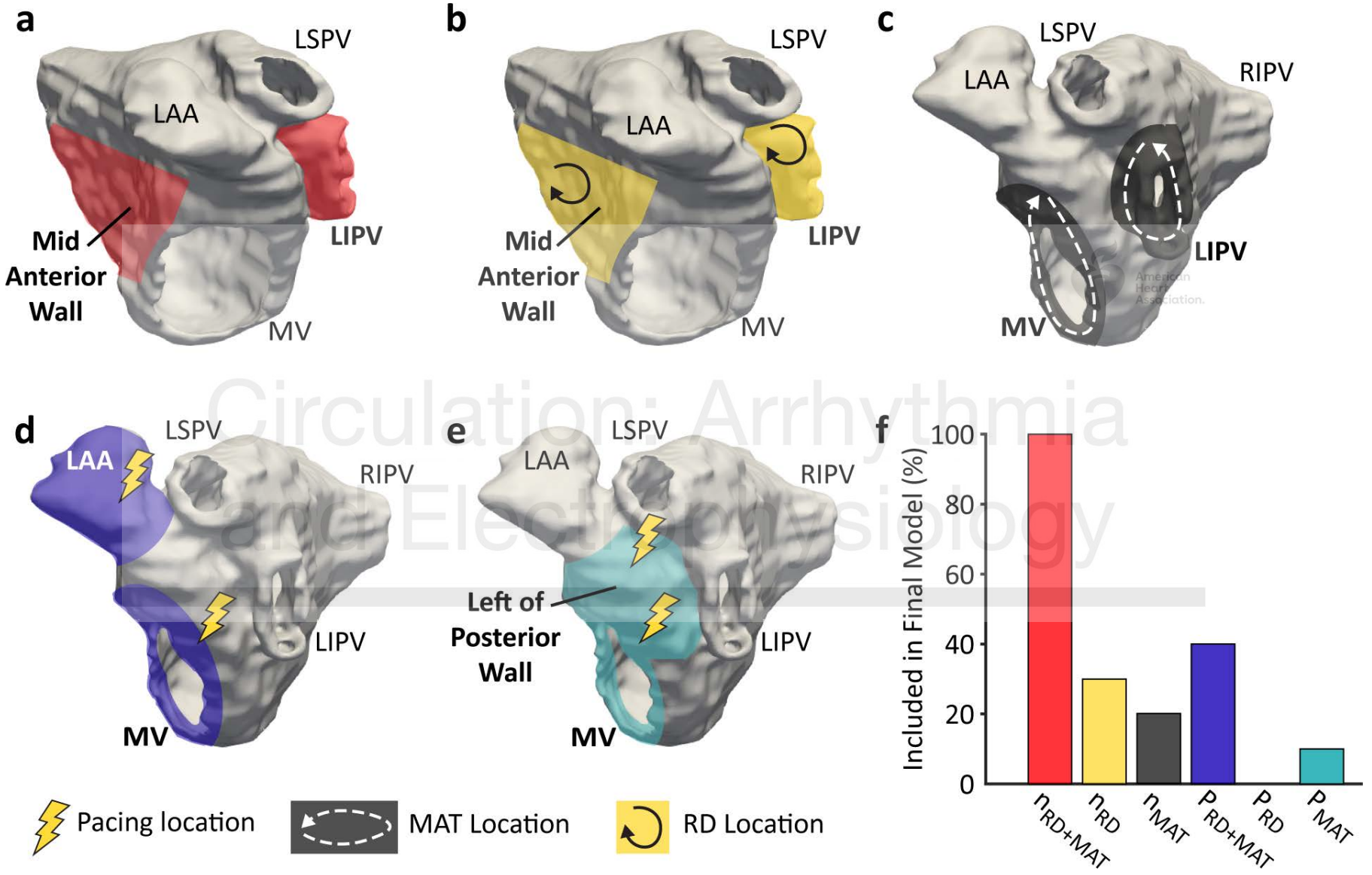
What this study adds:

- A personalized approach to predict pre-procedure the probability of AF recurrence after PVI is developed, which combines machine learning and MRI-image-based computational modeling of AF inducibility in the fibrotic atrial substrate.
- The machine learning (ML) algorithm uses as input features from the patient's LGE-MRI and from the results of the personalized mechanistic simulations.
- The optimized ML algorithm predicts AF recurrence following PVI with an average validation sensitivity of 82%, specificity of 89%, and area under the curve of 0.82.
- The inclusion, in the ML algorithm, of features extracted from the results of personalized mechanistic simulations of AF inducibility results in a highly generalizable AF recurrence prediction even for a small training data set.









Graphic Abstract