Predicting Postoperative Atrial Fibrillation from Independent ECG Components

Chih-Chun Chia, James Blum, Zahi Karam, Satinder Singh, Zeeshan Syed

University of Michigan, Ann Arbor, MI

Abstract

Postoperative atrial fibrillation (PAF) occurs in 10% to 65% of the patients undergoing cardiothoracic surgery. It is associated with increased post-surgical mortality and morbidity, and results in longer and more expensive hospital stays. Accurately stratifying patients for PAF allows for selective use of prophylactic therapies (e.g., amiodarone). Unfortunately, existing tools to stratify patients for PAF fail to provide clinically adequate discrimination. Our research addresses this situation through the development of novel electrocardiographic (ECG) markers to identify patients at risk of PAF. As a first step, we explore an eigen-decomposition approach that partitions ECG signals into atrial and ventricular components by exploiting knowledge of the underlying cardiac cycle. We then quantify electrical instability in the myocardium manifesting as probabilistic variations in atrial ECG morphology to assess the risk of PAF. When evaluated on 385 patients undergoing cardiac surgery, this approach of stratifying patients for PAF through an analysis of morphologic variability within decoupled atrial ECG demonstrated substantial promise and improved net reclassification by over 53% relative to the use of baseline clinical characteristics.

Introduction

Postoperative atrial fibrillation (PAF) occurs in 10% to 65% of the patients undergoing cardiothoracic surgery (Zaman et al. 2000; Ommen, Odell, and Stanton 1997; Maisel, Rawn, and Stevenson 2001; Asher et al. 1998). In PAF, the upper chambers (atria) of the heart *fibrillate* or contract rapidly and irregularly, preventing successful emptying of blood into the lower (ventricular) chambers. This causes blood to pool in the heart and clot, producing strokes and other morbidities that increase risk of postoperative mortality (Mathew et al. 2004). In addition to adversely affecting patients, PAF also imposes a substantial burden on the healthcare system by resulting in longer and more expensive hospital stays (Maisel, Rawn, and Stevenson 2001).

Prophylactic use of beta-adrenergic blockers and amiodarone has been shown to reduce the incidence of PAF (Mitchell et al. 2005; Gottlieb et al. 1999; Guarnieri

1999). However, while considering these treatments and other options such as anti-coagulation, the benefits of therapy must be balanced against adverse side effects[1]. This creates the need for clinical tools that can accurately stratify patients for PAF and guide fine-grained prophylactic administration of pharmacological therapy.

Existing Approaches to PAF Prediction. A number of different clinical metrics have been proposed to predict PAF. Older age has been shown to be consistently associated with a higher incidence of PAF (Zaman et al. 2000; Mathew et al. 1996; Hogue Jr, Hyder, and others 2000), most likely due to increased atrial fibrosis and dilation in older patients. Large observational studies have also found an association between other clinical characteristics and PAF, although the results of these studies have often been conflicting. Hypertension has been found to predict atrial fibrillation after cardiac surgery (Furberg et al. 1994), possibly due to fibrosis and dispersion of atrial refractoriness (Almassi et al. 1997; Aranki et al. 1996). Men also appear to be more likely than women to develop PAF after coronary artery bypass graft (CABG) (Zaman et al. 2000; Almassi et al. 1997; Aranki et al. 1996). It is believed that this effect may be due to differences in ion-channel expression and hormonal effects on autonomic tone. Previous atrial fibrillation and previous congestive heart failure have also shown an association with PAF (Mathew et al. 1996). In addition, procedural information such as aortic cross-clamp time and location of venuous cannulation have been found in some, but not all, studies to have predictive value for PAF (Maisel, Rawn, and Stevenson 2001). Postoperative factors such as respiratory compromise and prolonged ventilation have also been suggested (Aranki et al. 1996).

There is an extensive literature on electrocardiographic (ECG) markers to stratify patients for PAF. The majority of this work foucses on detecting abnormal P-wave duration on the surface ECG, as a way to identify intra-atrial conduction defects (Buxton and Josephson 1981). The use of this infor-

---

[1]For example, blinded therapy in the Prophylactic Amiodarone for the Prevention of Arrhythmias that Begin Early After Revascularization, Valve Replacement, or Repair (PAPABEAR) trial was more likely to be withdrawn in patients treated with amiodarone, largely because of a 3-fold increase in bradycardia requiring pacing and QTc interval prolongation greater than 650 milliseconds (Podgoreanu and Mathew 2005).

mation, however, is made challenging by the need for expert hand labeling of P-waves in the ECG. Moreover, features based on P-wave duration also suffer from inadequate precision and recall for clinical use. The small patient populations (64 to 240 patients) also make it difficult to generalize the findings from these studies to larger populations. Lack of data regarding specific treatment implications also limits the routine incorporation into clinical practice.

Predicting PAF from Atrial ECG. Our research focuses on developing and assessing novel ECG markers that can be clinically deployed in a fully-automated setting to identify patients at risk of PAF. To achieve this, we build off recent advances in predicting ventricular arrhythmias by quantifying probabilistic variations in the shape of the ECG signal over time (Syed et al. 2008; 2009a; 2009b; 2011). Investigations within multiple cohorts have shown that increased variability within the shape of the ECG is associated with myocardial instability. In this study, we adopt a similar approach but note that since the ECG is heavily dominated by ventricular activity, it may be more appropriate to decouple ECG signals into atrial and ventricular components that can be separately assessed for electrical instability. Specifically, we hypothesize that variability in the atrial ECG may reflect specific instability associated with PAF rather than other kinds of arrhythmias. To achieve this, we propose a novel eigen-decomposition algorithm for ECG time-series that leverages information about the underlying cardiac cycle to separate ECG signals into atrial and ventricular components. In a clinical trial on a representative cohort of patients undergoing cardiac surgery, we demonstrate how an analysis of morphologic variability in atrial components derived using such an approach can significantly improve stratification for PAF relative to existing automated approaches.

The contributions of this paper are as follow:

- We introduce the idea of stratifying patients for PAF by using information exclusively available in the atrial ECG;

- We propose a new approach to obtain an atrial ECG signal from the composite surface ECG through an eigen-decomposition algorithm informed by cardiac physiology;

- We describe how morphologic variability based on atrial ECG can be used as a novel marker of PAF;

- We rigorously evaluate the utility of this marker of PAF in a real-world cohort of patients undergoing cardiothoracic surgery; and

- We compare the relative merits of our eigen-decomposition algorithm leveraging cardiac physiology for source separation with an independent components analysis (ICA)-based approach.

Background

Electrocardiogram (ECG)

The ECG is a continuous recording of the electrical activity of the heart muscle or myocardium. At rest, each cardiac muscle cell maintains a voltage difference across its cell membrane. During depolarization (i.e., the 'firing' of the heart muscle), this voltage increases. Consequently, when depolarization is propagating through a cell, there exists a potential difference on the membrane between the part of the cell that has been depolarized and the part of the cell at resting potential. After the cell is completely depolarized, its membrane is uniformly charged again, but at a more positive voltage than initially. The reverse situation takes place during repolarization, which returns the cell to baseline. These changes in potential, summed over many cells, can be measured by electrodes placed on the surface of the body, leading to the ECG time-series.

The ECG is a quasi-periodic signal (i.e., corresponding to the quasi-periodic nature of cardiac activity). As shown in Figure 1(a) three major segments can be identified in a normal ECG. The P-wave is associated with depolarization of cardiac cells in the upper two chambers of the heart (i.e., the atria). The QRS complex (comprising the Q, R and S waves) is associated with depolarization of cardiac cells in the lower two chambers of the heart (i.e., the ventricles). The T-wave is associated with repolarization of the cardiac cells in the ventricles. The QRS complex is larger than the P-wave because the ventricles are much larger than the atria. The QRS complex also coincides with the repolarization of the atria, which is therefore usually not seen on the ECG. The T-wave has a larger width and smaller amplitude than the QRS complex because repolarization takes longer than depolarization (Lilly 2010). Figure 1 shows the relationship between the ECG signal (P, QRS, T-waves) and atrial/ventricular polarization. We emphasize that the P-wave corresponds to almost exclusively atrial activity, the T-wave corresponds to almost exclusively ventricular activity, while the QRS complex reflects both atrial and ventricular activity (see Figures 1(b) and (c) for detail); we will use these facts crucially in developing our approach to extracting the atrial ECG.

Morphological Variability (MV)

Recent work on stratifying patients for ventricular arrhythmias has shown that increased variation in the shape of the ECG waveform is a useful marker of myocardial instability (Syed et al. 2008; 2009a; 2009b; 2011). In this study, we build upon these results and focus on how a similar approach can be applied to atrial components of the ECG signal as a way of stratifying patients for PAF. We defer the question of how to derive a separation of the ECG into atrial and ventricular signals to the subsequent section. In what follows here, we briefly review the major principles associated with measuring MV.

For every pair of consecutively occurring beats in an ECG time-series, MV starts by quantifying how the shapes of the beats differ using a variant of dynamic time warping (DTW). This allows the original ECG signal to be transformed into a sequence of instantaneous morphology differences between consecutive pairs of beats. The spectral characteristics of this sequence are then studied, with energy between 0.30 and 0.55 Hz (as estimated using a Lomb-Scargle periodogram approach) being used as a marker of myocardial instability. In the remainder of this paper, we adopt an identical approach to measuring MV. A more detailed exposition of the

(a) ECG

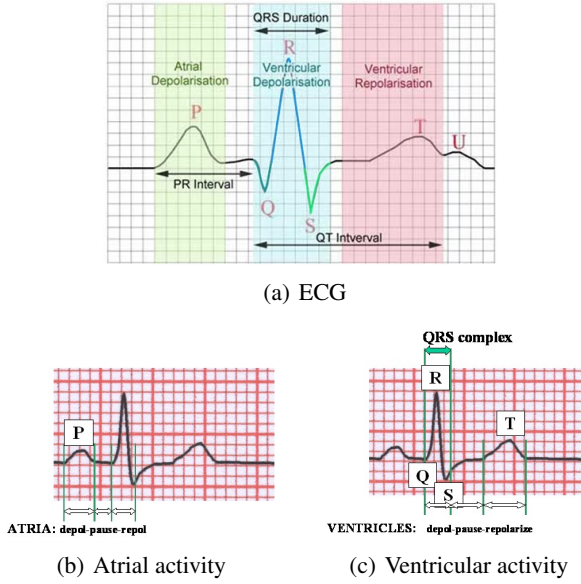(b) Atrial activity

(c) Ventricular activity

Figure 1: Atrial and ventricular components of the ECG (Grier 2008)

process of measuring MV can be found in (Syed et al. 2011).

ECG Decomposition

Consistent with the hypothesis proposed earlier, the focus of our work is to study morphological variability in atrial activation as a means of stratifying patients for PAF. Since observing atrial activity over the entire cardiac cycle is made difficult by the presence of ventricular activity, our approach requires first extracting the atrial components of the ECG waveform from the surface ECG. Traditional filtering-based methods are insufficient for this task since the ventricular activity is both higher amplitude, and occupies the same frequency ranges as atrial activity. Instead, we plan to formulate the separation of atrial and ventricular activity as a blind source separation problem, where the aim is to extract the atrial component of the ECG waveform from ventricular activity and noise.

The surface ECG measures electrical activity at different parts of the body and as shown by (Naït-Ali 2009) follows a linear instantaneous model, i.e., each recording of an ECG lead is a weighted linear combination of the atrial and ventricular components. Thus, the source separation problem that we are trying to solve can be viewed as an instance of the cocktail party problem. More formally, given n unknown signal sources $s_1(t), s_2(t), \cdots, s_n(t)$ and m observed signal mixtures $x_1(t), x_2(t), \cdots, x_m(t)$ the goal is to estimate $\mathbf{W}$ where $\mathbf{s}(t) = \mathbf{W} * \mathbf{x}(t)$ and $\mathbf{W}$ is the unknown m-by-n unmixing matrix with entries w_{ij} representing the contribution of observation $x_i(t)$ back onto source $s_j(t)$; note that the unmixing matrix is not a function of time. For our task of separating out the atrial component and the ventricular component, we need to learn the unmixing matrix that recovers the desired atrial and ventricular sources when applied to the ECG recordings (observations) across leads.

Independent Component Analysis (ICA): Most of the existing work on atrial component extraction has focused on the surface ECG extracted during atrial fibrillation episodes. The literature suggests that *during atrial fibrillation episodes*, atrial activity consists of small and continuous wavelets (a sawtooth form (Castells et al. 2003)) with a cycle of around 160ms. This has been modeled as a random variable with a distribution described by its histogram, i.e., a subgaussian signal. Based on this model, the atrial signal is said to have negative kurtosis (being subgaussian), while the ventricular signal has positive kurtosis (as it is assumed to be supergaussian). When such assumptions hold, i.e., during atrial fibrillation episodes, independent component analysis (ICA), which is capable of extracting independent non-Gaussian sources, has been shown to successfully extract atrial activity.

We note, however, that the goal of our work is to predict rather than detect atrial fibrillation. Therefore the assumptions upon which the ICA method are based do not apply to our research (since we intend to separate atrial and ventricular activity during normal sinus rhythm). Nevertheless, for completeness we consider the use of ICA for atrial component extraction on ECG in our experiments. Specifically, we make use of RobustICA, a variant of ICA based on using kurtosis as a contrast function. The component with the most positive kurtosis is considered to be the ventricular component, while the one with the most negative kurtosis is considered to be the atrial component.

Silence-energy-minimization (SEM): Our work differs from the standard cocktail party problem in that we have additional *a priori* knowledge of the time frames where *only one* of the speakers is speaking. In other words, based on cardiac physiology we know that for each heartbeat: the P-wave is associated exclusively with atrial depolarization while the T-wave relates only to ventricular repolarization. Thus there are periods within the ECG when only atrial (P-wave) or ventricular (T-wave) speaker's activity is present.

Let $s_A(t)$ and $s_V(t)$ denote the unknown A-beat (atrial) and V-beat (ventricular) source signals at time t. With m leads, let $\mathbf{x}(t)$ be the m-dimensional (vector) observed signal at time t. Assuming we can segment out the P- and T-waves in the observed signal we construct two new observed signals, $\mathbf{x_A}(t)$ and $\mathbf{x_V}(t)$ as follows: during P-wave activity set $\mathbf{x_A}(t) = \mathbf{x}(t)$ and set $\mathbf{x_V}(t) = 0$, during T-wave activity set $\mathbf{x_V}(t) = \mathbf{x(t)}$ and set $\mathbf{x_A}(t) = 0$, and everywhere else set $\mathbf{x_V}(t) = \mathbf{x_A}(t) = 0$. Collecting these new observed signals over k time steps into two $m \times k$ matrices X_A and X_V and collecting the unknown source signals into two $1 \times k$ vectors $\mathbf{s_A}$ and $\mathbf{s_V}$, we get the following two linear relationships: $\mathbf{s_A} = \mathbf{w_A^T} X_A$ and $\mathbf{s_V} = \mathbf{w_V^T} X_V$, where $\mathbf{w_A^T}$ and $\mathbf{w_V^T}$ are the $1 \times m$ unmixing vectors for the atrial and ventricular sources respectively. We solve for $\mathbf{w_A}$ and $\mathbf{w_V}$ using the following optimization function.

$$\max_{\mathbf{w_A}} ||\mathbf{w_A}^T X_A||^2 - c||\mathbf{w_A}^T X_V||^2 \quad \text{s.t.} \quad ||\mathbf{w_A}||^2 = 1$$

$$\max_{\mathbf{w_V}} ||\mathbf{w_V}^T X_V||^2 - c||\mathbf{w_V}^T X_A||^2 \quad \text{s.t.} \quad ||\mathbf{w_V}||^2 = 1$$

Here, we only derive the solution for the first optimization problem ($\mathbf{w_A}$) without loss of generality. Adding the

Lagrangian term, the problem becomes:

$$\max_{w_A} ||\mathbf{w_A}^T X_A||^2 - c||\mathbf{w_A}^T X_V||^2 - \lambda(||\mathbf{w_A}||^2 - 1).$$

Taking the derivative of the above with respect to $\mathbf{w_A}$ and setting it to zero, we get

$$X_A{}^T X_A \mathbf{w_A} - c \cdot X_V{}^T X_V \mathbf{w_A} - \lambda \mathbf{w_A} = 0$$
$$(X_A{}^T X_A - c \cdot X_V{}^T X_V) \mathbf{w_A} = \lambda \mathbf{w_A}.$$

Therefore, the resultant optimization problem is an eigen-problem. The solution can then be found by solving for the eigenvector of the following matrix $X_A{}^T X_A - c \cdot X_V{}^T X_V$, where c is a regularization term that controls the degree to which the unwanted ventricular part is attenuated. The derived unmixing vectors can be applied to all time frames, not just those corresponding to T- and P- waves, in the training and testing m-lead ECG signal to extract atrial and ventricular signals (at time t as $\mathbf{w_A^T}\mathbf{x(t)}$ and $\mathbf{w_V^T}\mathbf{x(t)}$ respectively).

We note that Weisman et al. proposed a method similar to ours to extract atrial electrical activity (Weissman, Katz, and Zigel 2009). However, their method makes an unrealistic assumption that the entire ECG signal can be cleanly segmented into segments of only pure atrial activity only and pure non-atrial activity. Moreover, their method tries to maximize the ratio of energy between atrial part vs non-atrial part, which requires an iterative algorithm when trying to solve the optimization function. Our approach, in contrast, has a closed form solution that is more applicable to real-time systems intended for continuous pre-operative and post-operative monitoring with prompt delivery of appropriate prophylaxis.

Extracting P and T-waves. The approach as described relies heavily on being able to identify the location of the P- and T-waves in the ECG. Many algorithms have been proposed in the literature to segment cardiac ECG beats into their corresponding P/Q/R/S/T-waves. However, these algorithms are generally unreliable at extracting the P/T-waves in real-world signals due to the relatively small magnitude of the P-wave and subtle changes marking the end of the T-wave. In addition, the real-world data employed in this paper are especially noisy, due to collection in an operating room (OR) setting, rendering these segmentation algorithms unsuitable for our purposes. As a result of this, we devised a heuristic based on physiology to establish the location of the P- and T-waves. Specifically, we attempted to relate the occurrence of these waves to the R-peak, which is the most prominent part of the beat (and therefore the easiest to detect). Our proposed heuristic is as follows: we make the general assumption that there is no ventricular activity (hence atrial only part) during $60 - 180$ms before an R-peak, while there is no atrial activity $80 - 480$ms after an R-peak (hence ventricular only part).

Experiments and Results

We evaluated our proposed methodology for PAF prediction on both synthetic and real-world data. We first study the ability of the two atrial extraction approaches described above, ICA and SEM, to reliably recover atrial and ventricular components on synthetically created ECG data. Next we study

the utility of atrial and ventricular separation in predicting PAF within a representative real-world clinical cohort. Details of the experiments and results are presented below.
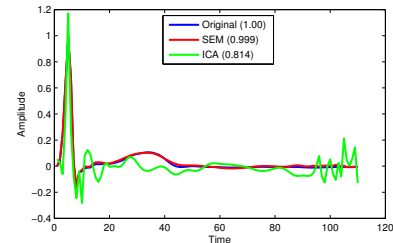
Synthetic Data
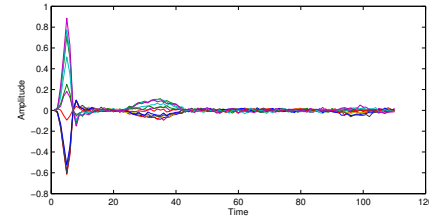
(a) Overlayed synthetic ECG data
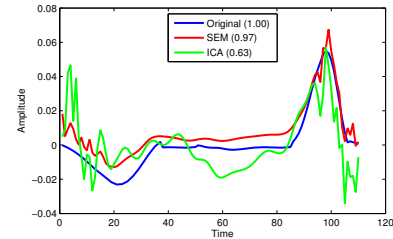
(b) Original vs. recovered atrial

(c) Original vs. recovered ventricular

Figure 2: Comparison of atrial and ventricular components extracted using ICA and SEM on synthetic data ($C = 10$ for SEM).

We created synthetic ECG beats by combining textbook templates of atrial activity (defined as the P-wave and T_A-wave) with ventricular activity (defined as the remaining waves). Specifically, we simulated multi-channel ECG data by using the linear instantaneous model proposed in (Naït-Ali 2009) to combine the atrial and ventricular components with randomly selected weights and additive white Gaussian noise. ICA and SEM were applied to this generated multi-channel ECG to obtain candidate atrial and ventricular components. These components were compared with the ground truth atrial and ventricular activity to assess the ability of ICA and SEM to reliably recover the original signals (using correlation as a performance criteria).

Figure 2 presents the results of this experiment. The synthetic multi-channel ECG data created using the approach above is illustrated in Figure 2(a). When separated into atrial and ventricular components (Figures 2(b) and (c)), the use of SEM for separation provided consistent improvements in the

recovery of both atrial and ventricular activity relative to the use of ICA. In particular, the use of prior knowledge about the relative absence of atrial and ventricular activity in SEM yielded a correlation coefficient of greater than 0.97. This was in contrast to the use of ICA, which failed to achieve any reasonable recovery of the atrial component and achieved marginal success dealing with ventricular activity (correlation coefficient of 0.81). Visually, the use of ICA also led to substantially more ripple in the extracted components than the use of SEM.
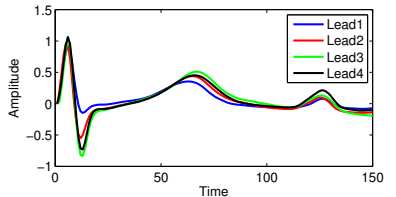
(a) Overlayed data from multiple ECG leads
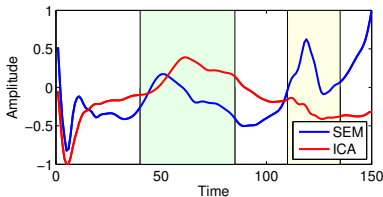
(b) Atrial components extracted by ICA and SEM
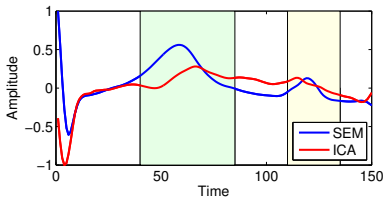
(c) Ventricular components extracted by ICA and SEM

Figure 3: Comparison of atrial and ventricular components extracted using ICA and SEM on actual ECG data ($C = 10$ for SEM). Shaded bands correspond to portions of the cardiac cycle corresponding to ventricular (green) and atrial (yellow) activity.

Real-World Data

Although we did not rigorously compare the ability of ICA and SEM to separate ECG into atrial and ventricular components on real patient data (owing largely to the absence of known ground truth in real data versus synthetic data), we note than in many cases the use of SEM provided qualitatively better results. For example, as shown in Figure 3, the use of SEM on 4-lead ECG data (Figure 3(a)) resulted in atrial components with substantially increased energy in the P-wave and PR-interval as opposed to ventricular components with substantially increased energy in the ST-segment and T-wave. This was in contrast to ICA, where the absence

of prior knowledge led to a comparatively poorer separation of the signal (Figures 3(b) and (c)).

We supplemented our analysis investigating the abilities of ICA and SEM to separate ECG into atrial and ventricular components with an evaluation of the clinical utility of such a separation on a real-world representative cohort of patients undergoing cardiac surgery. Data from 385 patients undergoing CABG, aortic, or other valvular surgeries at the University of Michigan Hospital were collected in 2013. The size of this cohort was considerably larger than previous studies investigating the use of ECG-based metrics to predict PAF largely because a focus on exploring a fully-automated approach to predict PAF (as opposed to one requiring substantial expert input) allowed us to evaluate our approach more rigorously in a larger cohort. For each patient, at least two sets of ECG waveforms were available; one recorded during surgery in the operating room (OR) and the other recorded during the intensive care unit (ICU) stay after surgery. All ECG waveforms were recorded at 240Hz, with 4-leads of recordings available (Lead 1, 2, 3, and a generic V-lead that we refer to as Lead 4). Expert review of the ECG data in the ICU following surgery was used to annotate the endpoint of PAF (90 events).

The goal of our investigation was to study the ability of markers based on MV, deriving from atrial and ventricular components of the ECG waveform in the OR, to predict PAF. We note that since recordings collected once surgery has started are typically too noisy for meaningful analysis, only the first 30 minutes of data in the OR preceding the operation were used. Moreover, since a key question while determining clinical utility is the extent to which any novel markers add information beyond existing variables, we also compared the use of MV measured from atrial and ventricular components of the ECG to baseline non-ECG clinical features available in the patient electronic health record (demographics, history and physical exam findings, laboratory reports, and type of surgery) and also based on the unseparated ECG signal (MV measured on each of Leads 1-4).

The models we evaluated include logistic regression models trained using stepwise backward elimination applied to: (Model 1) non-ECG features;
(Model 2) non-ECG features and features based on the complete ECG signal;
(Model 3) non-ECG features and features based on both the complete ECG signal and components derived using ICA;
(Model 4) non-ECG features and features based on both the complete ECG signal and components derived using SEM;
(Model 5) non-ECG features and features based on both the complete ECG signal and components derived using both ICA and SEM.
In all of these experiments, the stepwise backward elimination process removed one feature during each iteration based on cross-validated AUC results for each step. The reported AUCs are averaged across 50 trials that randomly divided data into 50% training and 50% test sets.

The evaluation metrics considered were: the discrimination (as assessed by the area under the ROC curve [AUC] and integrated discrimination improvement [IDI]) and reclassification between models (as assessed by the net re-

Method	AUC
Model 1	0.66
Model 2	0.69
Model 3	0.70
Model 4	0.70
Model 5	0.70

Table 1: AUC values for logistic regression models trained using stepwise backward elimination on different groups of features. See text for details of the different feature sets used for training the models above.

Comparison	IDI (P-value)	NRI (P-value)
Model 4 vs. Model 1	0.048 (<0.001)	53.4% (<0.001)
Model 4 vs. Model 2	0.017 (0.026)	25.6% (0.017)
Model 4 vs. Model 3	0.004 (0.275)	16.1% (0.091)

Table 2: IDI and NRI values comparing a logistic regression model trained using stepwise backward elimination on non-ECG features and features based on both the complete ECG signal and components derived through SEM (Model 4) to logistic regression models trained using stepwise backward elimination on different baseline feature sets (Models 1 to 3). See text for details of the different feature sets used for the models above. Note: IDI and NRI values are not presented for Model 4 vs. Model 5 since stepwise backward elimination resulted in the same features being retained in these models.

classification improvement [NRI]). NRI represents the proportion of patients appropriately assigned a higher or lower risk categorization under a new model relative to an old one, while IDI represents the difference of mean predicted probabilities of patients experiencing events and those remaining event free (Pencina, D'Agostino, and Vasan 2008).

Tables 1 and 2 present the results of these experiments. The results in Table 1 show that the inclusion of MV without source separation (Model 2) substantially improved performance relative to the use of baseline clinical features by themselves (Model 1). This performance was further improved with the addition of MV based on atrial and ventricular components derived through both ICA (Model 3) and SEM (Model 4). The improvement was marginally larger when SEM was used for separation than when ICA was used. Specifically, we note that when MV markers based on both ICA- and SEM-separated ECG components were included together (Model 5), the backward stepwise elimination process retained MV based on atrial activity derived using SEM in preference to MV based on all ICA derived components.

The IDI and NRI metrics in Table 2 using SEM (Model 4) were also positive relative to the use of the baseline clinical features by themselves (Model 1), the additional use of MV without source separation (Model 2), and the further inclusion of MV using ICA (Model 3). The smallest of these improvements corresponded to a positive net reclassification of over 16% with statistical significance at the 10% level.

Conclusion

We focused on the question of developing novel markers that can be used to stratify patients undergoing cardiac surgery for PAF. Given the substantial burden that PAF imposes post-operatively, the ability to identify patients most likely to experience PAF can substantially improve mortality and morbidity, and also reduce healthcare costs, by creating the opportunity to deliver prophylaxis in a timely and personalized manner. The challenge to realizing this, however, is that there are currently no established metrics for PAF risk stratification. To address this need, we explored the development of ECG-based markers in our work that can be deployed in an inexpensive, non-invasive, and fully-automated manner to evaluate patients undergoing cardiac surgery. We focused, in particular, on extending advances in stratifying patients for ventricular arrhythmias (that quantify excessive variability in the ECG waveform) to similarly evaluating the health of the electrical activity of the atria. Central to this is the ability to distinguish lower amplitude atrial activity from higher amplitude ventricular activity. To decompose the ECG into separate components corresponding to both atrial and ventricular activity, we proposed a novel eigen-decomposition approach based on silence energy minimization, which partitions ECG time-series into atrial and ventricular components by exploiting knowledge of the underlying cardiac cycle. Using this, we measured atrial instability by studying probabilistic variations in atrial ECG morphology as a means of determining risk for PAF.

We evaluated our work on both synthetic and real-world data. Although our cohort size is not large in absolute terms; it is larger than previously conducted studies for predicting PAF (our ongoing data collection will ultimately yield over a 1000 patients allowing more comprehensive evaluation and sharing with the clinical community in future work).

Our results on synthetic data showed that the use of additional knowledge based on physiology to distinguish between atrial and ventricular activity during the ECG decoupling process substantially improved performance relative to physiology-agnostic approaches such as ICA. When further evaluated on data from a well-characterized cohort of patients undergoing cardiac surgery, we further observed that the use of physiology to guide ECG separation into atrial and ventricular components achieved better results than the use of a purely statistical approach such as ICA. Moreover, the development of markers based on an analysis of atrial ECG significantly improved models based on baseline clinical features and an assessment of variability within the entire (unseparated) ECG. In particular, our results show that relative to the combination of baseline clinical features and ECG features without separation, our proposed approach can improve classification by over 25% with statistically significant improvements in discrimination.

Knowing which patients will or will not develop atrial fibrillation post-operatively provides the opportunity to deliver prophylaxis (e.g., amiodarone) in a selective manner. Also, morphologic variability of atrial ECG improves our understanding of the pathophysiology of PAF and may lead to better therapies. These results thus have the potential to improve the care of tens of thousands of patients each year.

Acknowledgment

References

Almassi, G.; Schowalter, T.; Nicolosi, A.; Aggarwal, A.; Moritz, T.; Henderson, W.; Tarazi, R.; Shroyer, A.; Sethi, G.; Grover, F.; et al. 1997. Atrial fibrillation after cardiac surgery: a major morbid event? *Annals of surgery* 226(4):501.

Aranki, S.; Shaw, D.; Adams, D.; Rizzo, R.; Couper, G.; VanderVliet, M.; Collins, J.; Cohn, L.; and Burstin, H. 1996. Predictors of atrial fibrillation after coronary artery surgery: current trends and impact on hospital resources. *Circulation* 94(3):390–397.

Asher, C.; Miller, D.; Grimm, R.; Cosgrove 3rd, D.; Chung, M.; et al. 1998. Analysis of risk factors for development of atrial fibrillation early after cardiac valvular surgery. *The American journal of cardiology* 82(7):892.

Buxton, A., and Josephson, M. 1981. The role of p wave duration as a predictor of postoperative atrial arrhythmias. *CHEST Journal* 80(1):68–73.

Castells, F.; Igual, J.; Rieta, J.; Sanchez, C.; and Millet, J. 2003. Atrial fibrillation analysis based on ica including statistical and temporal source information. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on*, volume 5, V–93. IEEE.

Furberg, C.; Psaty, B.; Manolio, T.; Gardin, J.; Smith, V.; and Rautaharju, P. 1994. Prevalence of atrial fibrillation in elderly subjects (the cardiovascular health study). *The American journal of cardiology* 74(3):236–241.

Gottlieb, S.; Dudek, A.; Lowry, D.; Nolan, S.; and Guarnieri, T. 1999. Intravenous amiodarone for the prevention of atrial fibrillation after open heart surgery: the amiodarone reduction in coronary heart (arch) trial. *Journal of the American College of Cardiology* 34(2):343–347.

Grier, J. M. 2008. Eheart: Introduction to ecg ekg. http://www.ndsu.edu/pubweb/ grier/eheart.html.

Guarnieri, T. 1999. Intravenous antiarrhythmic regimens with focus on amiodarone for prophylaxis of atrial fibrillation after open heart surgery. *The American journal of cardiology* 84(9):152–155.

Hogue Jr, C.; Hyder, M.; et al. 2000. Atrial fibrillation after cardiac operation: risks, mechanisms, and treatment. *The Annals of thoracic surgery* 69(1):300.

Lilly, L. 2010. *Pathophysiology of Heart Disease:: A Collaborative Project of Medical Students and Faculty*. Lippincott Williams & Wilkins.

Maisel, W.; Rawn, J.; and Stevenson, W. 2001. Atrial fibrillation after cardiac surgery. *Transplantation* 151136:11.

Mathew, J.; Parks, R.; Savino, J.; Friedman, A.; Koch, C.; Mangano, D.; and Browner, W. 1996. Atrial fibrillation following coronary artery bypass graft surgery. *JAMA: the journal of the American Medical Association* 276(4):300–306.

Mathew, J.; Fontes, M.; Tudor, I.; Ramsay, J.; Duke, P.; Mazer, C.; Barash, P.; Hsu, P.; Mangano, D.; et al. 2004. A multicenter risk index for atrial fibrillation after cardiac surgery. *JAMA: the journal of the American Medical Association* 291(14):1720–1729.

Mitchell, L.; Exner, D.; Wyse, D.; Connolly, C.; Prystai, G.; Bayes, A.; Kidd, W.; Kieser, T.; Burgess, J.; Ferland, A.; et al. 2005. Prophylactic oral amiodarone for the prevention of arrhythmias that begin early after revascularization, valve replacement, or repair. *JAMA: the journal of the American Medical Association* 294(24):3093–3100.

Naït-Ali, A. 2009. *Advanced biosignal processing*. Springer.

Ommen, S.; Odell, J.; and Stanton, M. 1997. Atrial arrhythmias after cardiothoracic surgery. *New England Journal of Medicine* 336(20):1429–1434.

Pencina, M. J.; D'Agostino, R. B.; and Vasan, R. S. 2008. Evaluating the added predictive ability of a new marker: from area under the roc curve to reclassification and beyond. *Statistics in medicine* 27(2):157–172.

Podgoreanu, M., and Mathew, J. 2005. Prophylaxis against postoperative atrial fibrillation. *JAMA: the journal of the American Medical Association* 294(24):3140–3142.

Syed, Z.; Scirica, B. M.; Stultz, C. M.; and Guttag, J. V. 2008. Risk-stratification following acute coronary syndromes using a novel electrocardiographic technique to measure variability in morphology. In *Computers in Cardiology, 2008*, 13–16. IEEE.

Syed, Z.; Scirica, B. M.; Mohanavelu, S.; Sung, P.; Michelson, E. L.; Cannon, C. P.; Stone, P. H.; Stultz, C. M.; and Guttag, J. V. 2009a. Relation of death within 90 days of non-st-elevation acute coronary syndromes to variability in electrocardiographic morphology. *The American journal of cardiology* 103(3):307–311.

Syed, Z.; Sung, P.; Scirica, B. M.; Morrow, D. A.; Stultz, C. M.; and Guttag, J. V. 2009b. Spectral energy of ecg morphologic differences to predict death. *Cardiovascular Engineering* 9(1):18–26.

Syed, Z.; Stultz, C. M.; Scirica, B. M.; and Guttag, J. V. 2011. Computationally generated cardiac biomarkers for risk stratification after acute coronary syndrome. *Science translational medicine* 3(102):102ra95.

Weissman, N.; Katz, A.; and Zigel, Y. 2009. A new method for atrial electrical activity analysis from surface ecg signals using an energy ratio measure. In *Computers in Cardiology, 2009*, 573–576. IEEE.

Zaman, A.; Archbold, R.; Helft, G.; Paul, E.; Curzen, N.; and Mills, P. 2000. Atrial fibrillation after coronary artery bypass surgery: a model for preoperative risk stratification. *Circulation* 101(12):1403–1408.