Heart Rate Topic Models

Alexander Van Esbroeck, Chih-Chun Chia, Zeeshan Syed

University of Michigan

Ann Arbor, MI 48109

Abstract

A key challenge in reducing the burden of cardiovascular disease is matching patients to treatments that are most appropriate for them. Different cardiac assessment tools have been developed to address this goal. Recent research has focused on heart rate motifs, i.e., short-term heart rate sequences that are over- or under-represented in long-term electrocardiogram (ECG) recordings of patients experiencing cardiovascular outcomes, which provide novel and valuable information for risk stratification. However, this approach can leverage only a small number of motifs for prediction and results in difficult to interpret models. We address these limitations by identifying latent structure in the large numbers of motifs found in long-term ECG recordings. In particular, we explore the application of topic models to heart rate time series to identify functional sets of heart rate sequences and to concisely describe patients using task-independent features for various cardiovascular outcomes. We evaluate the approach on a large collection of real-world ECG data, and investigate the performance of topic mixture features for the prediction of cardiovascular mortality. The topics provided an interpretable representation of the recordings and maintained valuable information for clinical assessment when compared with motif frequencies, even after accounting for commonly used clinical risk scores.

Introduction

Cardiovascular disease affects 7% of adults in the United States and is the single leading cause of mortality (Roger et al. 2011). It is estimated that in 2011, there were over 1 million coronary attacks in the U.S., with nearly one-third of individuals suffering a coronary attack dying of it. A similar situation exists in other parts of the world, in particular developing countries, where it is estimated that by the year 2020 cardiovascular disease will be responsible for 40% of all deaths (Reddy 2004). The increasing prevalence of cardiovascular disease, combined with its high risk of mortality, has resulted in a variety of treatments (e.g., procedures, medications, devices) being developed to prevent future events and reduce an individual's risk of mortality.

Despite the wide array of these treatment options, determining the appropriate level of therapy for an individual remains challenging. Underestimating a patient's risk for cardiovascular death (CVD) can preclude the use of potentially life-saving preventative treatment. Conversely, overestimation of risk can lead to application of expensive and often risky procedures which provide little or no benefit to the patient. One example of this is the implantable cardiac defibrillator (ICD), an expensive device requiring surgical implantation capable of preventing fatal arrhythmias. The outstanding majority of the patients receiving an ICD under current guidelines do not benefit from it, while most patients who die of arrhythmias are not prescribed one (Bardy et al. 2005). Accurate risk stratification is vital for matching patients to appropriate treatments, with the potential to both save lives and reduce health care costs.

A number of different biomarkers and clinical scores have been generated to quantify a patient's cardiac risk and to guide clinical decision making. These include biochemical markers such as brain natriuretic protein and C-reactive protein; measures derived from echo- and electrocardiography; and clinical risk scores such as the Thrombolysis in Myocardial Infarction (TIMI) risk score, which integrate information related to different patient characteristics. While these metrics can identify many high-risk patients, they utilize instantaneous data and lose potentially valuable information about the performance of the heart over time. Recent work on cardiovascular risk stratification has focused on addressing this discrepancy and improving patient assessment through novel cardiac biomarkers derived from long-term physiological signals such as the ECG. Of particular interest is studying changes in the heart rate over long periods of time, as a way of characterizing abnormal autonomic nervous regulation of the heart predisposing to fatal arrhythmias. Much of the initial work in this space (e.g., heart rate variability (Malik 1996), heart rate turbulence (Barthel et al. 2003), and deceleration capacity (Bauer et al. 2006)) relates either aggregate changes in heart rate or the presence of a specific heart rate pattern to an increase in cardiovascular risk. Most recently, this work has been generalized to identify and integrate information in multiple heart rate motifs (i.e., short heart rate sequences discovered in a data-driven manner from long-term ECG) that are over- or underrepresented in patients experiencing cardiovascular outcomes

(Chia and Syed 2011).

While heart rate motifs have been found to be predictive of several cardiac outcomes (Syed et al. 2011), the approach presents several challenges. The number of motifs present in the data is overwhelming, and relatively small datasets necessitate feature selection to prevent overfitting. By selecting a sparse set of useful motifs, the current approach risks losing valuable information in the many remaining motifs and their interactions. Even if severely reducing the size of the feature set was unnecessary, a representation consisting of tens of thousands of features creates challenges for both model interpretation and identification of trends in the data. Additionally, reducing the size of the representation is typically done with respect to a particular clinical condition (e.g., predicting cardiovascular mortality). There are a variety of cardiac endpoints for which prediction is valuable, and different outcomes (e.g. ventricular arrhythmias, atrial arrhythmias, myocardial infarction, recurrent ischemia) require selection of distinct sets of predictive motifs to achieve optimal performance. This results in multiple representations that are challenging to relate across cardiac events.

In this paper, we hypothesize that there is information useful for prediction to be gained by considering all of the patterns in the data at once. Leveraging larger sets of motifs, as well as the interactions and relationships between different patterns, may provide additional information about the functioning of the heart relevant to the prediction of cardiac outcomes. We further hypothesize that the higher-level organization of motifs across patients, for example in related sets of motifs and their trends across a population, can allow for greater interpretability of heart rate motifs. Our study builds on these hypotheses and addresses the shortcomings of the heart rate motifs approach by identifying generative structure in the full set of motifs throughout the population. We learn topics consisting of related heart rate motifs capable of characterizing longitudinal ECG recordings across a population. These data-derived topics achieve an interpretable abstraction of the heart rate motifs present in a recording, while providing a concise, task-independent representation which can be used to better assess overall cardiac health.

The contributions of this paper are:

- we present a novel approach for accurate and interpretable cardiovascular risk stratification that identifies higher-level structure in short-term heart rate structure across a population;

- we propose topic models as a possible implementation of this approach, and describe how an existing method for identification of latent structure in text documents can be applied to physiological time series to assess cardiovascular risk; and

- we evaluate the performance of the topics alongside a popular risk measure on a real-world dataset consisting of long-term ECG recordings with nearly year-long patient follow-ups.

The paper is organized as follows: first, we describe the ECG, cardiac risk stratification, and motivate the application of topic models to heart rate motifs. We then detail the preprocessing, symbolization, and model estimation components of the proposed approach. We investigate the generated topics and specify a set of experiments to evaluate their utility for prediction, and finally present and discuss the results.

Background

The ECG measures the electrical potential of the heart on the surface of the chest. This simple measure of electrical activity contains a vast amount of information related to the timing of cardiac activity as well as the structure and functioning of the components of the heart and its related systems. This information, combined with the ease with which ECG can be recorded and its prevalence in health care settings, has made it a common data source for development of cardiac risk stratification metrics. With the recent development of better storage technology and practices, the ability to apply more sophisticated analysis to long-term ECG recordings has become possible.

Most cardiac biomarkers derived from the ECG focus on the idea that an inability for the heart to adjust its rate to compensate for different physiological situations is indicative of a high risk for cardiovascular problems. In recent work on heart rate motifs, Chia and Syed explored identifying and integrating information from the frequencies with which different short-term heart rate patterns occur in long-term ECG to assess cardiac patients. The approach converted 24 hour heart rate time series measured from long-term ECG data into symbolic sequences, corresponding to low through high heart rate ranges, and discovered short approximate symbolic motifs associated with high or low risk individuals (Chia and Syed 2011). This approach identified a small set of over- or under-represented symbolic motifs in patients suffering cardiovascular death, and integrated the frequencies of these motifs over an ECG recording into a predictive model. This approach found that by studying the short-term heart rate patterns over long-term data it was possible to identify information complementary to other commonly used clinical variables.

While the heart rate motifs method improves upon existing methods for risk stratification, there are several challenges with the use of heart rate motifs. For even a moderate choice of motif lengths, the number of motifs in the data grows unreasonably (with 4 symbols and motifs of length 8 there are over 60,000 motifs). The use of real-world long-term ECG datasets, which contain limited numbers of patients, necessitates the use of only a sparse subset of features to prevent overfitting. This has two effects: first, this prevents the leveraging of large numbers of patterns and their relationships for prediction. Second, this model describes only a small percentage of the original recordings, and in some cases an individual may have none of the predictive motifs present in their ECG. This provides a limited representation of the data, and can be challenging to interpret. Another difficulty is that while the resulting model has good predictive power, it provides limited insight into the generative structure of these motifs across a recording or a population. The use of individual motif frequencies identifies

no structure in the relationships between motifs, and cannot distinguish cases where one heart rate motif can appear in multiple contexts. Understanding the higher-level structure of these motifs may provide novel insights into the data and correspondingly into cardiac physiology. Additionally, because many cardiac outcomes exist, the use of motifs requires selection of different sets of predictors for each one. This makes it challenging to compare predictors across types of cardiac events, because there is no easy way to reconcile the two models.

By extracting underlying generative structure independent of any particular endpoint, it may be possible to retain useful information present in the full set of motifs relevant to overall cardiac frailty, without tailoring the representation to a particular task. We investigate one approach to uncovering latent structure, topic modeling, to heart rate motifs. Topic models have been used in a variety of application domains to identify sets of semantically related words capable of describing a set of texts. In this paper, we extend the use of topic models to ECG data. In particular, we investigate advancing the work of Chia and Syed by first symbolizing and identifying "words" from a physiological time series, and then identifying latent topics in the data. Instead of independently comparing each motif frequency between the ECG recordings for two outcomes, we consider the motifs as words, constituents of a single document corresponding to a recording. After defining a document as the frequencies of all motifs occurring in a recording, we can use this text-like representation to extract higher-level structure in the documents by learning topics over the motifs.

The application of topic models to heart rate motifs has the ability to address the challenges mentioned. By modeling the heart rate motifs in a recording with a set of generative topics, every motif in the data is explained by a contribution from some topic, accounting for the entirety of every patient's recording. By relating symbolically different but functionally similar motifs, all patterns in the data can be leveraged to identify useful predictive information. This allows the previously unwieldy motif frequency representation to be condensed into a mixture over a small set of topics, providing a similarly sparse representation but with the potential for greater predictive power.

This representation of the data as a mixture of topics also serves to identify generative structure behind the motifs. It allows the analysis of related sets of patterns, and easier identification of trends across recordings or patients by assessing the variability among a few topics rather than a large number of motif frequencies. The unsupervised nature of topic models means that this condensed representation provides information about the full set of motifs without focusing on a particular outcome. After identifying latent heart rate topics, the patient mixtures over topics can be used across a variety of cardiac events, allowing for a quantification of the relationships between their differing predictors. For these reasons, we believe that topic models provides a better interpretation and understanding of heart rate motifs, while maintaining and potentially expanding upon the predictive power of these features for cardiac assessment.

Other work has investigated the discovery of states in physiological time series. Cohen et al. modeled physiological states in multivariate intensive care unit (ICU) time series using a hierarchical clustering approach (2010). Saria et al. used a hierachical graphical model to simultaneously model multivariate time series and learn a set of disease "topics" in neonatal intensive care unit time series (2010). In contrast, the proposed approach leverages time series motifs in a "bag of motifs" representation analogous to the "bag of words" representation in text analysis.

Methods

The proposed method consists of four stages. Initially, a heart rate time series is generated from the ECG signal. Then, this time series is converted into a symbolic sequence. From this sequence we extract motifs, each of which describes a heart rate pattern. Finally, we use the motif frequencies for each patient across a population to train a topic model than can be used for cardiovascular risk stratification.

We start this process by extracting the heart rate time series from the raw ECG recording. This involves identifying QRS complexes and then calculating time between adjacent complexes. This is done using the open source QRS detection algorithms proposed by Hamilton et al. (1986) and Zong et al. (2003) to identify QRS complexes at time instants where both algorithms agree. The instantaneous heart rate was defined as the time between all pairs of normal sinus beats.

We then convert the instantaneous heart rate values to symbolic sequences, to normalize the data and to transform it into a text-like representation appropriate for analysis with topic models. We use symbolic aggregate approximation (SAX) to achieve this representation (Lin et al. 2003). Given an alphabet size A, SAX first divides the input values into A equiprobable bins. SAX then assigns each value from the original series to a symbol based on the bin in which it sits. In the hypothetical case of $A = 3$, the symbols may be interpreted as low, medium, and high heart rates. We apply SAX to each patient's time series separately. This provides robustness to inter-patient variability in baseline heart rates in that a symbol corresponding to high heart rate has a consistent meaning across individuals, an essential property when learning structure across a population.

Each patient's symbolic sequence can be abstractly represented as a text document. We consider all substrings of a given length in this symbolic sequence as words within the document representation, characterizing long-term ECG as a bag of motifs and their respective frequencies of occurrence. For a motif length n, we calculate the frequency of all n-length substrings in the symbolic sequence, allowing for overlaps. In earlier work on heart rate motifs, a formulation of approximate motifs was used, in order to group functionally equivalent heart rate sequences as well as to account for the hard symbolization boundaries. This definition of motifs required an extensive framework to efficiently compute motif frequencies on large datasets. The application of topic models, which provide a natural and robust way to identify groups of related motifs, allows us to sidestep the challenges involved in computing approximate motif frequencies. Given a corpus of patient ECG motif frequencies,

we can then train a topic model that identifies latent structure in the occurrence of these large numbers of motifs throughout the course of a day. As a result, in our work, we do not consider approximate motifs and use a simple linear time algorithm based on hashing to identify motifs and their frequencies in symbolic sequences.

To train a topic model on the abstract textual representation of long-term ECG, we applied Latent Dirichlet Allocation (LDA) (Blei, Ng, and Jordan 2003), a commonly used topic model, to the collection of heart rate motif frequencies. LDA is a generative model that assumes that a collection of documents was derived from a set of topics (e.g., the proceedings of an artificial intelligence conference may consist of topics such as planning, knowledge representation, or machine learning). Each topic underlying the corpus can be characterized by a distribution over words in the vocabulary: a topic about machine learning may have high probabilities of words like "unsupervised" or "classfier". Given only a set of documents and their respective word frequencies, the LDA generative model can be used to learn a set of topics to explain the corpus, and attribute each document to some mixture of these topics.

More concretely, LDA defines an underlying set of K topics, where each topic k can be defined by a distribution over all of the words in the vocabulary. Each document d is itself generated by a distribution over topics, where the generative process assumes a two part process for each word w_{di} in the document. First, a topic z_{di} is chosen by sampling a topic from that document's topic distribution θ_d. Then the word w_{di} is sampled from that topic's distribution over words β_k. More concretely, the model defines for a document d the prior distributions over observed words w_{di}, their latent topics z_{di}, and the document distributions over topics θ_d as:

$$\theta_d \sim Dir(\alpha)$$
$$z_{di} \sim Multi(\theta_d)$$
$$w_{di} \sim Multi(\beta_{z_{di}})$$

Where α parameterizes a symmetric Dirichlet prior on topic distributions. In training the model, the parameters of interest are the θ_d and β_k values, which characterize the topics of semantically related words and the proportions of these topics with respect to each document.

In our work, we consider one patient's ECG recording as analogous to a document, with each heart rate motif corresponding to a word. It is then possible to treat the set of symbolized heart rate time series as a corpus, over which we can identify topics of heart rate motifs. The LDA parameters were estimated through variational Bayesian inference as described in Blei et al. (Blei, Ng, and Jordan 2003). Rather than identifying topics representing semantically related words, the derived topics contain sets of functionally related motifs. We used open source software for variational inference-based estimation of the model parameters from the symbolized heart rate corpus, resulting in a set of heart rate topics and each patient's topic distribution.

Experiments

We evaluated the potential clinical utility of heart rate topic models on data from 4,557 patients admitted to hospitals with non-ST-elevation acute coronary syndrome. For each patient, the first 24 hours of continuous ECG were used in our experiments.

The unsupervised topic modeling process provides a higher-level representation of the data, i.e., in terms of a distribution of topics rather than a distribution over motifs. In addition to studying the improvement in visualization and interpretability provided by this unsupervised approach, we also quantified the ability of a topic model representation to preserve clinically useful information by performing the following supervised learning experiment. We divided the patient population at random into evenly sized separate training and testing cohorts. We developed two sets of supervised models on the training set that transformed the topic model representations of patients into a quantitative predictions for CVD. The first of these models used l_2-regularized logistic regression with the parameters of each document's distribution over topics as features. The predictive ability of the model was measured using the area under the ROC curve (AUC). To identify whether the improvement in AUC between two choices of model features was significant, we further evaluated the integrated discriminative improvement (IDI) of using one model versus another (Pencina et al. 2008). The second of the supervised models used in our study was a Cox proportional hazards regression model, which incorporates the time until occurrence of an event into the analysis, allowing for a more robust analysis of patient outcomes. Proportional hazards models assume a baseline risk function that quantifies the risk of an event at a given time. The hazard ratio statistic derived from these models represents the multiplicative effect of a predictive variable on baseline risk levels. A hazard ratio of 2 can be roughly interpreted as a two-fold relatively increased risk of an event per unit time, while a ratio of 1 indicates no effect on risk. To achieve a dichotomized predictor (a binary value corresponding to low or high risk) for use with the proportional hazards model, the predictions from the logistic regression model were dichotomized at the top quintile consistent with the cardiovascular literature.

The topic model estimation was performed on the full set of patient recordings. We experimented with varying values of motif length (n=2,4,6,8), alphabet size (A=2,4,6,8), and number of topics (K=1,...,10), accounting for variability in model estimation by repeating the LDA parameter estimation process 10 times. The choice of n and A was done using cross-validation, while selection of K and choice of model over parameter estimation replicates was done using the Bayesian information criterion (BIC). This corresponded to a motif length of 6, an alphabet size of 4, and 10 topics, which was used for the full set of experiments.

We also compared the performance of our heart rate topic models approach to an approach based on the heart rate motifs work of Chia and Syed. The goal of comparing to individual motifs is to assess whether the additional structure derived by the unsupervised approach provides comparable or additional information to the raw motif frequencies. To provide a fair comparison given the large number of features for the classification task, a search over regularization coefficients was conducted, using the same approach imple-

mented for motif length and alphabet size selection with the topic mixture features. A motif length of 6 and an alphabet size of 4 were used for the motifs to provide a direct comparison with the topic models.

Finally, to study the extent to which heart rate topic models provided information that was complementary to existing clinical measures, we also evaluated the performance of topic models with the inclusion of the TIMI risk score (TRS) (Antman et al. 2000). The TRS is commonly used to evaluate a patient's risk of acute coronary syndrome, and incorporates an array of clinical risk factors (e.g., age, weight, blood pressure), serum cardiac biomarker levels, number of anginal events in the previous 24 hours, and several other factors. Patients are assigned a score between 0 and 7, which can be categorized into low (0-2), moderate (3-4), and high (5-7) risk. TRS provides an easily calculated score that a clinician can use to assess a patient's cardiac risk. Due to the ease of use and prevalance of the metric, the TRS can be considered a standard for risk stratification. In order for a new biomarker to demonstrate potential for genuine clinical impact in cardiovascular health, it must provide predictive information that goes beyond what TRS already captures. Evaluation with TRS indicates whether the topic model contains useful information that is not captured by standard clinical measures.

Results

Topic Analysis

The topic distributions for individual patients varied significantly across the population, as shown in Figure 1. This indicates that despite a number of similarities in the symbolized heart rate time series between patients (e.g., many of the patients had long runs of the same symbols, such as "1111"), the model still managed to identify topics that vary significantly across the patient population. As Figure 1 shows, the patient mixtures over topics characterize their day-long recordings in a way that is easy to visualize and compare across individuals and subgroups. Importantly, the ability to reduce each patient into a small number of topics that can be visualized provides a representation in terms of higher-level inferred state that is more compact and easier to understand (i.e., in terms of what topics differentiate patients) than simply displaying the frequencies of a very large number of motifs simultaneously.

In a traditional scenario, understanding the meaning of a topic is done by investigating its most probable words. For example, a topic about cardiology may have words like "heart", "ECG", or "myocardial" as some of its most likely words. To gain some understanding of the higher-level structure the model captures, we investigated the most probable words in several of the topics. Figure 2 depicts the 10 most likely words in four topics, two corresponding to high risk of CVD (left) and two corresponding to low risk (right) under a Cox proportional hazards model. Topic 1 includes a number of words consisting of a high heart rate (denoted by 4s) interrupted by a single low rate interval. Topic 2 consists of oscillations between heart rates, particularly sharp changes between low and medium-high to high rates. In
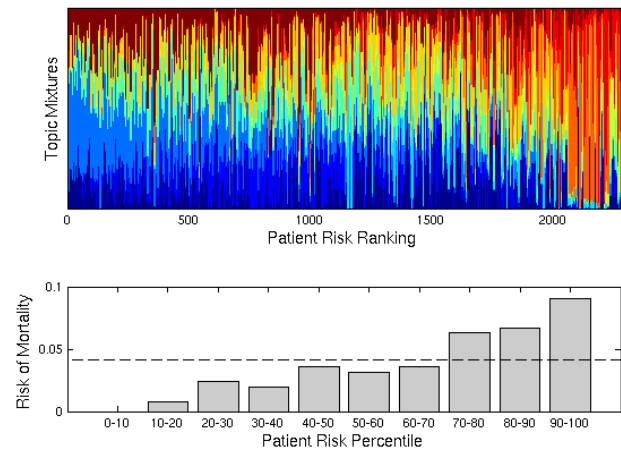

Figure 1: Topic mixtures for the testing set of patients. The top diagram shows each patient's mixture of topics, with each patient represented by a vertical strip, ordered by their decision values with estimated higher risk patients on the left and lower risk patients on the right. The bottom diagram shows risk of mortality calculated over 10 bins corresponding to 10 percentile ranges of the decision values, with the overall population risk (dotted line).

contrast, the low-risk topics show more stability, with deviations from the baseline having a low amplitude. Long runs of the same symbols occur commonly in several of the topics, showing the ability of the model to differentiate between instances of the same motif depending on their surrounding patterns. Rather than identifying only whether or not a motif is used, the topics model approach uses context to select the most appropriate topic for a given motif instance.

Logistic Regression

The model trained using topic mixture features provided a slightly higher AUC compared to those trained using TRS or motif frequencies. Training a model with both topic features and TRS improves performance substantially, as does inclusion of motif frequencies with TRS. The increase in AUC was higher when combining the TRS with topics than with motifs, even after selecting an appropriate regularization parameter. The IDI values indicated that both topics and motifs gave a statistically significant improvement (p ¡ 0.05) when combined with TRS over using TRS on its own. The comparison between the TRS/Topics and TRS/Motifs models also showed an improvement when using topic models, although in this case the improvement was marginally significant (p = 0.07).

Cox Proportional Hazards

When considering the timing of events and the censoring of patients, all three measures showed a significant association with time to death, with identified high-risk patients showing a more than two-fold increased risk of CVD. After accounting for the TRS, both motifs and topics remained significant, with topic models achieving a marginally higher hazard ratio.
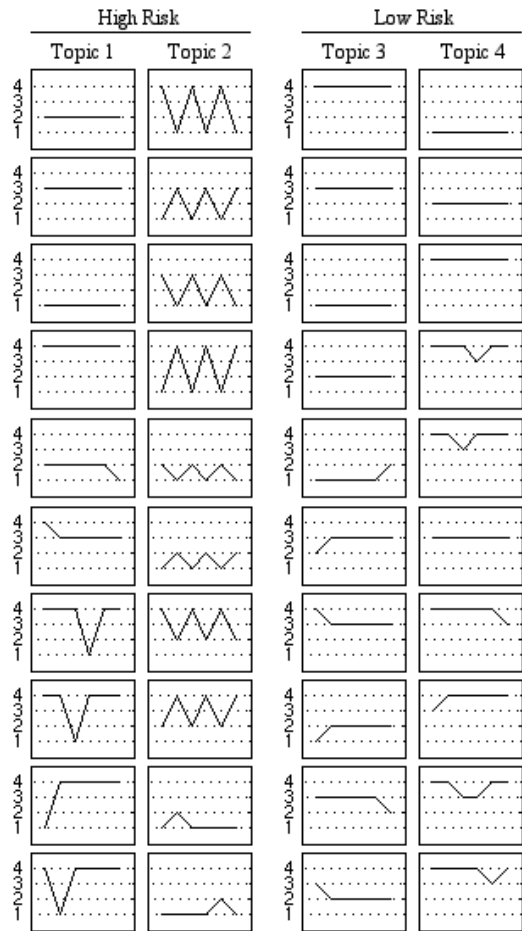
Figure 2: Ten most likely motifs for two of the topics most associated with high risk (left two columns) and low risk (right two columns).

Features	AUC
Motifs	0.652
Topics	0.674
TRS	0.659
TRS + Motifs	0.694
TRS + Topics	0.718

Table 1: AUC of logistic regression models trained with different feature sets with respect to incidence of 90-day cardiovascular death

Model 1	Model 2	IDI	p
TRS	TRS + Motifs	0.01	0.004
TRS	TRS + Topics	0.02	<0.001
TRS + Motifs	TRS + Topics	0.01	0.070

Table 2: IDI and corresponding p-values indicating the performance increase gained by using model 2 instead of model 1

Features	Hazard Ratio	p
Motifs	2.84	<0.001
Topics	2.78	<0.001
TRS	2.54	<0.001
Motifs (+TRS)	2.36	<0.001
Topics (+TRS)	2.40	<0.001

Table 3: Hazard ratios of different predictors with 90-day cardiovascular death times under a Cox proportional hazards model

Conclusions

This paper explores the application of topic models to heart rate time series inferred from long-term ECG. In contrast to prior work on cardiovascular assessment using heart rate motifs, which identified a small set of motifs predictive of a particular outcome, this approach leverages large numbers of patterns and uncovers higher-level structure in the data without information about patient outcomes, resulting in a sparse, task-independent representation. The topics provide novel insight into heart rate motifs, allow for a better understanding of short-term cardiac dynamics, and contributes novel information about patient health.

We evaluated these heart rate topics on a set of day-long ECG recordings from real-world data from patients with acute coronary syndrome. The topics performed well at identifying patients at high risk of cardiovascular death. This predictive power was independent of TRS, a commonly used clinical measure for risk stratification, meaning that integration of the two approaches has the potential to improve cardiac risk stratification. Inclusion of these topic features with TRS improved performance more than the inclusion of motif frequencies (in terms of the AUC, IDI and hazard ratio), indicating that topics, by identifying generative structure behind the short-term heart rate patterns, may be more useful for risk stratification than motifs. Importantly, this improvement in performance is accompanied by an increase in the ability to visualize and interpret cardiac activity with topic models.

There are several limitations to our study. First, we trained the topic models without supervision to allow for a task-independent representation. Integration of patient labels into the topic learning process, using a method such as the one described in (Blei and McAuliffe 2007) could improve upon the discriminative ability of the topics and give a more favorable comparison with alternative approaches. Second, further investigation of the methods on a larger dataset is needed to validate the approach's performance. Using data with longer follow-ups would also provide a more accurate analysis of the risk stratification performance. Finally, attributing physiological significance to the heart rate topics requires a more thorough evaluation that is beyond the scope of the current paper.

Acknowledgments

References

Antman, E.; Cohen, M.; Bernink, P.; McCabe, C.; Horacek, T.; Papuchis, G.; Mautner, B.; Corbalan, R.; Radley, D.; and Braunwald, E. 2000. The timi risk score for unstable angina/non–st elevation mi. *JAMA: the journal of the American Medical Association* 284(7):835.

Bardy, G.; Lee, K.; Mark, D.; Poole, J.; Packer, D.; Boineau, R.; Domanski, M.; Troutman, C.; Anderson, J.; Johnson, G.; et al. 2005. Amiodarone or an implantable cardioverter–defibrillator for congestive heart failure. *New England Journal of Medicine* 352(3):225–237.

Barthel, P.; Schneider, R.; Bauer, A.; Ulm, K.; Schmitt, C.; Schömig, A.; and Schmidt, G. 2003. Risk stratification after acute myocardial infarction by heart rate turbulence. *Circulation* 108(10):1221–1226.

Bauer, A.; Kantelhardt, J.; Barthel, P.; Schneider, R.; Mäkikallio, T.; Ulm, K.; Hnatkova, K.; Schömig, A.; Huikuri, H.; Bunde, A.; et al. 2006. Deceleration capacity of heart rate as a predictor of mortality after myocardial infarction: cohort study. *The Lancet* 367(9523):1674–1681.

Blei, D., and McAuliffe, J. 2007. Supervised topic models. In *Neural Information Processing Systems*.

Blei, D.; Ng, A.; and Jordan, M. 2003. Latent dirichlet allocation. *The Journal of Machine Learning Research* 3:993–1022.

Chia, C., and Syed, Z. 2011. Computationally generated cardiac biomarkers: Heart rate patterns to predict death following coronary attacks. In *SIAM International Conference on Data Mining*.

Cohen, M.; Grossman, A.; D., M.; M.M., K.; A.J., B.; and G.T., M. 2010. Identification of complex metabolic states in critically injured patients using bioinformatic cluster analysis. *Critical Care* 14(1):R10.

Hamilton, P., and Tompkins, W. 1986. Quantitative investigation of qrs detection rules using the mit/bih arrhythmia database. *Biomedical Engineering, IEEE Transactions on* (12):1157–1165.

Lin, J.; Keogh, E.; Lonardi, S.; and Chiu, B. 2003. A symbolic representation of time series, with implications for streaming algorithms. In *Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*, 2–11. ACM.

Malik, M. 1996. Heart rate variability. *Annals of Noninvasive Electrocardiology* 1(2):151–181.

Pencina, M.; D'Agostino Sr, R.; D'Agostino Jr, R.; and Vasan, R. 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.

Reddy, K. 2004. Cardiovascular disease in non-western countries. *New England Journal of Medicine* 350(24):2438–2440.

Roger, V.; Go, A.; Lloyd-Jones, D.; Adams, R.; Berry, J.; Brown, T.; Carnethon, M.; Dai, S.; de Simone, G.; Ford, E.; et al. 2011. Heart disease and stroke statistics2011 update. a report from the american heart association. *Circulation* 123(4):e18–e209.

Saria, S.; Koller, D.; and Penn, A. 2010. Learning individual and population level traits from clinical temporal data. In *Neural Information Processing Systems, Predictive Models in Personalized Medicine Workshop*.

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

Zong, W.; Moody, G.; and Jiang, D. 2003. A robust open-source algorithm to detect onset and duration of qrs complexes. In *Computers in Cardiology*, 737–740. Ieee.