Seeking predictors for paroxysmal atrial fibrillation in stroke with an online clinical database

C Sefik Evren Erdener
Hacettepe University Institute of Neurological Sciences and Psychiatry, Ankara, Turkey

ABSTRACT
OBJECTIVE: A considerable fraction of ischemic stroke cases remain cryptogenic and there is increasing data suggesting the role of missed paroxysmal atrial fibrillations (pAF) in at least a number of these cases. Since electrophysiological identification of pAFs can be challenging, there has been an accumulation of proposed predictors and biomarkers for pAFs. The predictive values of these is varying and sometimes conflicting among studies. Therefore, we aimed to verify a fraction of previously reported parameters for pAF detection by investigating an independent clinical sample.

METHODS: Using a publicly available data downloaded from the MIMIC-3 intensive care unit database, we tested the predictive role of particular risk factors and biomarkers for pAF detection after ischemic stroke in 124 patients with ischemic stroke admitted within 24 hours of stroke onset.

RESULTS: Our evaluation revealed a strong association of older age in women, as well as admission National Institutes of Health Stroke Scale (NIHSS) and discharge modified Rankin Scores (mRS) in both sexes for pAFs, in patients that were in sinus rhythm on admission. We also detected a trend for lower gender-adjusted hemoglobin in patients with pAF, although the difference was insignificant. On the other hand, we did not find any significant association of pAF detection with some other previously reported biomarkers: serum magnesium level, leukocyte count, neutrophil/lymphocyte ratio or left atrial dilatation.

CONCLUSION: Even though our analysis did not reveal a strong and specific biomarker to predict pAFs after stroke, it identified key risk factors. It may be necessary to consider the possibility of pAFs and perform rigorous evaluation to prevent further events of embolic stroke in female patients older than 75 years with more severe neurological deficits on admission and higher disability on discharge and also with relatively lower hemoglobin level. This first study from Turkey using clinical data from the MIMIC-3 database also demonstrates the value of publicized clinical data for confirmatory studies on various medical fields across the World.

Keywords: Clinical markers; database; embolism; paroxysmal atrial fibrillation; stroke.

Cite this article as: Erdener SE. Seeking predictors for paroxysmal atrial fibrillation in stroke with an online clinical database. North Clin Istanb

Ischemic stroke is a leading cause of death and disability worldwide and is a heterogeneous phenomenon with a range of different etiological factors [1]. While in clinical practice, it can usually be identified as thrombotic or embolic nature, many cases remain cryptogenic with no obvious etiology to be determined [2, 3]. Etiological identification is crucial for treatment decisions on secondary prophylaxis to prevent recurrent stroke events. Atrial fibrillation (AF), either persistent or intermittent (i.e., paroxysmal) is one of the most common factors leading to cardioembolic stroke, especially in the elderly [4, 5]. Many cryptogenic strokes are suspected to result from paroxysmal episodes of atrial fibrillation, and identification of AF can be critical as it shifts the stroke pro-
Phylaxis to an anticoagulation approach [4, 5]. Confirming the presence of paroxysmal atrial fibrillation (pAF) can be challenging due to its temporary and sometimes short-lasting nature. Approximately half of patients with AF have no structural heart disease, and therefore echocardiography has limited value although left atrial dilatation has been associated with AF [6]. Most of the standard diagnostic procedures have low sensitivity for pAF detection [7]. Extended durations of ECG recording, sometimes up to several weeks, may be necessary for detection of some episodes [8, 9], but these resources are not widely available. Therefore, there are many attempts to identify novel and easy-to-test biomarkers for pAF [10]. These risk factors and biomarkers can at least identify a fraction of patients that need to be extensively evaluated for pAF episodes, allowing allocation of clinical resources more efficiently. A number of predictors, such as advanced age and female sex [11] or blood parameters like anemia [11, 12], iron deficiency [13], hyperuricemia [14–16], serum potassium and magnesium [17, 18], leukocyte count [19], neutrophil-to-lymphocyte ratio [20, 21] and vitamin-D level [22] have been associated with AF previously. Specifically for ischemic stroke patients, N-terminal Pro-Brain Natriuretic Peptide (BNP) levels have been reported to predict pAF episodes with high sensitivity and specificity [23]. However, there are also conflicting studies with some of these factors, like discrepant results on association of neutrophil/lymphocyte ratio and serum magnesium levels with AF [24, 25]. Thus, it can be useful to confirm the predictive value of previously proposed factors for pAF detection specifically for stroke patients acquired from an independent clinical population.

In this study, we aimed to test the predictive power of certain parameters for pAF detection in acute ischemic stroke patients utilizing an online clinical database. Medical Information Mart for Intensive Care-3 (MIMIC-3) is the updated version an intensive care unit research database freely open to the medical research community [26, 27]. It was published in 2006 by the Computational Physiology Laboratory of the Massachusetts Institute of Technology (MIT), the Beth Israel Deaconess Medical Center (BIDMC) and the Philips Medical Center in the United States. A number of retrospective studies have been performed on this clinical dataset from several thousand intensive care unit (ICU) patients with multidisciplinary medical problems [28–34]. To our knowledge, our work is the first published study using data from stroke patients in MIMIC database. Besides being a confirmatory study for previously published predictors for AF, this work is an example for how these open-access clinical resources can be exploited for clinical research.

**MATERIALS AND METHODS**

The contents of the MIMIC database were downloaded from the Physionet website (http://physionet.org/physiobank/database/mimic3cdb) after receiving the necessary online trainings and signing the data use agreement. Ethical approval and consent to participate was not necessary and applicable for this study. The database contained clinical data, including medical history, physical examination, inpatient progress notes, diagnoses, laboratory and radiology results from 46520 ICU patients and 58976 admissions at the time it was accessed for this study. Diagnosis tables were searched to identify patients hospitalized for ischemic stroke, using keywords “stroke” “cerebrovascular accident” and “cva”. Initial list returned 1021 admitted patients. After initial checks on clinical records, 169 cases were excluded since they had in fact other diagnoses such as subarachnoid hemorrhage, intraparenchymal hemorrhage or stroke mimics. The remaining 852 cases were randomized and 150 of them were selected for analysis. Patients younger than 50 years of age, admitted beyond the first 24 hours of stroke onset, with no evidence of major arterial occlusion or with arterial dissection were excluded (26 cases). Analyses proceeded for the remaining 124 patients.

The database provided details on stroke etiological evaluation, including but not limited to magnetic resonance imaging (with angiography), Doppler ultrasound, electrocardiography (ECG), continuous telemetric monitoring and echocardiography. Blood workup in all cases included complete blood count (CBC), renal function tests, lipid profiles and serum electrolytes. We included hemoglobin, hematocrit, erythrocyte and leukocyte counts, serum potassium, magnesium, calcium for predictive tests in our study since these measurements were available in all admitted cases. Other previously reported predictive parameters, like BNP, vitamin-D or uric acid were not tested or results were not reported in database, and therefore they could not be analyzed in this study.

**Statistical Analyses**

Analyses were performed using IBM SPSS Statistics 23 software. All data are reported as mean±standard de-
RESULTS

Clinical characteristics of patients are presented in Table 1. Middle cerebral artery (MCA) was occluded in 103 cases (83.0%), internal carotid artery (ICA) in 14 (11.3%) while the remaining 7 patients had occlusions in anterior cerebral, posterior cerebral or basilar arteries. Mean hospitalization duration was 9.7±10.7 days. All patients were admitted within 24 hours of stroke onset. 42 patients (33.9%) had documented AF on admission. These patients also had AF rhythm in their subsequent ECGs indicating that they had persistent AF. On the other hand, telemetry monitoring revealed paroxysmal AF (pAF) in 31 (25.0%) patients who initially had sinus rhythm in their ECGs. Hence, in total, 73 cases (58.9%) had either persistent or paroxysmal AF documented during their hospital stay.

Etiological causes of stroke in these patients are shown in Table 2. Only one patient in AF group had a clearly atherosclerotic cause of ischemic stroke while others were classified as cardioembolic. In the group without AF (persistent or paroxysmal), 16 patients had other sources of cardioembolism while 17 patients had atherosclerotic stoke. Eighteen patients had no identifiable etiology and were classified as cryptogenic. It should be noted that if there was no telemetry monitoring, 31 patients with pAF would also have remained
cryptogenic. In patients with AF (both persistent and paroxysmal), we detected a significantly higher fraction of female sex, hypertension, left atrial dilatation, as well as older age, admission National Institutes of Health Stroke Scale (NIHSS) score and modified Rankin Score (mRS) (Table 1). There were no significant differences in leukocyte counts, lymphocyte/neutrophil ratios, serum potassium or magnesium levels between AF and no-AF groups, unlike a number of previous studies [17, 18, 21, 35]. Meanwhile, we detected a relatively lower hemoglobin levels after adjusting for sex in line with previous work [36] although the difference in our analysis was not significant.

Next, focusing on the parameters showing difference between AF and non-AF groups, we tested the predictive potential of those parameters for subsequent pAFs in patients with sinus rhythm on admission (n=82). These results are shown in Table 3. Univariate logistic regression revealed a significant association of pAF detection with patient age in females (OR=1.068, \( p=0.035 \)) and a near-significant association in males (OR=1.068, \( p=0.07 \)). Receiver operating characteristic (ROC) analysis showed a cut-off of 75 for females for 80% Sensitivity and 73% Specificity (Area under the curve (AUC)=0.699) (Fig. 1A, B). pAFs were more than twice as common in female patients (OR=2.562, \( p=0.044 \)). The effect of gender on pAF detection was dependent on age, however (Fig. 1C). Despite the lack of significance between ages of men and women (Fig. 1C), age-adjusted sex effect on pAF detection was not significant in logistic regression analysis either, (\( p=0.126 \)), supporting the contributing role of age on sex-related differences. After adjustments for age and sex, patients with pAF tended to have a significantly higher NIHSS score on admission (OR=1.222, \( p=0.025 \); AUC=0.767) and nearly-significantly higher mRS (OR=1.464, \( p=0.080 \)) on discharge although age was also found as a contributing factor for higher mRS (Pearson coefficient: 0.314, \( p=0.004 \)). ROC curve analysis for NIHSS scores showed a 70% sensitivity and 61% specificity for a cut-off level of 13 (AUC=0.767) (Fig. 1D).On the other hand, correlation analysis was significant between age and admission NIHSS (Pearson coefficient: 0.216, \( p=0.251 \)), signifying that, unlike the disability score of mRS, the symptomatic score of NIHSS was independent from age and sex.

Sex-adjusted hemoglobin levels again had a tendency to be lower in pAF groups, but the association did not reach statistical significance. At least in females, a contributing effect of hemoglobin on pAF cannot be excluded, and this is independent from age as hemoglobin level did not correlate with age in females (Pearson coefficient: 0.064, \( p=0.689 \)). In males, hemoglobin level was significantly correlated with older age (Pearson coefficient: -0.324, \( p=0.042 \)). Left atrial dilatation, in contrast, was not significantly different between pAF and non-pAF groups, again conflicting with previous reports [37]. This finding supports the view that LA diameter primarily serves as a marker of persistent AF [6]. Finally, past medical history of hypertension was also not significantly associated with pAF detection (\( p=0.157 \)), another factor differing from persistent AF.

<table>
<thead>
<tr>
<th>No pAF (N=51)</th>
<th>pAF (n=31)</th>
<th>P</th>
<th>OR</th>
<th>95% CI</th>
<th>ROC curve- AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (female) (%)</td>
<td>22 (43.1)</td>
<td>20 (64.5)</td>
<td>0.044</td>
<td>2.397</td>
<td>0.954–6.020</td>
</tr>
<tr>
<td>Age (male)</td>
<td>70.6±11.5</td>
<td>77.6±9.3</td>
<td>0.070</td>
<td>1.068</td>
<td>0.995–1.148</td>
</tr>
<tr>
<td>Age (female)</td>
<td>72.4±11.2</td>
<td>79.5±9.6</td>
<td>0.035</td>
<td>1.068</td>
<td>1.002–1.140</td>
</tr>
<tr>
<td>LA dilatation (%)</td>
<td>29 (61.7) (n=47)</td>
<td>17 (65.4) (n=26)</td>
<td>0.805</td>
<td>1.172</td>
<td>0.431–3.184</td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>35 (68.6)</td>
<td>26 (83.9)</td>
<td>0.131</td>
<td>2.375</td>
<td>0.771–7.299</td>
</tr>
<tr>
<td>Hemoglobin (male) (g/dL)</td>
<td>13.5±2.1</td>
<td>13±2.5</td>
<td>0.571</td>
<td>0.895</td>
<td>0.653–1.225</td>
</tr>
<tr>
<td>Hemoglobin (female) (g/dL)</td>
<td>12.9±1.5</td>
<td>12.4±1.5</td>
<td>0.304</td>
<td>0.798</td>
<td>0.521–1.222</td>
</tr>
<tr>
<td>Admission NIHSS score</td>
<td>11.7±5.6</td>
<td>17±4.9</td>
<td>0.01</td>
<td>1.219</td>
<td>1.029–1.443</td>
</tr>
<tr>
<td>Discharge mRS</td>
<td>3.86±1.4</td>
<td>4.6±0.9</td>
<td>0.02</td>
<td>1.567</td>
<td>1.054–2.331</td>
</tr>
</tbody>
</table>

NIHSS: National Institutes of Health Stroke Scale; mRS: Modified Rankin Score; pAF: Paroxysmal atrial fibrillation; OR: Odds ratio; CI: Confidence interval; ROC: Receiver operating characteristic; AUC: Area under the curve.
DISCUSSION

This study evaluated the predictive potential of a number of previously introduced serum biomarkers and risk factors for AF and pAF utilizing an online database from an independent clinical population. Specifically, we tested the readily available parameters in a stroke patient hospitalized within 24 hours of onset to predict subse-
quent paroxysmal atrial fibrillations. This is a significant clinical problem because a prominent number of stroke cases remain cryptogenic for etiology with no apparent causative factors despite rigorous evaluation. Detection of a single pAF episode in that setting can influence the lifelong treatment decisions as anticoagulation would be mandatory to prevent recurrent strokes. It may be necessary to acquire long-term continuous ECG monitoring to catch electrophysiologic evidences of pAFs, which is not always possible in routine clinical practice. Therefore, identification of high-risk patients based on clinical factors or biomarkers would be extremely helpful to focus advanced investigations on these cases rather than all cryptogenic strokes.

Our work showed the predictive potential of age, admission NIHSS and discharge MRS to predict pAFs (as detected by telemetric monitoring) in patients with ischemic stroke hospitalized within 24 hours of onset and with no AF on admission. The effect of age on pAF detection is augmented in females, possibly due to age-dependent sex-specific biological effects, like estrogen levels [38, 39]. Serum biochemical or hematological parameters, like magnesium, potassium and leukocyte counts or neutrophil/lymphocyte ratios had no significant association with pAF. Also, although the difference was not significant in our analysis, it would be helpful to consider hemoglobin level during clinical evaluation as relatively lower hemoglobin level is usually found in pAF groups being neither confounded by age or gender. This association can be better evaluated in larger samples with higher statistical power. Overall, our results suggest that especially in female patients over 75 years of age with an admission NIHSS over 13 would have a pAF episode within approximately 10 days after stroke onset, which would be missed if no ECG monitoring was performed. 83% of such cases had pAF in our patient population.

Atrial fibrillations are definitive causative factors for ischemic stroke [4], especially if detected before stroke [40], and it is highly recommended to include appropriate cardiac evaluation in the clinical work-up of stroke patients. However, one has to be aware of a possible overestimation of their role in recurrent strokes since a number of recent studies suggest that atrial fibrillation detected in the immediate post-stroke setting (either persistent or paroxysmal) could rather arise from a neurogenic autonomic dysregulation, questioning the cause-effect relationship of atrial fibrillations and stroke in such cases [40, 41]. Future studies need to address the problem of neurogenic versus cardiogenic AFs to validate their association with recurrent stroke events when not anticoagulated, and to identify biochemical or structural cardiac markers to distinguish those two groups clinically. In addition to a neurogenic autonomic dysregulation, electrolyte imbalances, especially abnormalities in potassium and magnesium [17, 18, 42], triggered after the stroke onset may also predispose the patients to arrhythmic events, including atrial fibrillations. As noted above, we did not find an association between such electrolyte abnormalities and the presence of AF in our sample.

Lack of association for potassium and magnesium levels or leukocyte parameters can arise for a number of reasons. First of all, this clinical database was prepared from the records of multiple institutions in the USA, and laboratory procedural standards might differ among hospitals, affecting the absolute values reported. Racial information was not systematically provided in those databases, but it can be a prominent confounder for blood measurements, and it is likely that data from different racial populations were pooled in this patient sample considering the residence of multiple ethnic groups in North America. Although the patients were monitored during their hospitalization for pAFs, it is possible that some very brief episodes were missed. Moreover, none of the patients were reported to have a Holter monitoring during inpatient phase or after discharge. Therefore, it is likely that a number of pAF cases were missed and still in the cryptogenic stroke group, which can result in this such discrepancies.

Despite the lack of an identified specific biomarker according to our study to predict pAF episodes, it has been helpful to confirm certain risk factors like sex, age, admission NIHSS and other contributing factors like hemoglobin level in an independent clinical sample. Ambitious long-term ECG monitoring in cryptogenic stroke cases can be beneficial especially such high-risk groups. This work serves as the first study on stroke patients of MIMIC database, at least to our knowledge, and also underlines the value of online clinical databases for evaluation of predictive role of certain factors for specific patient populations. There has been a progressive accumulation of suggested biomarkers for pAF and other stroke etiologies over the last several years and it would be helpful to test novel markers in such independent clinical databases to confirm their predictive potential to build a collective medical research environment.
Conclusion

There is a higher probability of pAFs after stroke, which may be the etiological cause, especially in older females admitted with higher NIHSS scores and discharged with higher disability. Additionally, low hemoglobin level can potentially be helpful for consideration of patients selected for advanced electrophysiologic evaluation. This study is a prime example of the utilization of public clinical databases for independent testing of various biomarkers identified in other clinical populations.

Acknowledgements: The author would like to thank Ellen Witkowski, Ph.D. (Boston University, USA) for English language editing of the manuscript.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study has received no financial support.

REFERENCES


