Symmetric Projection Attractor Reconstruction analysis of murine electrocardiograms: Retrospective prediction of Scn5a$^{+/-}$ genetic mutation attributable to Brugada syndrome

Esther Bonet-Luz, MMath, PhD,* Jane V. Lyle, BSc,*
Christopher L.-H. Huang, MA, BMBCh, DM, DSc, PhD, MD, ScD,†‡§
Yanmin Zhang, MD, PhD,‖ Manasi Nandi, BSc, PhD,¶
Kamalan Jeevaratnam, DAHP, DVM, MMedSC, PhD, MRCVS,†‡ Philip J. Aston, BSc, PhD*

*From the *Department of Mathematics, University of Surrey, Guildford, United Kingdom, †Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom, ‡Physiological Laboratory, University of Cambridge, Cambridge, United Kingdom, §Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom, ‖Department of Paediatric Cardiology, Shaanxi Institute for Pediatric Diseases, Affiliate Children's Hospital of Xi'an Jiaotong University, Xi'an, China, and ¶School of Cancer and Pharmaceutical Sciences, Faculty of Life Sciences & Medicine, King's College London, London, United Kingdom.*

BACKGROUND Life-threatening arrhythmias resulting from genetic mutations are often missed in current electrocardiogram (ECG) analysis. We combined a new method for ECG analysis that uses all the waveform data with machine learning to improve detection of such mutations from short ECG signals in a mouse model.

OBJECTIVE We sought to detect consequences of Na^+ channel deficiencies known to compromise action potential conduction in comparisons of Scn5a$^{+/-}$ mutant and wild-type mice using short ECG signals, examining novel and standard features derived from lead I and II ECG recordings by machine learning algorithms.

METHODS Lead I and II ECG signals from anesthetized wild-type and Scn5a$^{+/-}$ mutant mice of length 130 seconds were analyzed by extracting various groups of features, which were used by machine learning to classify the mice as wild-type or mutant. The features used were standard ECG intervals and amplitudes, as well as features derived from attractors generated using the novel Symmetric Projection Attractor Reconstruction method, which reformulates the whole signal as a bounded, symmetric 2-dimensional attractor. All the features were also combined as a single feature group.

RESULTS Classification of genotype using the attractor features gave higher accuracy than using either the ECG intervals or the intervals and amplitudes. However, the highest accuracy (96%) was obtained using all the features. Accuracies for different subgroups of the data were obtained and compared.

CONCLUSION Detection of the Scn5a$^{+/-}$ mutation from short mouse ECG signals with high accuracy is possible using our Symmetric Projection Attractor Reconstruction method.

KEYWORDS Brugada syndrome; ECG signals; Machine learning; Scn5a$^{+/-}$ mutation; Symmetric Projection Attractor Reconstruction

(Heart Rhythm O^2 2020;1:368–375)

Introduction

The electrocardiogram (ECG) is an investigational tool fundamental to much electrophysiological study and clinical diagnosis of abnormal cardiac activity. These often arise from reentrant arrhythmias resulting from slowed action potential conduction. The latter process depends on Na^+ channel activation.[1] Deficiencies in these form the major cause, accounting for 15%–30% of cases, of the hereditary Brugada syndrome (BrS) associated with 4%–12% of clinically reported sudden cardiac deaths.[2] In the clinical condition, visible electrocardiographic features, including right precordial ST-segment elevation,[2,3] occur in only some BrS cases, necessitating invasive drug challenge testing that nevertheless produces false-negative results,[4] unmasking only 33% of asymptomatic patients in 1 study.[2] This suggests either that the ECG is not a reliable signal to be used for identification of BrS or that the ECG does carry the signature of BrS but that current analysis techniques are insufficient to

Address reprint requests and correspondence: Prof Philip J. Aston, Department of Mathematics, University of Surrey, Guildford GU2 7XH, UK. E-mail address: P.Aston@surrey.ac.uk.

2666-5018/

https://doi.org/10.1016/j.hroo.2020.08.007

KEY FINDINGS

- We considered the problem of detecting an Scn5a$^{+/-}$ mutation in mice from their electrocardiogram (ECG) signal. It was not clear whether the ECG carries the signature of the mutation. Our results show that using enhanced analysis techniques, which combine our Symmetric Projection Attractor Reconstruction method with machine learning, mice with the Scn5a$^{+/-}$ mutation can be distinguished from wild-type mice with high accuracy (96%) by analyzing their ECG signal.
- For mice, sex (M/F) and age (old/young) can be determined from the ECG signal but with lower accuracy than for genotype. For these classifications, the standard ECG interval and amplitude measures performed poorly and there was a significant improvement using Symmetric Projection Attractor Reconstruction features.
- Symmetric Projection Attractor Reconstruction is a robust technique and always performs better than standard ECG interval and amplitude measures in the context of discriminating between Scn5a$^{+/-}$ genotype, sex, and age in mice. However, the best classification performance is typically obtained by combining both sets of features.

reliably detect it. The aim of this paper is to consider new ECG analysis techniques to determine whether consistent detection from the ECG is possible.

A number of mathematical approaches have been developed to provide concise representations of the ECG waveform, and thereby simplify analysis.[5] Heart rate variability methods, which analyze beat-to-beat (R-R) intervals, are extensively used.[6] However, these reflect only the single, albeit important, aspect of heart rhythm, but omit information concerning the many remaining processes associated with arrhythmias. Although it has been suggested that heart rate variability metrics are more variable in symptomatic than asymptomatic BrS patients, this has not been shown to have diagnostic value.[7] Another mainstay of ECG signal analysis determines particular points on the signal from which various time intervals and amplitudes can be derived. This approach still does not incorporate all of the available waveform morphology. Some differences have been identified in interval measures for BrS patients,[8] but they do not appear to be sensitive or specific enough to improve on current identification, and further novel metrics are suggested.[9] In populations of wild-type and Scn5a$^{+/-}$ mutant mice, only 2 out of 6 ECG intervals were found to differ significantly between the 2 groups.[10] Thus, it appears that identification of BrS patients from their ECG would be challenging.

We have recently described a new approach, Symmetric Projection Attractor Reconstruction (SPAR), for analyzing the morphology of "approximately periodic" signals,[11,12] which we apply here to the problem of distinguishing between wild-type and BrS hearts from short ECG signals. This technique has a fundamentally different approach, as it employs all of the available high-fidelity waveform data. It thus preserves more information about the underlying processes being measured. The SPAR method compactly encapsulates the entire signal in the form of a symmetric 2-dimensional attractor, providing both a simple visualization of the waveform and a means of easy quantification. The method also reduces the effect of baseline variation and factors out changes in heart rate in order to concentrate on changes in the waveform morphology. In a physiological context, we have previously applied our SPAR method to various signals, including blood pressure,[11,13] photoplethysmogram signals,[14] and ECG,[15,16] where it was shown that SPAR could accurately discriminate male from female sex in human ECG signals.

Previous studies have modeled proarrhythmic changes and their possible underlying mechanisms in BrS in populations of Na^+ channel–deficient hearts with a single clear-cut Scn5a$^{+/-}$ mutation.[17] The latter provide genetically homogeneous experimental systems with specific monogenic ion channel abnormalities, not available to clinical investigations. Despite the anatomically smaller size and higher heart rate, mice share with humans considerable genetic homology, similar cardiac conduction systems, and arrhythmogenic features related to genetic change. They are therefore amenable to exploration of the SPAR technique for examining the physiological effects of a single clear-cut mutation bearing on the fundamental property of action potential conduction on ECG waveforms obtained from whole animals.[17]

Methods

Data

The research reported in this paper used data acquired from a previous study[10] that conformed to the ARRIVE guidelines. All procedures in that study complied with the UK Home Office regulations (Animal [Scientific Procedures] Act 1986) following ethical review by the University of Cambridge Animal Welfare and Ethical Review Body. The mice were housed in an animal facility at 21°C with 12-hour light/dark cycles and were fed sterile chow (RM3 Maintenance Diet; SDS, Witham, Essex, UK) with free access to water. Both the wild-type (WT) and Scn5a$^{+/-}$ (SCN5A) mice used were bred on a 129/sv background to avoid strain-related variations. Experimental procedures, fully described earlier,[10] placed the anesthetized mice supine on a temperature-regulated platform maintaining a 37°C body temperature, with small strips of adhesive tape lightly attached to the limbs to reduce small movements to avoid ECG recording artefacts. Two-millimeter-diameter needle recording electrodes (MLA1204; ADInstruments, Colorado Springs, CO) were connected to an input box leading to amplification (NL104) and filter (NL126) units (settings: low-frequency cut-off: 50 Hz; high-frequency cut-off: 500 Hz) mounted within an NL900D chassis and power supply

(NeuroLog-Digitimer, Hertfordshire, UK). Our approach to the placement of needle electrodes has been described previously.[10] To reduce variability in placement of needle electrodes, the same individual placed the needle electrodes for all the animals used in the study. This was undertaken by a veterinarian with previous clinical experience. A key reason for changes in amplitude is if (1) there are any small limb twitches (typically occurring when animals are not adequately anesthetized) or (2) the needle electrode becomes loose when limbs move. In our study all animals were adequately anesthetized and the limbs were secured with adhesive tape to ensure no movements, thus reducing motion artefacts. Moreover, our analysis normalizes the amplitude of the signals, so any variability in the magnitude of the signal is removed at this stage.

Recordings were performed within a grounded Faraday cage. Analog-to-digital conversion of ECG signals at 5000 Hz used a CED 1401 series interface (Cambridge Electronic Design, Cambridge, UK) connected to a computer. Spike II software (Cambridge Electronic Design) was used to record and perform initial processing of ECG recordings.

Signals were collected from a total of 42 anesthetized mice. Lead I and lead II recordings were obtained from 36 mice while a single lead recording was obtained from the remaining 6 mice (3 lead I and 3 lead II), giving 78 animal lead recordings in total. The duration of the recordings ranged from 21 seconds to 10 minutes. Close to equal numbers of old and young, male and female, and Scn5a$^{+/-}$ and WT mice were studied. The 78 animal lead signals could be grouped by lead (39 lead I, 39 lead II), sex (42 male, 36 female), age (39 young, 39 old), and genotype (42 WT, 36 SCN5A). Young mice were up to 3 months old while the old mice were at least 12 months old. The number of mice in the young female, old female, and young male groups was 5 and in the old male group was 6 for both the WT and SCN5A classes. Two lead recordings were made for all the WT mice, giving 42 signals. For the SCN5A mice, only 1 lead was recorded for 6 of the mice, resulting in 36 signals.

Symmetric Projection Attractor Reconstruction analysis of ECG signals

The SPAR approach is a new mathematical method that examines the ECG using a fundamentally different approach from other techniques.[11,12] The method replots all of the high-fidelity time series data of the ECG in a bounded box. The resulting attractor is an easily visualized 2-dimensional representation of the signal. (An animation that illustrates the method for generating an attractor is included as supplementary material.) It is overlaid with a density to clarify which parts are visited more or less frequently, and are therefore of higher or lower density. We can then extract metrics from the attractor to quantify the morphology and variability of the underlying ECG. In this study we used these metrics to classify whether an ECG signal is from a wild-type or an Scn5a$^{+/-}$ mutant animal.

For the SPAR analysis, we did not perform any further filtering or other preprocessing of the ECG signals. The full data with 5000 Hz sampling frequency was used, although the wave profile would still be clearly defined with data downsampled to a lower frequency, so we would anticipate similar results in this case. The first step was to find the average cardiac cycle length for the ECG data in a given time window. Peak detection (with outlier correction) was used to determine the R-R intervals, and the mean of these was taken as the average cycle length of the window. (This is different from the method proposed in Aston et al[11] for blood pressure data, as the prominent R peaks in clean ECG data make this a viable alternative.) The interval between 3 equally spaced points traversing the signal is the time delay parameter τ, which was taken as one-third of the average cycle length.[11,12] Signals were also normalized such that the average height of the R peaks was 1. The size of the attractor is dependent on the amplitude of the signal, so this normalization ensured that the generated attractors are of comparable size. The scaling factor of the normalization for each window was also taken as an attractor measure.

In contrast to the generally triangular nature of the attractor generated from blood pressure or photoplethysmogram signals,[11,14] the attractor of a lead I or II ECG typically has 3 long arms, predominantly representing the R peak in the signal, and a central core region, whose features depend more on the T- and P-wave morphologies (Figure 1; see also ref. 15). Mouse lead I and II ECG data may have deep S peaks, and in some cases deep Q peaks as well. These give rise to shorter arms of the attractor in the opposite direction from the long arms associated with the R peaks, so that a typical mouse attractor has 6 arms in total (Figure 1).

In addition to the average heart rate and the vertical normalization scaling factor, a further 72 manually defined features were derived from the attractor, giving 74 features in total that related to the density, size, and symmetry of the attractor. The derived features mostly relate to either the central core or the arms of the attractor, as these are associated with different parts of the signal.

Figure 1 shows a sample of lead I ECG data from young, female wild-type and mutant mice and their corresponding attractors generated from a 10-second window of data. This comparison demonstrates clear qualitative differences between the attractors. Of course these are only small samples from 2 mice. The problem we address is to achieve high accuracy in predicting these differences using the data from all the mice.

ECG intervals and amplitudes

We sought to benchmark results obtained using our SPAR method against those derived using more standard features. Thus, we also derived standard ECG intervals and amplitudes using LabChart (ADInstruments, Oxford, UK). Again, no further filtering or other preprocessing was used. Deriving these intervals required setting various parameters for each signal individually. Moreover, the same parameters could

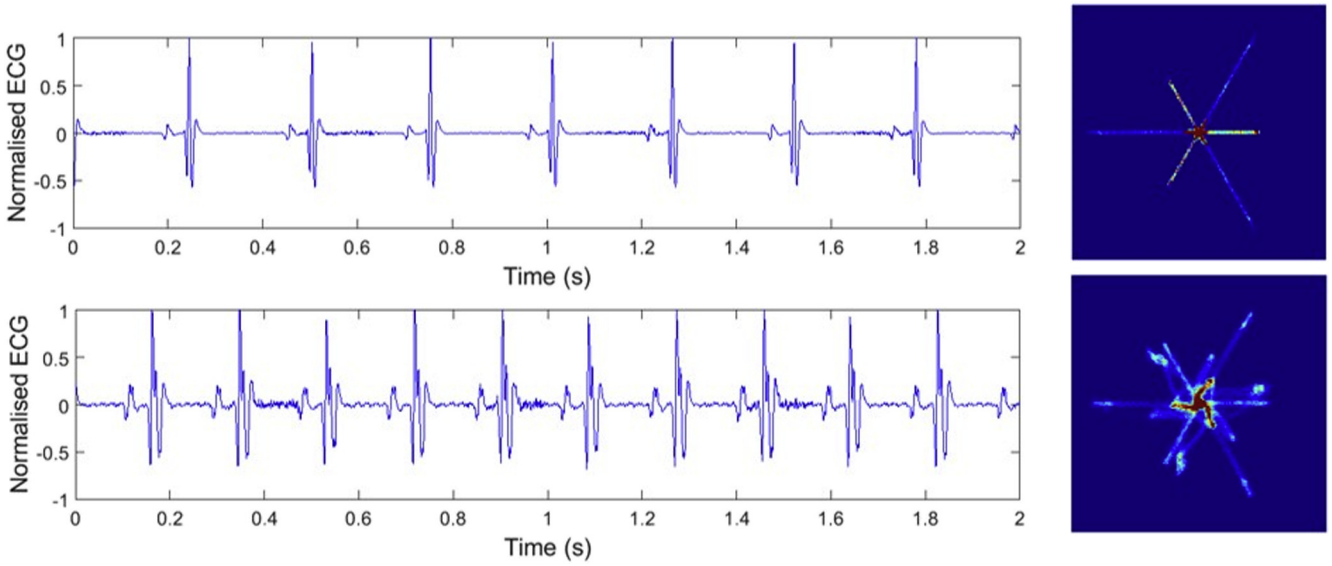

Figure 1 A sample of lead I electrocardiogram (ECG) data from a young, female, wild-type mouse (top) and from a young, female, Scn5a$^{+/-}$ mutant mouse (bottom) with the corresponding attractors.

not be used for all signals, as the standard parameters resulted in incorrect intervals being identified in some cases, so manual adjustments were made where necessary. The features extracted were:

- ECG intervals: R-R, PR, P duration, QRS, QTc (Bazett), JT, T_{peak} to T_{end}
- ECG amplitudes: P amplitude, Q amplitude, R amplitude, S amplitude, ST height, T amplitude

Machine learning

Our study sought to determine whether it was possible to distinguish between WT and Scn5a$^{+/-}$ mutant mice from short ECG signals, and was therefore addressing a simple binary classification of genotype. It also undertook binary classifications of sex and age from the same records.

For each of the ECG signals, we chose for analysis a 130-second interval of data free from artefactual noise, which was divided into 13 nonoverlapping 10-second windows. The only exception was 1 dataset that had a length of 21 seconds (lead I, Female, Old, Wild-Type), which was divided into 13 windows of length 1.6 seconds. (The average heart rate for this animal was just under 200 beats per minute, so each window contained just over 5 complete cycles, which was sufficient to generate an attractor as well as the intervals and amplitudes.) Thus, there were 13 windows of data for each ECG signal. For each window, an attractor was generated and the 74 features extracted together with the 7 ECG intervals and 6 ECG amplitudes. Thus, we had 3 sets of features, each containing $13 \times 78 = 1014$ records. We also generated a further feature set by combining all of the other feature groups; thus we present results for the 4 feature groups:

- ECG intervals (Int)
- ECG intervals + amplitudes (Int+Amp)
- SPAR features (SPAR)
- SPAR + intervals + amplitudes (SPAR+Int+Amp)

Our machine learning approach for the classification used a k-nearest neighbors algorithm with the common choice of $k = 3$.[18] We explored other values of k but overall $k = 3$ gave the highest accuracy of classification. We also explored a similar process that employed a (nonlinear with Gaussian kernel) support vector machine (SVM).

We applied a "leave 1 animal out" cross-validation, where we removed the 26 records (13 from lead I and 13 from lead II) for an animal and trained a machine on the remaining 988 records. The 26 removed records were then tested individually. This process was repeated for all 36 animals for which there were lead I and II recordings. For the remaining 6 mice, we formed a further 3 test sets, each also consisting of 26 records, by combining a lead I and lead II recording from different animals. The accuracies from the 39 cross-validation cycles were then averaged. For a small dataset, this gives a more consistent result than would be obtained by holding out 1 small test dataset and also helps to avoid overfitting the data.

For each group of features we used forward feature selection[19] to reduce the number of features up to a maximum of 20, selecting the number of features that gave the highest classification accuracy from the cross-validation. For the k-nearest neighbors approach, the highest classification accuracy was taken from an ensemble of 20 runs in which the order of the features was shuffled randomly before each run, as this allowed selection of different variables that have equal highest accuracy on a particular run. We report the classification accuracy both for individual records ($n = 1014$) and for a majority vote in which the classification for each animal lead is obtained as the majority prediction from the 13 individual records ($n = 78$). We did not perform multiple runs for the SVM method owing to the significantly greater computational time required for this method.

Table 1 Machine learning results for binary classifications

Classification	Int	Int+Amp	SPAR	SPAR+Int+Amp
Genotype	70.9%	83.0%	85.9%	90.9%
	74.3%	85.9%	87.2%	96.2%
Sex	60.4%	67.7%	76.7%	78.3%
	59.0%	71.8%	83.3%	80.8%
Age	66.2%	70.2%	79.2%	83.0%
	71.8%	73.1%	84.6%	88.5%

For each classification, the top row is the accuracy of individual records (n = 1014) and the bottom row is the accuracy for animal leads using majority vote (n = 78).

Int = electrocardiogram intervals; Int+Amp = electrocardiogram intervals + amplitudes; SPAR = Symmetric Projection Attractor Reconstruction features; SPAR+Int+Amp = Symmetric Projection Attractor Reconstruction features + intervals + amplitudes.

Results

We used machine learning with a *k*-nearest neighbors classifier to perform binary classifications to distinguish genotype, sex, and age. The results for individual records and majority vote are shown in Table 1. These results demonstrate a higher accuracy for genotype than for sex and age for all sets of features.

We considered the 41 animal leads that were misclassified by majority vote for genotype using either the Int, Int+Amp, or SPAR features. Of these, only 2 animal leads were misclassified by all 3 feature groups, but both of these were correctly classified by the combined features. There were a further 9 animal leads that were misclassified by 2 of the 3 groups and the remaining 17 were only misclassified by 1 of the groups, so there is little consistency in the misclassifications, which possibly explains why the combined features did well.

Although many results can be generated from these data, we then concentrated on the classification of genotype, given that the other 2 categories constitute known quantities.

We first considered the accuracy of the genotype classification for WT and SCN5A mice to assess consistency of prediction across the 2 classes. The confusion matrix for the majority vote classification of the animal ECG lead for the 4 sets of features is shown in Figure 2. The row-normalized table in Figure 2 demonstrates that the true prediction rate for mutant mice is slightly higher than for wild-type mice in all cases, except for the SPAR features, in which there is a small difference in the opposite direction.

A summary of the accuracy of the majority vote classification of genotype for various subgroups of the data is shown in Figure 3. When comparing the results for lead I and lead II, we note that the interval features give identical results, which is to be expected, as the intervals derived from the 2 leads should be very similar. For the other feature groups, the Int+Amp features performed best on lead II, whereas SPAR and the combined features performed best on lead I. The accuracy for males was higher than for females for all feature groups except the combined group. However, the accuracy for young animals was poorer than for old animals for 2 feature groups, with the reverse for the other 2 groups. Overall, the ECG intervals gave the worst results and all the features combined gave the best results in all cases, as expected.

For the smaller groups (lower panel in Figure 3), the results for young and old males were quite consistent across all feature groups, whereas the accuracy for young females was lower than for old females for the Int and Int+Amp features, with the reverse for the other 2 feature groups. We note that the Young Female group with the combined features was the only group for which the classification achieved 100% accuracy.

With the majority vote classification for animals, it was possible to have some individual records misclassified but still correctly classify the animal by ECG lead. A summary of the number of votes out of 13 for the classification of genotype is shown in Figure 4. These charts demonstrate that many animal leads had all 13 records correctly classified, with only a relatively small number of animal leads having between 7 and 12 correctly classified records. This was not unexpected, as we were using fairly clean data from anesthetized mice and clearly the category of genotype is consistent across all 13 windows. The standard interval features had the most animal ECG leads with 0 out of 13 correct

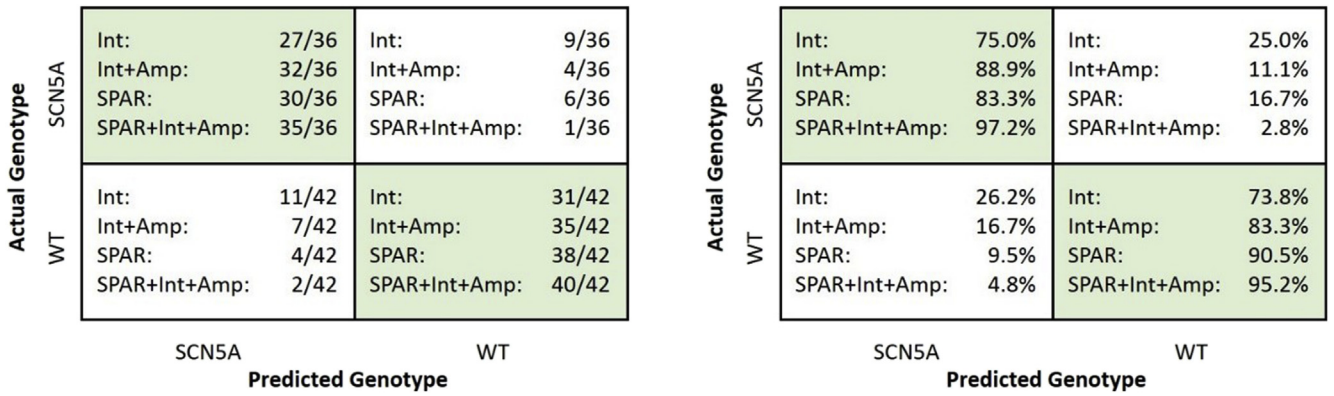

Figure 2 The confusion matrix for the majority vote classification of genotype (left) and with row normalization (right). Int = electrocardiogram intervals; Int+Amp = electrocardiogram intervals + amplitudes; SCN5A = Scn5a$^{+/-}$; SPAR = Symmetric Projection Attractor Reconstruction features; SPAR+Int+Amp = Symmetric Projection Attractor Reconstruction features + intervals + amplitudes; WT = wild-type.

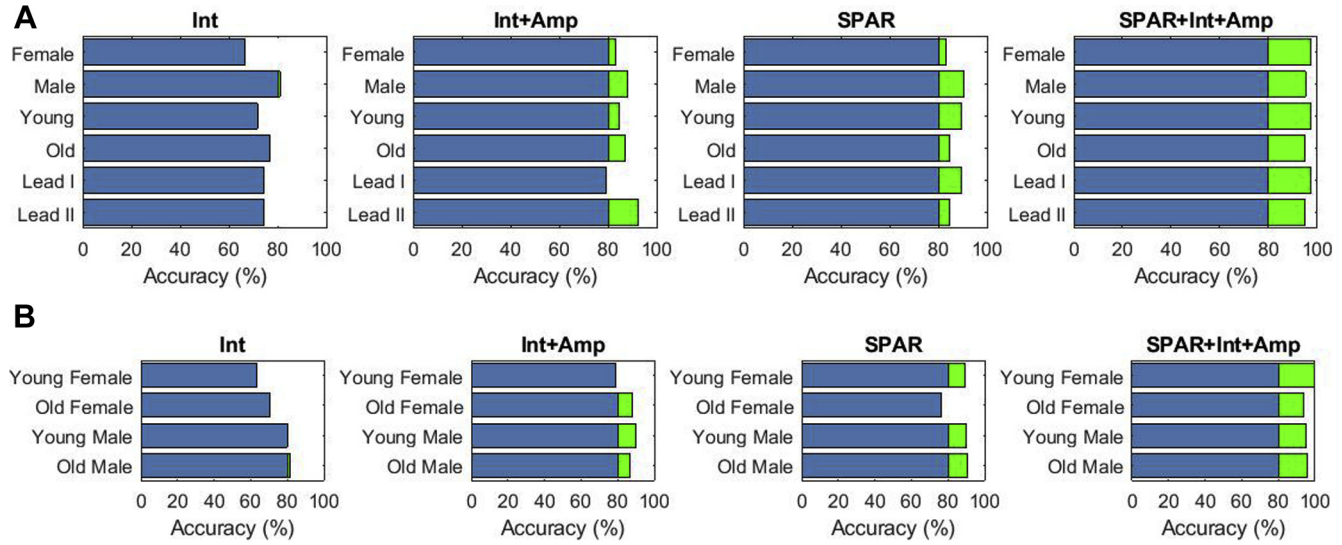

Figure 3 The accuracy of majority vote prediction of genotype for the different feature groups and for various subgroups of animals and electrocardiogram (ECG) leads. The green regions indicate accuracies that exceed 80%. Int = ECG intervals; Int+Amp = ECG intervals + amplitudes; SPAR = Symmetric Projection Attractor Reconstruction features; SPAR+Int+Amp = Symmetric Projection Attractor Reconstruction features + intervals + amplitudes.

votes. In contrast, the SPAR method classified at least some of the records correctly in the majority of cases, so it is interesting to note that the 3 misclassified animal ECG leads for the SPAR+Int+Amp features all have 0 out of 13 correct votes.

We also considered the features that were selected in the classification. When classifying for genotype, all the interval measures other than the QRS and R-R intervals were selected. Thus, the features used in the classification reflect atrial and repolarization activity. For the interval and amplitude measures together, features reflecting repolarization remained key, with amplitudes and intervals involving the T wave being selected, as well as the S-wave amplitude and ST height. For the SPAR features, the classification selected 7 features, predominantly those reflecting the density distribution and symmetry in the arm regions of the attractor. When SPAR, interval and amplitude measures were combined, machine learning predominantly selected SPAR features (16 out of 20), with the remaining features being the ST height, P and R amplitudes, and the PR interval. One of the classic indicators for BrS is ST-segment elevation in the precordial leads V_1–V_3,[1] so it is interesting to see that the ST height is selected in both feature groups that include the amplitudes.

The results obtained using an SVM classifier had slightly lower accuracies on average than those obtained using the *k*-nearest neighbors approach. The SVM method achieved classification with only a small number of variables in each case, and much less than the *k*-nearest neighbors method used.

Discussion

Overview

The present study applied a SPAR approach combined with machine learning to classify short ECG signals from murine wild-type and Scn5a$^{+/-}$ hearts by genotype, age, and sex for the first time. The main limitation of this study was the small number of mice from which the ECGs were collected. However, this is typical of studies involving animals, as it is labor intensive to collect these data. Nevertheless, our

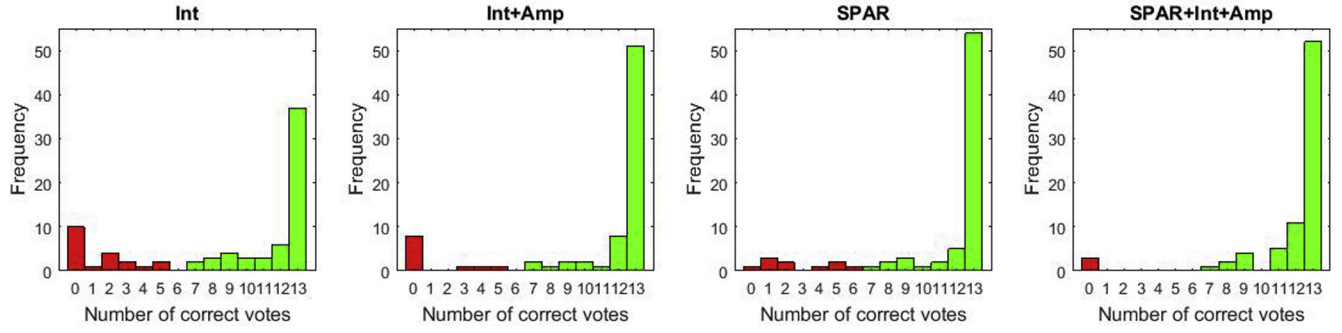

Figure 4 Histograms showing the distribution of votes for the classification of genotype. The green/red bars indicate votes that give correct/incorrect results. Int = electrocardiogram intervals; Int+Amp = electrocardiogram intervals + amplitudes; SPAR = Symmetric Projection Attractor Reconstruction features; SPAR+Int+Amp = Symmetric Projection Attractor Reconstruction features + intervals + amplitudes.

available data were sufficient to demonstrate that the SPAR approach better distinguishes signals from wild-type and Scn5a$^{+/-}$ mice than did standard ECG intervals or intervals and amplitudes. It also provides a higher accuracy of classification of sex (male/female) and age (young/old). However, the best results in all cases were obtained by combining all the features. These results confirm the robustness of the SPAR method.

Methodological issues

Deriving the ECG intervals and amplitudes required setting parameters in LabChart for each dataset, with some manual adjustment of the standard parameters being necessary in some cases; thus, the extraction of these features was not automatic. The SPAR features were obtained by setting parameters initially and then processing all the signals, and so did not require the manual effort involved in obtaining the intervals and amplitudes.

It may seem unfair to have extracted so many measures from the attractor in comparison to the 13 interval and amplitude features. However, finding individual points on an ECG in a consistent way from noisy signals is not an easy task, which is why these few points, associated with the dominant features of the signal, are commonly used in practice. With the SPAR method, because the averaging process of finding the attractor is done before extracting features, it is much easier to evaluate a variety of features from the attractor in a stable and consistent way.

When trying to understand which features of the attractor, and hence of the signal, are important for the classification, it may be better to work with the SVM features, as this was typically a much smaller set than was selected by the *k*-nearest neighbors method.

Clinical implications

The present results suggest that applications of SPAR to distinguish human cardiac electrophysiological conditions merit testing. The most direct extension might examine human BrS ECG data, bearing in mind the greater genetic heterogeneity of human populations and clinical BrS compared to the monogenic murine model explored here. Broader applications could screen and attempt SPAR characterizations of conditions with discrete conduction abnormalities exemplified by right or left bundle branch block extending, or accessory bundles, as in Wolff-Parkinson-White syndrome, shortening conduction times. They could extend to SPAR analysis of ECG recordings beyond leads I and II; BrS is most frequently diagnosed using ECG leads V_1–V_3,[2] but valuable information may be extracted from other leads. ECG analysis might then provide simple, potentially valuable, noninvasive clinical screening detection of proarrhythmic tendency, amenable to continuous monitoring for intermittent cardiac abnormalities.

Conclusions

We have shown that it is possible to distinguish between wild-type and Scn5a$^{+/-}$ mutant mice with high accuracy (96%) using short ECG signals. This suggests that the answer to the question we posed in the introduction is that there is indeed sufficient information in the ECG to detect this specific genetic mutation in mice using enhanced analysis techniques.

Funding

The early stages of this work were supported by EPSRC Impact Acceleration Account funding.

Disclosures

Esther Bonet-Luz was funded by the EPSRC Impact Acceleration Account during the early stages of this work. Philip Aston and Manasi Nandi have a patent (WO2015121679A1 "Delay coordinate analysis of periodic data"), which covers the foundations of the SPAR method used in this paper.

Appendix

Supplementary data

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.hroo.2020.08.007.

References

1. Jeevaratnam K, Guzadhur L, Goh YM, Grace AA, Huang CLH. Sodium channel haploinsufficiency and structural change in ventricular arrhythmogenesis. Acta Physiol 2016;216:186–202.
2. Brugada J, Brugada R, Antzelevitch C, Towbin J, Nademanee K, Brugada P. Long-term follow-up of individuals with the electrocardiographic pattern of right bundle-branch block and ST-segment elevation in precordial leads V1 to V3. Circulation 2002;105:73–78.
3. Hoogendijk MG, Opthof T, Postema PG, Wilde AAM, de Bakker JMT, Coronel R. The Brugada ECG pattern. Circ Arrhythm Electrophysiol 2010;3:283–290.
4. Pappone C, Santinelli V. Brugada syndrome: Progress in diagnosis and management. Arrhythm Electrophysiol Rev 2019;8:13–18.
5. Clifford GD, Azuaje F, McSharry PE. Advanced Methods and Tools for ECG Data Analysis. Artech House Publishers; 2006.
6. Acharya UR, Joseph KP, Kannathal N, Lim CM, Suri JS. Heart rate variability: a review. Med Bio Eng Comput 2006;44:1031–1051.
7. Behar N, Petit B, Probst V, et al. Heart rate variability and repolarization characteristics in symptomatic and asymptomatic Brugada syndrome. Europace 2017;19:1730–1736.
8. Mizumaki K, Fujiki A, Nishida K, et al. Bradycardia-dependent ECG changes in Brugada syndrome. Circ J 2006;70:896–901.
9. Batchvarov VN. The Brugada syndrome - Diagnosis, clinical implications and risk stratification. Eur Cardiol 2014;9:82–87.
10. Jeevaratnam K, Zhang Y, Guzadhur L, et al. Differences in sino-atrial and atrio-ventricular function with age and sex attributable to the Scn5a$^{+/-}$ mutation in a murine cardiac model. Acta Physiol 2010;200:23–33.
11. Aston PJ, Christie MI, Huang YH, Nandi M. Beyond HRV: Attractor reconstruction using the entire cardiovascular waveform data for novel feature extraction. Phys Meas 2018;39:024001.
12. Nandi M, Venton J, Aston PJ. A novel method to quantify arterial pulse waveform morphology: Attractor reconstruction for physiologists and clinicians. Phys Meas 2018;39:104008.
13. Aston PJ, Nandi M, Christie MI, Huang YH. Comparison of attractor reconstruction and HRV methods for analysing blood pressure data. Comp Cardiol 2014;41:437–440.

14. Charlton PH, Camporota L, Smith J, et al. Measurement of cardiovascular state using attractor reconstruction analysis. Nice: Proc 23rd European Signal Processing Conference (EUSIPCO); 2015. p. 444–448.
15. Lyle JV, Charlton PH, Bonet-Luz E, et al. Beyond HRV: Analysis of ECG signals using attractor reconstruction. Comp Cardiol 2017;44:091–096.
16. Lyle JV, Nandi M, Aston PJ. Investigating the response to dofetilide with Symmetric Projection Attractor Reconstruction of the electrocardiogram. Comp Cardiol 2019;46:073.
17. Martin CA, Zhang Y, Grace AA, Huang CLH. In vivo studies of Scn5a$^{+/-}$ mice modeling Brugada syndrome demonstrate both conduction and repolarization abnormalities. J Electrocardiol 2010; 43:433–439.
18. Cunningham P, Delaney SJ. k-nearest neighbours classifiers. Technical Report UCD-CSI-2007-4. University College Dublin 2007.
19. Guyon I, Elisseeff A. An introduction to variable and feature selection. J Mach Learn Res 2003;3:1157–1182.