Data mining experiments on the Angiotensin II-Antagonist in Paroxysmal Atrial Fibrillation (ANTIPAF-AFNET 2) trial: 'exposing the invisible'

Sercan Okutucu[1]*, Deniz Katircioglu-Öztürk[2], Emre Oto[2], H. Altay Güvenir[3], Ergun Karaagaoglu[4], Ali Oto[1], Thomas Meinertz[5], and Andreas Goette[5,6]

[1]Department of Cardiology, Memorial Ankara Hospital, Ankara, Turkey; [2]Medical Information Technology Solutions (MITS), Bilkent University Cyberpark, Ankara, Turkey; [3]Department of Computer Engineering, Faculty of Engineering, Bilkent University, Ankara, Turkey; [4]Department of Biostatistics, Faculty of Medicine, Hacettepe University, Ankara, Turkey; [5]Atrial Fibrillation Network Association, Münster, Germany; and [6]Department of Cardiology, Vincenz-Krankenhaus, Paderborn, Germany

Aims	The aims of this study include (i) pursuing data-mining experiments on the Angiotensin II-Antagonist in Paroxysmal Atrial Fibrillation (ANTIPAF-AFNET 2) trial dataset containing atrial fibrillation (AF) burden scores of patients with many clinical parameters and (ii) revealing possible correlations between the estimated risk factors of AF and other clinical findings or measurements provided in the dataset.
Methods	*Ranking Instances by Maximizing the Area under a Receiver Operating Characteristics* (ROC) *Curve* (RIMARC) is used to determine the predictive weights (P_w) of baseline variables on the primary endpoint. Chi-square automatic interaction detector algorithm is performed for comparing the results of RIMARC. The primary endpoint of the ANTIPAF-AFNET 2 trial was the percentage of days with documented episodes of paroxysmal AF or with suspected persistent AF.
Results	By means of the RIMARC analysis algorithm, baseline SF-12 mental component score ($P_w = 0.3597$), age ($P_w = 0.2865$), blood urea nitrogen (BUN) ($P_w = 0.2719$), systolic blood pressure ($P_w = 0.2240$), and creatinine level ($Pw = 0.1570$) of the patients were found to be predictors of AF burden. Atrial fibrillation burden increases as baseline SF-12 mental component score gets lower; systolic blood pressure, BUN and creatinine levels become higher; and the patient gets older. The AF burden increased significantly at age >76.
Conclusions	With the ANTIPAF-AFNET 2 dataset, the present data-mining analyses suggest that a baseline SF-12 mental component score, age, systolic blood pressure, BUN, and creatinine level of the patients are predictors of AF burden. Additional studies are necessary to understand the distinct kidney-specific pathophysiological pathways that contribute to AF burden.
Keywords	Atrial fibrillation • Blood urea nitrogen • Creatinine • Data mining • Machine learning • RIMARC • SF-12

Introduction

Atrial fibrillation (AF) is the most common sustained arrhythmia. It is associated with relevant excess morbidity and mortality.[1,2] So far, we are unable to prevent many of the severe complications associated with AF, despite antithrombotic therapy and management of concomitant heart disease.[1–3] Specifically, the perceived benefit of rhythm control therapy by antiarrhythmic drugs appears to be offset by proarrhythmic side effects. Recently, Angiotensin II-Antagonist in Paroxysmal Atrial Fibrillation (ANTIPAF-AFNET 2)

trial examined the hypothesis that blocking the angiotensin II type 1 receptor with olmesartan medoxomil reduces the incidence of episodes of AF in patients with paroxysmal AF during 12 months by 25% compared with standard medication without angiotensin receptor blocker (ARB) therapy in a prospective, randomized, placebo-controlled, double-blind trial.[4] This trial revealed that 1 year of ARB therapy did not reduce the number of AF episodes in patients with documented paroxysmal AF without structural heart disease.

Data mining is the computational process that takes much of its inspiration and methods from the intersection of artificial

* Corresponding author. Tel: +90 312 2536666; fax: +90 312 2536623. *E-mail address:* sercanokutucu@yahoo.com

---

What's new?

- Data-mining analyses of Angiotensin II-Antagonist in Paroxysmal Atrial Fibrillation (ANTIPAF-AFNET 2) trial dataset with *Ranking Instances by Maximizing the Area under a Receiver Operating Characteristics Curve* and chi-square automatic interaction detector algorithm suggest that:
    o AF burden increases as
        § Baseline SF-12 mental component score gets lower
        § Systolic blood pressure, BUN and creatinine levels become higher and the patient gets older
    o The AF burden increased significantly at age >76.

---

intelligence, machine learning, statistics, and database systems for discovering previously unknown patterns in large datasets.[5,6] The overall goal is to extract valuable knowledge from information rich yet knowledge poor datasets and transform it into human-readable and applicable rule-bases for further use in various domains, including the healthcare and decision support systems.[5,6] Being positioned slightly different from the hypotheses-dependent statistical analyses, but incorporating all the statistical methods within, data mining generates novel hypotheses in both a supervised and an unsupervised nature. Aside from the raw analysis step, it involves data management and pre-processing aspects, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures and generated hypotheses, visualization, and online updating.[5,6] In practice, however, most hypotheses generation tasks require automated intelligence to induce new knowledge from tacit relationships among observations. For that matter, data-mining approaches often refer to machine learning algorithms designed and optimized to extract knowledge in an unguided manner for surfacing the effects of relationships that have not been evaluated adequately and for the accurate prediction of the future observations in the applied domain.

Here, in the context of data-mining approach, we applied machine learning methods to determine predictors of AF burden in the ANTIPAF-AFNET 2 dataset. We primarily used the RIMARC [Ranking Instances by Maximizing the Area under a Receiver Operating Characteristics (ROC) Curve] algorithm[5] to determine the predictive weights of the clinical features (variables) on AF burden and also used the CHAID (CHi-squared Automatic Interaction Detection) decision tree algorithm as a supplementary approach.[7,8] RIMARC algorithm operates by ranking instances based on how likely they are to have a designated label. By means of these analyses, we tried to extract clinically relevant information from the ANTIPAF-AFNET 2 database and seek some factors that might affect AF burden, which is the primary outcome of the trial.

Methods

Study design and participants

The ANTIPAF-AFNET 2 was a prospective, randomized, placebo-controlled, multicentre trial analysing the AF burden (percentage of days with documented episodes of paroxysmal AF) during a 12-month follow-up as the primary study endpoint. The trial was conducted by the German AFNET as the sponsor.[9,10]

Four hundred thirty patients with documented paroxysmal AF without structural heart disease were randomized to placebo or 40 mg olmesartan per day. Concomitant therapy with ARBs, angiotensin-converting enzyme inhibitors, and antiarrhythmic drugs was prohibited. Patients were followed up using daily trans-telephonic ECG (tele-ECG) recordings independent of symptoms. Of note, more than 80 000 tele-ECGs were recorded in the 430 patients throughout the trial, which is a unique feature of that investigation. Details of the trial design have been reported previously. A full description of ANTIPAF-AFNET 2 trial can be found elsewhere.[1,4]

Data-mining experiments

To attain our aims on this dataset and extract patterns and relationships within, we pursued a data-mining approach with two different machine learning algorithms. First one is the RIMARC (*Ranking Instances by Maximizing the Area under a Receiver Operating Characteristics* (ROC) *Curve*) classification algorithm[4] that was used to assign 'predictive weights' (having values between [0, 1]) to the baseline clinical parameters in determining the class label, i.e. AF burden. To calculate these predictive weights, RIMARC basically learns a ranking function over the instances by maximizing the area under the ROC curve, as this is a commonly accepted metric for assessing the accuracy of the results produced by a classifier.[5,6] It comprises a method, MAD2C, that applies a discretization to the continuous (real-valued) parameters in the dataset and transforms them into categorical parameters with value ranges generating a maximal AUC. Thus, RIMARC algorithm starts by discretizing all the continuous (real-valued) parameters until the whole dataset is made up of categorical typed parameters. The emphasis laid on the robustness of RIMARC towards the missing values in a dataset is also attributed to the MAD2C method for discretizing the continuous (real-valued) parameters.[5,6] As the result of a RIMARC execution, predictive weight for each parameter (now discretized) is calculated and respective value ranges are provided for an optimal AUC.

In this study, the class variable is AF burden and baseline variables are all other baseline clinical parameters that affect AF burden. The primary endpoint of the study was the percentage of days with documented episodes of paroxysmal or with suspected persistent or permanent AF. The AF burden was calculated as the number of days with paroxysmal AF or with preceding documentation of suspected persistent AF (up to a maximum of 365 days) divided by the number of measurement days, that is, days in follow-up with at least one readable tele-ECG recording (up to a maximum 365 days). Regarding these, AF burden is a valid choice for the class variable to be used in our experiments. Apart from AF burden, the ANTIPAF-AFNET 2 dataset contains 23 baseline clinical parameters for a total of 425 patients. The clinical parameters are shown in *Table 1*. To build a classifier model, a categorical class variable is needed indicating a discriminative condition over the instances; therefore, we applied a thresholding that assigns the two categories of 'Normal' (N) and 'Patient' (P) for each sample based on their corresponding AF burden values. The AF burden score <0.10 is set as N and >0.10 is set as P. Following this class label discretization, RIMARC algorithm is applied to assign predictive weights to each clinical parameter for determining the AF burden of a patient.

In this dataset, clinical parameters have an overall missing value rate of ~23%. To compare RIMARC results with the widely used decision tree classification method, we must choose a technique that is also robust to

missing values like RIMARC. In the CHAID (Chi-Square Automatic Interaction Detector) decision tree classification algorithm within which a chi-square test is used to build the tree, missing values are treated as a set of separate predictor category.[8] Unlike many other decision tree classification approaches generating binary node splits, CHAID generates multi-node split decision nodes with categorization of continuous (real-valued) parameters.[8] Similar to the way RIMARC acts, the algorithm first generates the 'best' set of parameter categories performing chi-square tests using all non-missing values from the dataset. Next, not to disregard missing valued instances, it identifies the category that is most similar to the 'missing' category in hand. Finally, it decides whether to merge the missing category with its most similar category or to keep the missing category as a separate category.[7,8]

While building its decision tree with multiple node splits, CHAID executes a pre-pruning approach that ensures the elimination of any redundant nodes. A node is made to split further down only when a significance criterion is fulfilled and thus the incident of overfitting is prevented from happening right from the start. The output of a CHAID decision tree can also be interpreted as a rule base effectively leading an instance towards a prediction through the divided categories for each parameter defining the dataset.[5–8]

Table I Baseline clinical parameters in the analysis

Age	Ischemic heart disease (Y/N)
Gender (F/M)	Diabetes mellitus (Y/N)
SF-12 physical component score	Aspirin (Y/N)
SF-12 mental component score	Verapamil (Y/N)
Systolic blood pressure (mmHg)	Diltiazem (Y/N)
Diastolic blood pressure (mmHg)	Statin (Y/N)
Left ventricular ejection fraction (%)	Diuretic (Y/N)
Blood urea nitrogen (mg/dL)	Tri-tetracyclic antidepressant (Y/N)
Creatinine (mg/dL)	Oral anticoagulants (Y/N)
Glomerular filtration rate	Dihydropyridin (Y/N)
Hypertension (Y/N)	Nitrate (Y/N)
NYHA Class IV (Y/N)	

Results

The results of the application of the RIMARC algorithm on the ANTIPAF-AFNET 2 trial dataset are presented in Table 2, which tabulates the clinical parameters with their respective predictive weights. To assess the significance of the RIMARC's model, a 10-fold cross-validation is performed. For each patient instance in the dataset, the predictive probabilities regarding the class parameter AF burden is calculated. To measure the accuracy characterized by the sensitivity and specificity for this predictive model, an ROC curve is generated with the respective c-statistics as the AUC value (Figure 1). The AUC of the ROC curve with the value of 0.815 (standard deviation = 0.046, 95%CI, $P = 0.001$) can be interpreted as a decent result. This can be attributed to the robustness of the RIMARC algorithm towards the missing values in the datasets.

According to the CHAID decision tree classifier, the root parameter, which is the most discriminative among the other parameters, is found to be the SF-12 mental component score of the patients ($P_w = 0.3597$) (Figure 2). With this CHAID tree classifier built, a 10 fold cross-validation is performed and the prediction accuracy of the model is assessed by the AUC value of ROC curve generated. The AUC value is found as 0.614 (standard deviation = 0.060, 95%CI, $P = 0.001$).

Baseline SF-12 mental component score, age, BUN, systolic blood pressure, and creatinine level of the patients were found to be predictive of AF burden by the RIMARC algorithm. The CHAID decision tree technique also confirms the effect of baseline SF-12 mental component score on indicating AF burden as a single predictive parameter among all others.

As BUN ($P_w = 0.2719$), systolic blood pressure ($P_w = 0.2240$), and creatinine ($P_w = 0.1570$) levels of the patient increases, the AF burden also increases (Figure 3). Furthermore, the risk of AF burden increases as the patient gets older ($P_w = 0.2865$). The risk increases significantly at a higher rate after age of 76.

Discussion

Using a combination of explorative data-mining analyses, we identified that SF-12 mental component score, age, BUN, systolic blood

Table 2 RIMARC based predictive weights (P_w) of the parameters on atrial fibrillation burden

| Parameter	P_w	Parameter	P_w
Baseline SF-12 mental component score	0.3597	Aspirin	0.0418
Age	0.2865	Diuretic	0.0367
Blood urea nitrogen (mg/dL)	0.2719	Ischaemic heart disease	0.0358
Systolic blood pressure (mmHg)	0.2240	Nitrate	0.0306
Creatinine (mg/dL)	0.1570	Sex	0.0301
Left ventricular ejection fraction (%)	0.1453	NYHA class IV	0.0216
SF-12 physical component score	0.1379	Dihydropyridin	0.0200
Diastolic blood pressure (mmHg)	0.1244	Verapamil	0.0137
Oral anticoagulants	0.0882	Diltiazem	0.0079
Hypertension	0.0596	Tri-tetracyclic antidepressant	0.0079
Statin	0.0564	Glomerular filtration rate	0.0066
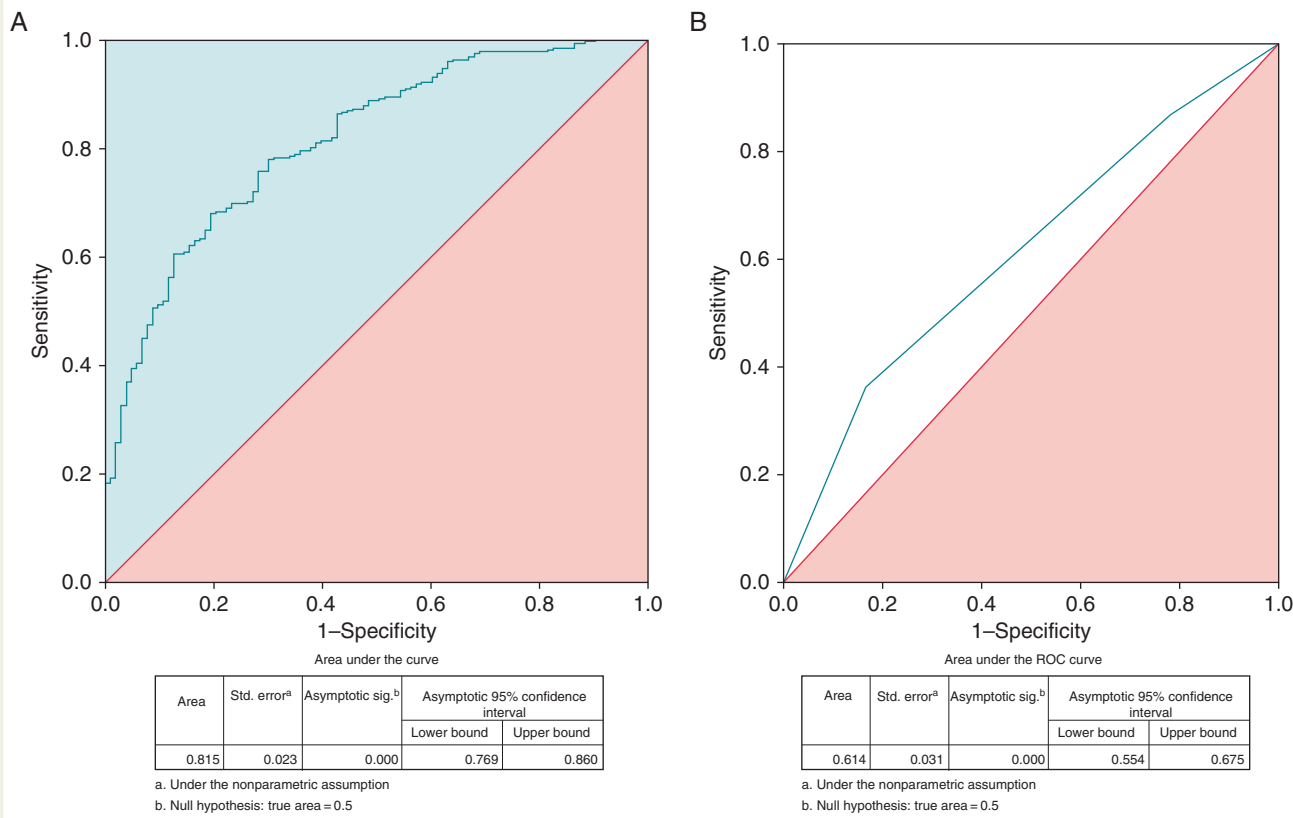
Figure 1 (*A*) Receiver operating characteristic curve with the respective AUC value for RIMARC results. (*B*) Receiver operating characteristic curve the respective AUC value for CHAID tree classifier.
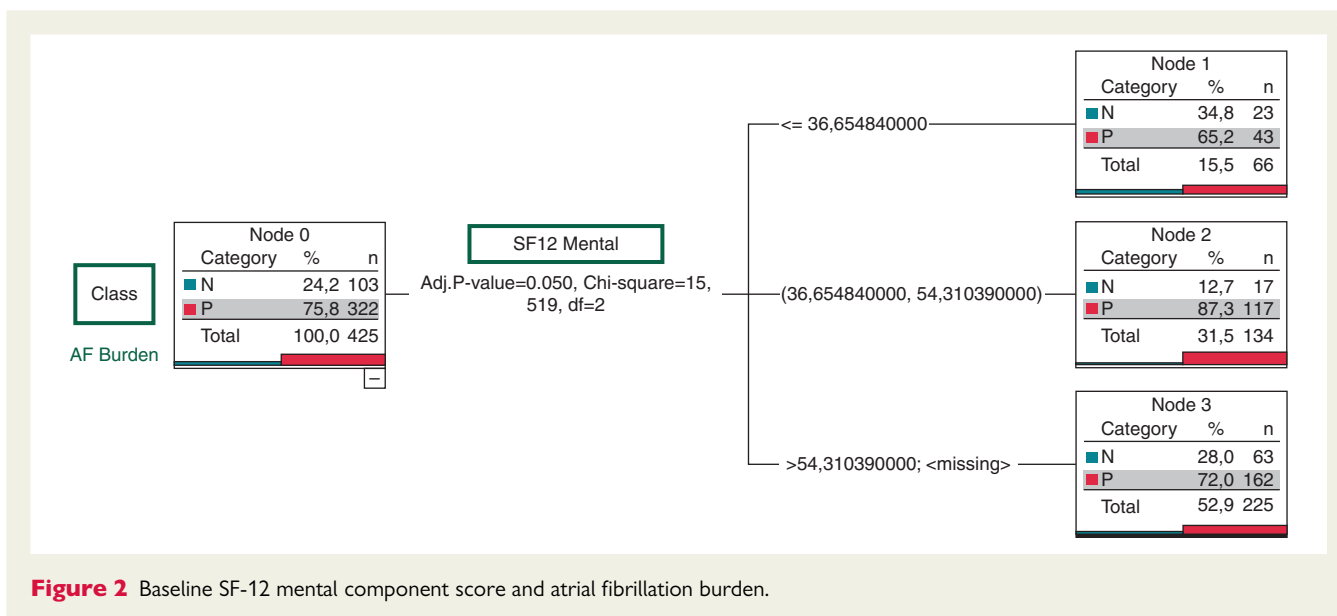

Figure 2 Baseline SF-12 mental component score and atrial fibrillation burden.

pressure, and creatinine level of the patients are predictive of AF burden. On the basis of these findings, some additional insight into the AF burden and AF treatment is obtained.

Major finding of this analysis is the predictive power of baseline SF-12 mental component score on AF burden. The Short-Form 12 Health Survey is a generic health-related quality-of-life (QOL)
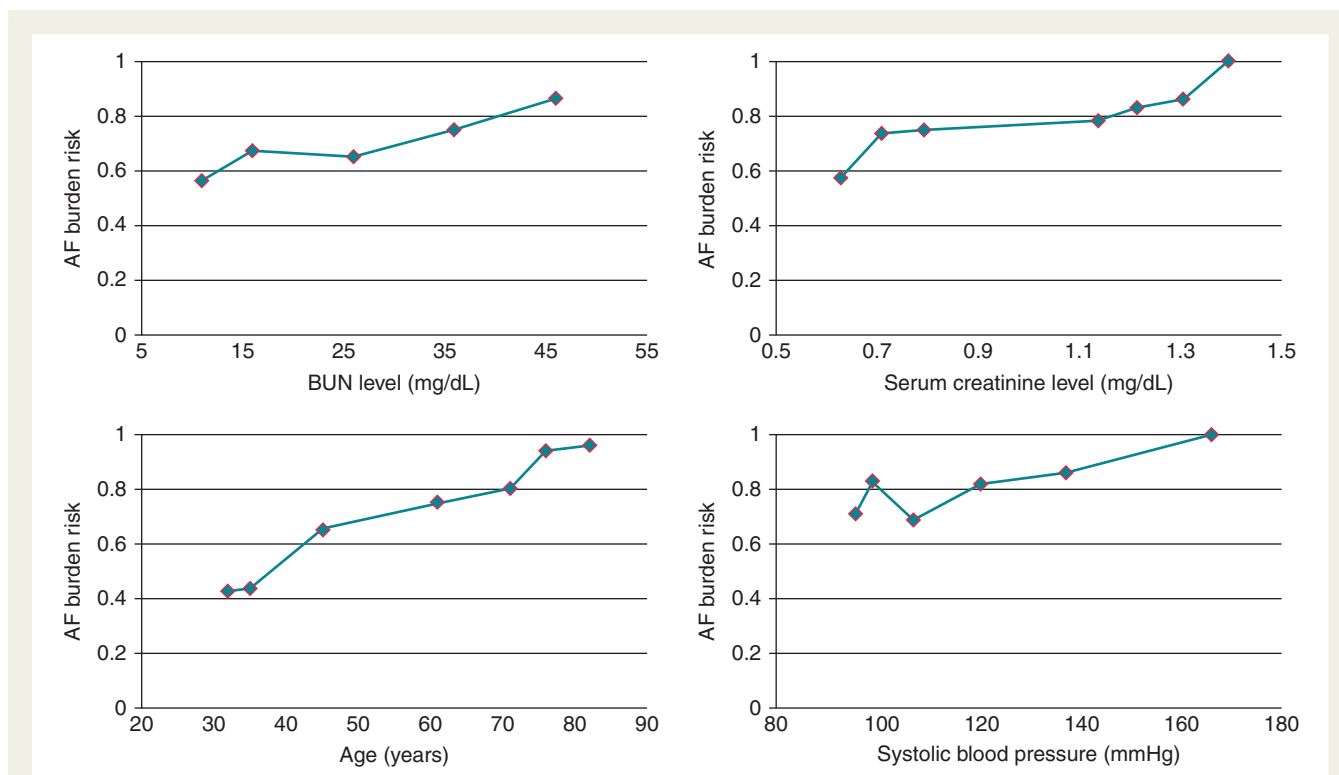
Figure 3 Atrial fibrillation burden risk chart according to clinical parameters.

instrument. The items of the SF-12 assess physical component and mental component.[11,12] Patients with AF have significantly poorer QOL compared with healthy controls, the general population, and other patients with coronary heart disease. Studies examining rate or rhythm-control strategies alone demonstrate improved QOL after intervention.[13] The cornerstone of treatment in patients with AF is to reduce symptoms and improve the QOL.[14] Three of the four large randomized control trials (STAF,[15] PIAF,[16] RACE[17]) comparing rate vs. rhythm control demonstrated a greater improvement in QOL in patients receiving rate control. However, the AF-FIRM trial[18] revealed a similar improvement in QOL for both rate and rhythm-control groups. In recent analysis of two large clinical trials, reported by von Eisenhart Rothe et al.,[11] AF patients prone to experiencing depressed mood, particularly in persistent ones. In accordance with our data-mining analysis, von Eisenhart Rothe et al.[12] reported association of depressed mood with AF symptom burden over 6 months after adjustment for perceived frequency and duration of AF episodes, pulmonary diseases, and gender. In current analysis, we obtained SF-12 mental component score as a predictor of AF burden by the RIMARC algorithm. Furthermore, baseline SF-12 mental component score was the only single predictor of AF burden among all others by CHAID decision tree technique. This finding denotes that a lower QOL at baseline is the predictor of high AF burden.

The second important finding of this analysis is related to the intersection of renal function and AF burden. By means of explorative analyses, BUN and serum creatinine level of the patients were found to be predictors of AF burden. Several possible mechanisms may explain the high rate of identified AF in patients with chronic kidney disease (CKD), including older age and a high burden of risk factors such as hypertension and cardiovascular disease, excessive inflammation which has been linked to both CKD and AF, larger left atrial and left ventricular sizes among CKD patients and activation of the renin−angiotensin−aldosterone system.[19] Other plausible pathways linking kidney disease and AF include abnormalities in mineral metabolism. It is possible that alterations in these pathways may also contribute to the risk of AF in patients with renal dysfunction through effects on cardiac structure, endothelial function, and vascular calcification.[19] The burden of AF is even greater in patients with concomitant kidney disease. Recently published studies have highlighted the often under recognized, yet highly prevalent relation between kidney disease and AF. Furthermore, evidence has suggested that the burden of AF will likely rise in this high-risk population, making the intersection of kidney disease and AF a highly relevant clinical problem.[20] Further investigations are needed to explore unique kidney-specific biological pathways linking AF and kidney disease, given the disproportionately high burden of disease in this population.

The prevalence of AF is related to age. The prevalence of AF is 2.5% in people older than 40 years and 6% in those older than 65 years. Approximately 70% of individuals with AF are between 65 and 85 years of age.[2] The relationship between AF burden and age was remarkable in our analysis. The risk of AF burden increases as patient gets older, and AF burden risk increases significantly at a higher rate after the age of 76.

Atrial fibrillation and hypertension are two prevalent, and often coexistent, conditions in the general population.[2,21] Both these conditions frequently coexist and their prevalence increases rapidly

with ageing. Hypertension is still the main risk factor for the development of AF. Hypertension is associated with left ventricular hypertrophy, impaired ventricular filling, left atrial enlargement, and slowing of atrial conduction velocity.[21] In our analysis, we observed that systolic blood pressure levels of patients predict AF burden.

Conclusions

In conclusion, on the ANTIPAF-AFNET 2 dataset, the RIMARC algorithm helped reveal the predictive power of various parameters on AF, along with the risk scores of categorical values and risk ranges for numerical parameters. Based on the highest weighted parameters found by RIMARC, some additional insight into the AF burden and AF treatment is obtained. QOL is of central importance in AF as both a treatment goal and an endpoint in the evaluation of therapies. A number of interventions for AF have been shown to improve QOL, including pharmacological and non-pharmacological rate control, antiarrhythmic drugs, and non-pharmacological rhythm control strategies. Collection of further data is needed to establish the role of QOL on the course of AF. Additional studies are necessary to understand the distinct kidney-specific pathophysiological pathways that contribute to the development of AF as well as the unique considerations in preventing and treating AF specific to patients with a broad range of renal dysfunction.

Funding

Conflict of interest: none declared.

References

1. Goette A, Breithardt G, Fetsch T, Hanrath P, Klein HU, Lehmacher W et al. Angiotensin II antagonist in paroxysmal atrial fibrillation (ANTIPAF) trial: rationale and study design. *Clin Drug Investig* 2007;**27**:697–705.
2. Camm AJ, Lip GY, De Caterina R, Savelieva I, Atar D, Hohnloser SH et al. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation—developed with the special contribution of the European Heart Rhythm Association. *Europace* 2012;**14**:1385–413.
3. Savelieva I, Kakouros N, Kourliouros A, Camm AJ. Upstream therapies for management of atrial fibrillation: review of clinical evidence and implications for European Society of Cardiology guidelines. Part II: secondary prevention. *Europace* 2011;**13**:610–25.
4. Goette A, Schon N, Kirchhof P, Breithardt G, Fetsch T, Hausler KG et al. Angiotensin II-antagonist in paroxysmal atrial fibrillation (ANTIPAF) trial. *Circ Arrhythm Electrophysiol* 2012;**5**:43–51.
5. Ravens U, Katircioglu-Ozturk D, Wettwer E, Christ T, Dobrev D, Voigt N et al. Application of the RIMARC algorithm to a large data set of action potentials and clinical parameters for risk prediction of atrial fibrillation. *Med Biol Eng Comp* 2015;**53**:263–73.
6. Feng Y, Wang Y, Guo F, Xu H. Applications of data mining methods in the integrative medical studies of coronary heart disease: progress and prospect. *Evid Based Complement Alternat Med* 2014;**2014**:791841.
7. Kobayashi D, Takahashi O, Arioka H, Koga S, Fukui T. A prediction rule for the development of delirium among patients in medical wards: Chi-Square Automatic Interaction Detector (CHAID) decision tree analysis model. *Am J Geriatr Psychiatry* 2013;**21**:957–62.
8. Miller B, Fridline M, Liu PY, Marino D. Use of CHAID decision trees to formulate pathways for the early detection of metabolic syndrome in young adults. *Comput Math Methods Med* 2014;**2014**:242717.
9. Breithardt G, Dobrev D, Doll N, Goette A, Hoffmann B, Kirchhof P et al. The German Competence Network on Atrial Fibrillation (AFNET). *Herz* 2008;**33**:548–55.
10. Leute A, Kirchhof P, Breithardt G, Goette A, Lewalter T, Meinertz T et al. [German Competence Network on Atrial Fibrillation (AFNET). A nationwide cooperation for better research and patient care]. *Med Klin (Munich)* 2006;**101**:662–6.
11. von Eisenhart Rothe AF, Goette A, Kirchhof P, Breithardt G, Limbourg T, Calvert M et al. Depression in paroxysmal and persistent atrial fibrillation patients: a cross-sectional comparison of patients enroled in two large clinical trials. *Europace* 2014;**16**:812–9.
12. von Eisenhart Rothe A, Hutt F, Baumert J, Breithardt G, Goette A, Kirchhof P et al. Depressed mood amplifies heart-related symptoms in persistent and paroxysmal atrial fibrillation patients: a longitudinal analysis—data from the German Competence Network on Atrial Fibrillation. *Europace* 2015;**17**:1354–62.
13. Freeman JV, Simon DN, Go AS, Spertus J, Fonarow GC, Gersh BJ et al. Association between atrial fibrillation symptoms, quality of life, and patient outcomes: results from the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF). *Circ Cardiovasc Qual Outcomes* 2015;**8**:393–402.
14. Camm AJ. Quality of life in patients with atrial fibrillation. *Rev Esp Cardiol* 2010;**63**:1393–5.
15. Carlsson J, Miketic S, Windeler J, Cuneo A, Haun S, Micus S et al. Randomized trial of rate-control versus rhythm-control in persistent atrial fibrillation: the Strategies of Treatment of Atrial Fibrillation (STAF) study. *J Am Coll Cardiol* 2003;**41**:1690–6.
16. Hohnloser SH, Kuck KH, Lilienthal J. Rhythm or rate control in atrial fibrillation—pharmacological Intervention in Atrial Fibrillation (PIAF): a randomised trial. *Lancet* 2000;**356**:1789–94.
17. Hagens VE, Vermeulen KM, TenVergert EM, Van Veldhuisen DJ, Bosker HA, Kamp O et al. Rate control is more cost-effective than rhythm control for patients with persistent atrial fibrillation—results from the RAte Control versus Electrical cardioversion (RACE) study. *Eur Heart J* 2004;**25**:1542–9.
18. Steinberg JS, Sadaniantz A, Kron J, Krahn A, Denny DM, Daubert J et al. Analysis of cause-specific mortality in the Atrial Fibrillation Follow-up Investigation of Rhythm Management (AFFIRM) study. *Circulation* 2004;**109**:1973–80.
19. Allison SJ. Chronic kidney disease: ESRD risk in CKD patients with incident atrial fibrillation. *Nat Rev Nephrol* 2013;**9**:125.
20. Bautista J, Bella A, Chaudhari A, Pekler G, Sapra KJ, Carbajal R et al. Advanced chronic kidney disease in non-valvular atrial fibrillation: extending the utility of R2CHADS2 to patients with advanced renal failure. *Clin Kidney J* 2015;**8**:226–31.
21. Tremblay-Gravel M, Khairy P, Roy D, Leduc H, Wyse DG, Cadrin-Tourigny J et al. Systolic blood pressure and mortality in patients with atrial fibrillation and heart failure: insights from the AFFIRM and AF-CHF studies. *Eur J Heart Failure* 2014;**16**:1168–74.