

BRAVEHEART: Open-source software for automated electrocardiographic and vectorcardiographic analysis

Hans Friedrich Stabenau, MD, PhD^a, Jonathan W. Waks, MD^a

^a*Harvard-Thorndike Electrophysiology Institute, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA*

Corresponding Author:

Jonathan W. Waks, MD

Beth Israel Deaconess Medical Center

185 Pilgrim Road, Baker 4

Boston, MA 02215

Phone: (617) 667-8800

Fax: (617) 632-7620

Email: jwaks@bidmc.harvard.edu

Brief Title: BRAVEHEART ECG/VCG Analysis Software

Disclosures: JWW is on the advisory board for HeartcoR Solutions and has received grant support from Anumana, both unrelated to this study.

Abstract

Background and Objectives: Electrocardiographic (ECG) and vectorcardiographic (VCG) analyses are used to diagnose current cardiovascular disease and for risk stratification for future adverse cardiovascular events. With increasing use of digital ECGs, research into novel ECG/VCG parameters has increased, but widespread computer-based ECG/VCG analysis is limited because there are no currently available, open-source, and easily customizable software packages designed for automated and reproducible analysis.

Methods and Results: We present BRAVEHEART, an open-source, modular, customizable, and easy to use software package implemented in the MATLAB programming language, for scientific analysis of standard 12-lead ECGs acquired in a digital format. BRAVEHEART accepts a wide variety of digital ECG formats and provides complete and automatic ECG/VCG processing with signal filtering to remove high- and low-frequency artifact, non-dominant beat identification and removal, accurate fiducial point annotation, VCG construction, median beat construction, customizable measurements on median beats, and output of measurements and results in numeric and graphical formats.

Conclusions: The BRAVEHEART software package provides easily customizable scientific analysis of ECGs and VCGs. We hope that making BRAVEHART available will allow other researchers to further the field of EVG/VCG analysis without having to spend significant time and resources developing their own ECG/VCG analysis software and will improve the reproducibility of future studies. Source code, compiled executables, and a detailed user guide can be found at <http://github.com/BIVectors/BRAVEHEART>. The source code is distributed under the GNU General Public License version 3.

Keywords: electrocardiogram, vectorcardiogram, software, open-source, annotation

1. Introduction

The electrocardiogram (ECG) is a critical tool for clinical cardiovascular medicine and cardiovascular medical research. ECG analysis is used to diagnose existing cardiovascular disease and for risk stratification of future cardiovascular events, both in clinical practice and for research purposes. Although historically ECGs have been read by trained physicians on paper printouts, as medicine becomes more computerized, there has been increasing use of computerized ECG annotation and interpretation [1]. With increased availability of digitized ECGs there is also increasing interest in measurements which require additional signal processing of the standard 12-lead ECG and which cannot be obtained by visual inspection of the standard 12-ECG alone.

Computerization and digitization of ECGs have also allowed widespread analysis of vectorcardiograms (VCGs), where the electrical activity of the heart is visualized in 3 dimensions, and associated vectorcardiographic measurements [2]. These measures, which include the spatial ventricular gradient (SVG) [3], sum absolute QRST integral (SAI QRST) [4], QRST angle [5], total cosine R to T (TCRT) [6], and quantification of vectorcardiographic loop morphology [7], have been helpful in diagnosing current cardiovascular disease and predicting future adverse cardiovascular events across multiple studies with diverse patient populations.

Although clinically promising, widespread study of these and other VCG measurements has been limited because they require additional computer processing and specialized software. Previously described software includes the ECGlib software library and ECGlab graphical user interface (GUI) [8, 9], the Leiden ECG analysis and decomposition software (LEADS) software package [10], and other programs [11, 12]. However, these programs are not publicly available [10, 11, 9], and available open-source software packages [8, 12, 13, 14] are libraries for specific parts of ECG/VCG processing, such as R peak detection, dominant beat labeling, or fiducial point annotation, rather than full programs that allow complete processing of ECG data from file to measurements.

ECG/VCG research therefore usually requires investigators to create custom software which can be time consuming to develop and test, and which makes reproducible

measurement difficult because variations in filtering, signal baseline definition, and signal processing preclude the direct comparison of measurements. Reproducible analysis is also a special priority as the reproducibility of results is of growing concern in scientific research [15]. We therefore set out to develop open-source, easily customizable software designed to support ECG/VCG research with high-throughput, automated, and reproducible ECG and VCG analysis with the Beth Israel Analysis of Vectors of the Heart (BRAVEHEART) software library.

BRAVEHEART addresses the limitations of existing software as a modular software library meant to perform reproducible annotations and ECG and VCG measurements for scientific analysis on standard 12-lead ECGs acquired in a digital format. The design goals of BRAVEHEART include:

- A customizable, modular, and easily extensible design makes it simple to add new measurements or ECG formats as needed.
- Signal filtering and baseline correction performed in a physiologically appropriate way.
- Automated operation that is robust to non-dominant beats (premature ventricular contractions [PVCs], aberration, ventricular pacing, etc.), artifact, and noise.
- Reproducible ECG/VCG annotation, signal segmentation, and measurements.
- Automatic and efficient batch processing of ECGs with parallel processing capability.
- Open source MATLAB (MathWorks, Natick, MA) code as well as compiled executable code for researchers without access to MATLAB.

2. Methods/Software Description

BRAVEHEART is a software package which performs automated ECG/VCG analysis for research purposes. It takes a digital 12-lead ECG signal in a variety of formats, and provides automatic signal filtering, QRST fiducial point annotation, median beat alignment and construction, and measurement of a variety of ECG/VCG parameters on the median beat. It is also designed to be easily customizable. BRAVEHEART currently runs on

MATLAB versions above R2022a. Matlab source code, compiled executables for Windows and Mac operating systems, and a detailed user guide can be found at <http://github.com/BIVectors/BRAVEHEART>. The source code is distributed under the General Public License (GPL) version 3.

2.1. BRAVEHEART Overview:

BRAVEHEART ECG processing proceeds via the following steps:

1. Reading in the digital ECG in a variety of formats.
2. Wavelet-based filtering for high-frequency and low-frequency (baseline wander) noise removal.
3. VCG construction.
4. Peak thresholding for heart rate (HR) estimation and locating QRS complexes.
5. Baseline offset correction of the VCG leads.
6. Median filtering for estimation of QRS width and automatic pacemaker spike detection and removal.
7. Preliminary heuristic fiducial point (QRS onset [Q_{on}], QRS offset [Q_{off}], and T wave offset [T_{off}]) annotation.
8. Non-dominant QRST morphology (PVC, ventricular pacing, or aberration) detection: non-dominant beats can be analyzed separately or automatically excluded from subsequent analyses.
9. Outlier beat detection: outlier beats can be automatically excluded.
10. Median beat VCG construction.
11. Median beat fiducial point (Q_{on} , Q_{off} , and T_{off}) annotation using a custom neural network (NN).
12. VCG signal and annotation quality assessment.
13. Graphical display and analysis of the VCG median beat.
14. Export of ECG/VCG measurement data to an external file for analysis.

2.2. Reading Digital ECG Formats:

BRAVEHEART is able to read a variety of common digital ECG formats, including extensible markup language (XML), from a variety of ECG recording system

manufacturers such as General Electric (GE), Philips, and Mortara Instruments. HL7 XML is also supported. Other currently readable formats include Digital Imaging and Communications in Medicine (DICOM), International Society for Holter and Noninvasive Electrocardiology (ISHNE) .ecg files, GE Marquee .mrq files, GE Prucka format .txt files, and unformatted numeric lead data in text files. Each file format has its own load module, and it is therefore easy to integrate new file formats as needed.

2.3. ECG Signal Filtering:

Once ECG signals are loaded, wavelet-based filtering is performed to remove noise/artifacts which can hamper ECG segmentation and processing. For high-pass filtering of each ECG lead, a discrete wavelet decomposition is computed. The approximated baseline signal is then reconstructed using wavelet decomposition levels below a configurable bandpass which is nominally set to 0.24 Hz. This low frequency signal which approximates baseline wander and respiration effects without influencing the main frequencies contained in the QRST complex (0.5-40 Hz) [16], is then subtracted from the original signal.

Low-pass filtering is achieved for each lead using a maximal overlap discrete wavelet decomposition using a level which is nominally set to the 62.5 Hz bandpass with a soft thresholding scheme to remove high frequency noise above this threshold. For both high- and low-pass filtering, the wavelet and level used as a frequency cutoff are adjustable by the user, although the nominal settings of high-pass at 0.24 Hz and low-pass of 62.5 Hz work well for most applications. Certain ECGs with significant baseline wander or high frequency artifact may require adjustment of these settings. Filtering can also be disabled if needed for specific applications. An example of ECG filtering and baseline wander removal using this methodology is shown in Figure 1.

2.4. Vectorcardiogram (VCG) Construction:

A $3 \times n$ VCG matrix (\mathbf{V}) with orthogonal X, Y, and Z leads as its rows is constructed from a $n \times 8$ ECG matrix (\mathbf{E}) with the 8 independent ECG leads (I, II, V1-V6) as its

columns using a 3×8 transformation matrix (\mathbf{M}):

$$\mathbf{V} = \mathbf{ME}^T \quad (1)$$

The Kors transformation matrix [17] is nominally used for \mathbf{M} , although other transformation matrices [2] (such as the inverse Dower transformation [18]) can be substituted. The vector magnitude (VM) lead is then constructed by taking the Euclidean norm of the VCG X, Y, and Z leads:

$$VM(t) = \sqrt{X(t)^2 + Y(t)^2 + Z(t)^2} \quad (2)$$

The VM signal is used for all subsequent annotations as it represents the global electrical activation of the heart as a combination of all 8 independent leads [19], and it is always positive which simplifies signal processing and annotation.

2.5. QRS Peak Thresholding:

QRS complexes are detected in the VM lead by looking for peaks nominally in the top 5% of values, subject to a maximum HR constraint which helps in cases where there is significant fractionation of the QRS complex or very large amplitude T waves. The threshold parameter and maximum HR constraint may be manually adjusted in order to process ECGs that feature high-amplitude measurement artifacts, very peaked T waves, large differences in amplitude between native QRS complexes and PVCs, or paced QRS complexes. Peak thresholding is illustrated in Figure 2.

2.6. Baseline Offset Correction of the Vectorcardiogram:

ECG machines usually assign the zero voltage reference to a point that is not necessarily the physiological zero voltage. This may be computationally advantageous during ECG signal acquisition using an ECG machine where visual inspection of the baseline rather than exact voltage measurement relative to zero voltage is performed in most cases and for most clinical purposes. However, when measurements are being made on ECG leads, especially measurements involving area under the QRST complex, an

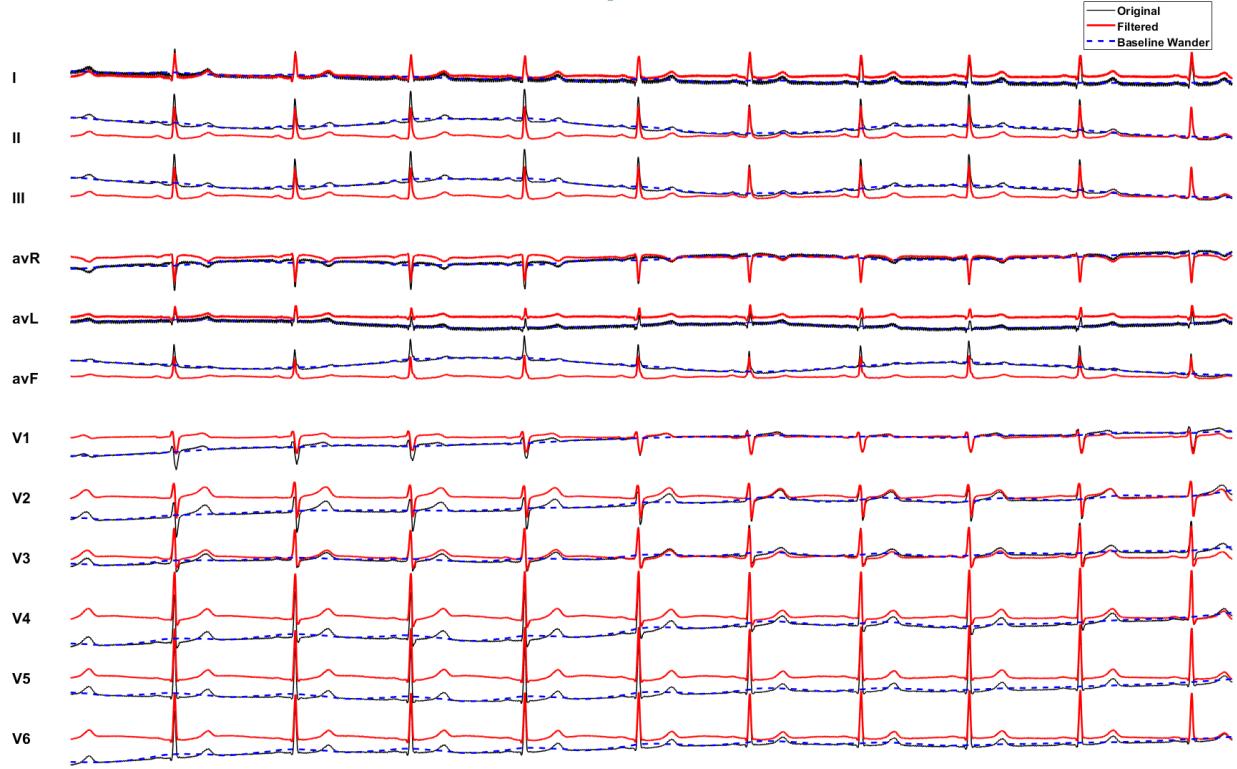


Figure 1: Example of utilizing wavelet decomposition to remove significant low-frequency artifact from an ECG with a sampling rate of 500 Hz. The original ECG signal is in black and the filtered signal is in red. In this case the significant baseline wander, as illustrated by the dashed blue line, is estimated and then subtracted out of the final signal so that the baseline is flat. In this example, modest high-frequency noise is also reduced.

inaccurate baseline zero voltage reference can lead to large errors in these measurements (See Supplemental Figure 1). Filtering can introduce baseline offsets as well. Therefore, in order to ensure accurate measurements, a method of finding the physiologic zero voltage reference after filtering is needed.

During the TP interval, after ventricular repolarization is complete and before start of atrial depolarization (in sinus rhythm), the heart is electrically silent, and therefore the voltage in all ECG leads should approximate 0 mV. While each ECG lead does undergo baseline wander correction during high-pass filtering as noted above, and this can remove large constant (0 Hz) shifts, there is often still a residual baseline offset which can be due to noise, respiration, or amplifier drift. The high-pass filter does not use any physiologic information and has no way of ensuring that the TP segment of the signal equals 0 mV. When these offset ECG leads are then projected onto X, Y, and Z coordinates, the residual

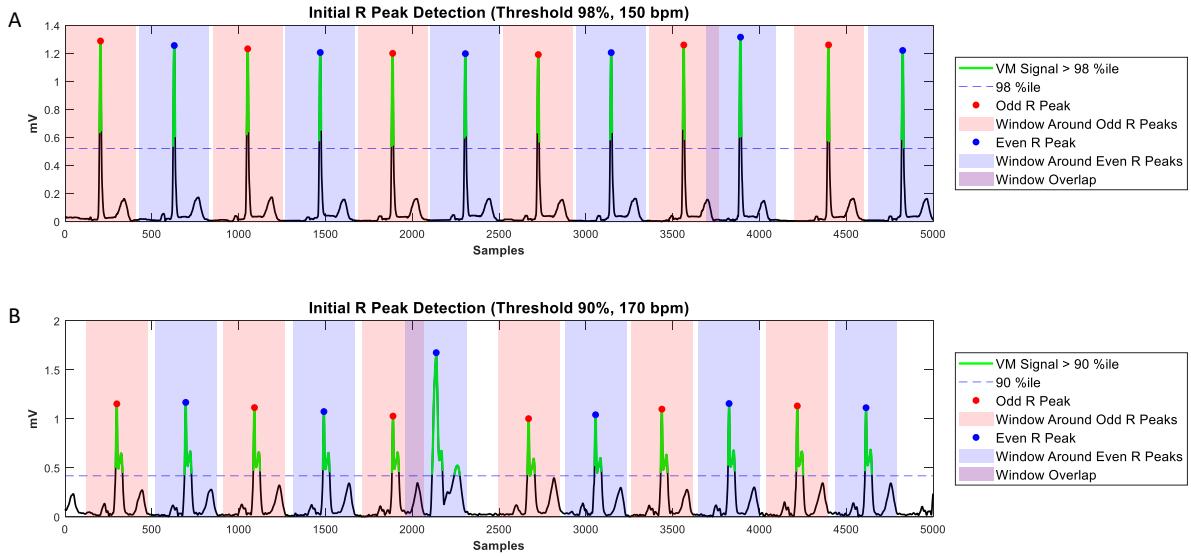


Figure 2: Example of VM lead peak thresholding with HR constraints. **A:** Parts of the signal which are above the 98th percentile are highlighted in green. R peaks are chosen as the dominant peak within a window (odd: red, even: blue, with width set by a chosen maximal HR). In this example the R peaks are significantly larger than the T waves and there are not any significant extra peaks in the QRS complex, and R peak detection is therefore straightforward. **B:** In this example there are 2 peaks in each QRS complex and a PVC. The threshold was artificially lowered below what would nominally be chosen for illustrative purposes. Even with a lower than normal threshold of 90%, only a single peak is found for each QRST complex, and the second minor QRS peak and the PVC T wave are ignored. Had the PVC been more tightly coupled the maximum HR constraint could be increased.

Abbreviations: PVC - premature ventricular contraction, VM - vector magnitude, HR - heart rate.

offsets can add up to an even larger baseline offset [20, 21].

In order to optimize the VCG for quantitative measurements such as area and angle measurements, it is essential to zero the ECG lead baselines in a physiologically meaningful way by adding a constant offset so that the TP segment approximates 0 mV voltage. This re-zeroing process can be performed at any time before or after transformation of the ECG into a VCG, but doing so after VCG construction is advantageous as fewer leads need to be corrected (3 vs 8), and the ECG baseline offsets have no effect on the transformation from ECG to VCG other than introducing a new offset in the transformed signals (matrix multiplication is distributive for constants).

For each lead, this offset is determined by the following procedure: First, the signal is smoothed with a 4th order Savitzky-Golay filter with window size equal to 10% of the average cardiac cycle length which has the effect of smoothing out sharp features of the

ECG signal. Next, the set of points corresponding to flat regions of the ECG signal is determined by finding the set of points where the slope of the filtered signal is less than 2% of the maximum slope. Finally, for each lead, the median of this set of points is determined, which is then the offset by which the lead is subsequently shifted. This baseline offset procedure is robust to noisy signals and atrial fibrillation, and functions as long as the HR is not so tachycardic that there is no appreciable TP segment (or TQ segment in atrial fibrillation) which would render the ECG difficult to quantitatively analyze for other reasons. An example of baseline correction is shown in Figure 3.

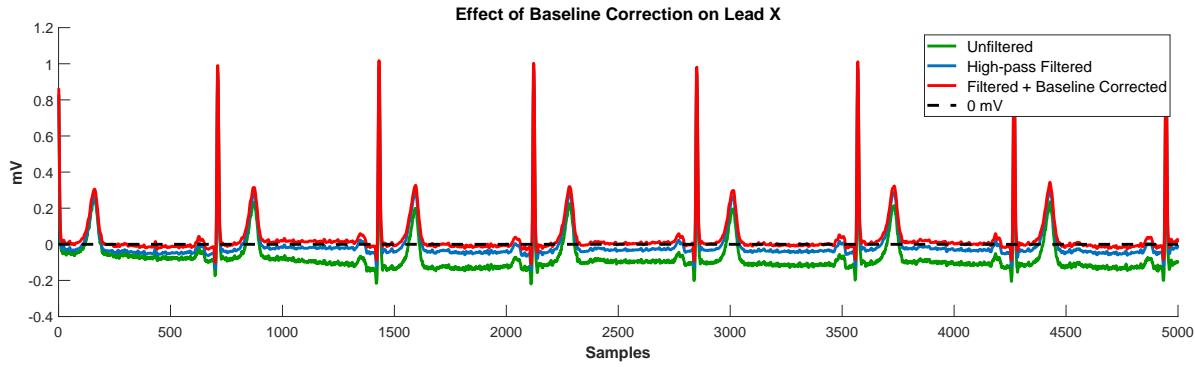


Figure 3: Example of baseline correction after filtering. The dashed black line corresponds to 0 mV. The green signal is the raw unfiltered ECG signal (lead X) with 0 mV voltage assigned by the ECG recording system. The blue signal is lead X after high-pass filtering; most of the baseline offset is removed by filtering, but the TP segment is not appropriately set to zero voltage. The red signal is lead X after additional baseline correction as described in the methods. After baseline correction the TP segment lead X now approximates 0 mV.

2.7. Median Filtering for Pacemaker Spike Detection:

For each peak detected in the VM lead, pacemaker spikes are identified (if present) with the following procedure: first, the signal is smoothed with a 4 ms median filter. The start and end of each R peak previously detected in the VM signal (see Section 2.5) is annotated by marching out from the peak forwards and backwards and marking where the signal first crosses an adjustable threshold which is nominally set at 20% of the peak height.

Candidate QRS peaks that are less than a certain threshold in width (default: 20 ms) are marked as pacing spikes, as these spikes occur over time intervals that are not physiologic for a QRS complex. If pacing spikes are found they are simply ignored and the

QRS-finding procedure is repeated. The HR is then estimated as the mean distance between R peaks (RR intervals) over the entire ECG.

2.8. Heuristic First Pass VCG Fiducial Point Annotation:

It has been shown that performing calculations on median beats is preferable to calculating a quantity on each beat individually and then finding the mean or median of measurements [19, 22]. Accurate annotation of each individual QRST complex is therefore important primarily to allow accurate alignment of each individual beat to create an accurate median beat which will be used for all subsequent measurements. For this reason, it is less critical that each individual beat be perfectly annotated, as small errors in fiducial point annotation do not preclude accurate median beat creation which relies primarily on aligning the major features of the QRS complex/peaks rather than aligning beats based on the exact location of fiducial points.

To allow rapid computation of fiducial points with a simple way of visualizing the annotation algorithm and troubleshooting problematic ECGs or ECGs with atypical QRST complexes, first pass QRST annotation utilizes standard signal processing methods on the VM lead. For each R peak that is present, the QRS width is first estimated (similarly to the previous step in Section 2.7) using a 40 ms median filter with an adjustable threshold nominally set at 20%. A QRS search window is then defined around each R peak. The adjustable QRS search window is set to be equal in duration to twice the estimated QRS width, centered around the R peak. The adjustable T wave search window starts by default 100 ms after the end of the located S-wave/ Q_{off} and nominally extends forward to 45% of the mean RR interval. Rarely, it is necessary to adjust this percentage based on HR, QT interval, and QRST morphology. The onset and offset of the QRS complex are determined by an algorithm that is described in detail in the “Heuristic Annotation Details” section of the Supplemental Methods (Supplemental Section 1.1). Details on the search windows and an example of first pass annotation is shown in Figure 4.

T-wave offset detection can use baseline crossing, the tangent method [23], or a validated method that is more robust to abnormal T wave morphology that estimates minima of the T wave “energy” and which is more accurate than these other methods,

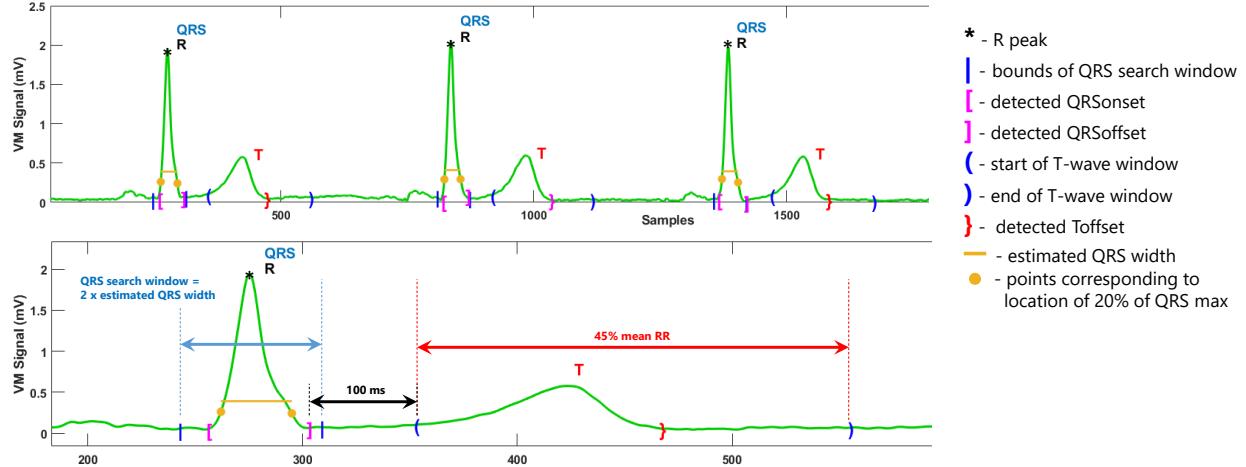


Figure 4: Example of first pass heuristic annotation. The upper panel shows 3 sequential beats, and the lower panel is zoomed in on beat 1. Search windows and the methodology for fiducial point detection are described in detail in the text. Using nominal settings, the QRS search window (blue arrow and blue | symbols) is set as 2 times the estimated QRS width (yellow bar) centered around the QRS peak. The T wave search window (red arrows and blue parentheses) nominally starts 100 ms after QRS offset, and extends forward 45% of the mean RR interval. These nominal parameters are adjustable.

especially in ECGs with low amplitude or notched T waves [24]. After first pass annotation, beats too close to the start or end of the ECG so that an entire QRST complex is not present are removed from analysis.

2.9. Non-dominant Beat (PVC, Pacing, and Aberration) Detection:

PVCs or other non-dominant beat classes, which include intermittent QRS aberration or intermittent ventricular pacing, are detrimental to median beat construction, especially when frequent non-dominant beats are present. For example, in cases of ventricular bigeminy where there are equal numbers of PVCs and normal QRST complexes, unless PVCs are removed, the median QRST complex will be a superposition of the PVC and normal QRST morphologies, with a morphology that is similar to neither (see Figure 5). For this reason it is also not always possible to simply compare individual beats to the median beat for non-dominant beat detection. BRAVEHEART automatically identifies and can remove non-dominant QRST morphologies. This process will identify any non-dominant QRST morphology including PVCs, intermittent ventricular pacing, and intermittent QRS aberration. All of these types of non-dominant QRST complexes are considered “PVCs” for purposes of the following method description.

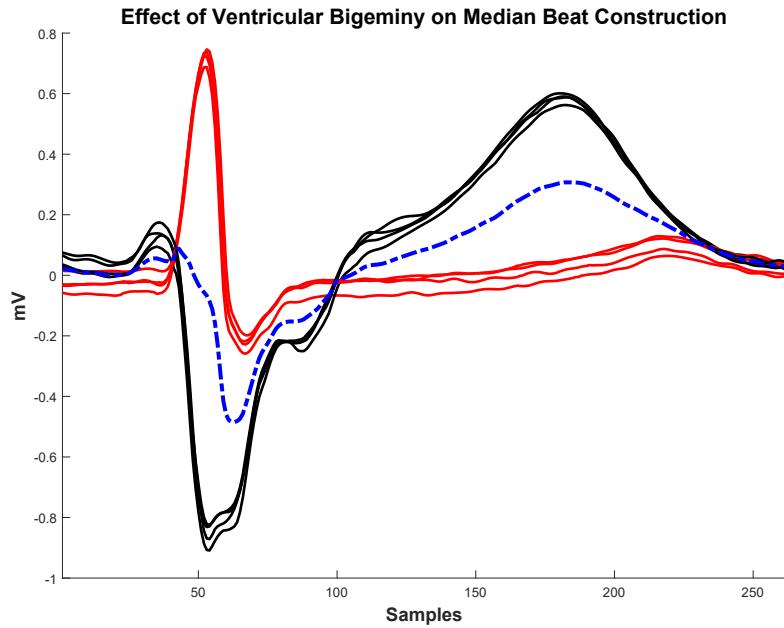


Figure 5: Effect of ventricular bigeminy on median beat construction. If ventricular bigeminy is present and the full ECG strip contains equal numbers of normal beats (red) and PVCs (black), the created median beat (dashed blue line) will be a superposition between the normal QRST complexes and PVC QRST complexes, and will result in incorrect measurements.

Abbreviations: PVC - premature ventricular contraction.

BRAVEHEART PVC detection utilizes normalized cross correlation (NCC) and root mean squared error (RMSE) to perform a form of “template matching” between individual QRST complexes. A detailed description of the PVC detection algorithm is available in the “PVC Detection Algorithm Details” section of the Supplemental Methods (Supplemental Section 1.2), Figure 6, and Supplemental Figures 2 and 3.

Using a single lead for PVC detection may limit the sensitivity and specificity of PVC detection because PVCs can be similar in morphology to dominant, non-PVC QRST complexes in one ECG lead, while having dramatically different morphology in other leads. BRAVEHEART therefore performs the PVC identification algorithm on the 3 orthogonal X, Y, and Z leads separately. A beat is ultimately identified as a PVC if it is identified as a PVC in at least 2 of the 3 orthogonal leads. This also helps avoid artifact in a single lead triggering incorrect PVC identification. The beats identified as PVCs can be removed from subsequent analyses, including median beat construction. The software also has the ability to also keep PVCs and remove non-PVC QRST complexes should analysis of PVCs be

desired.

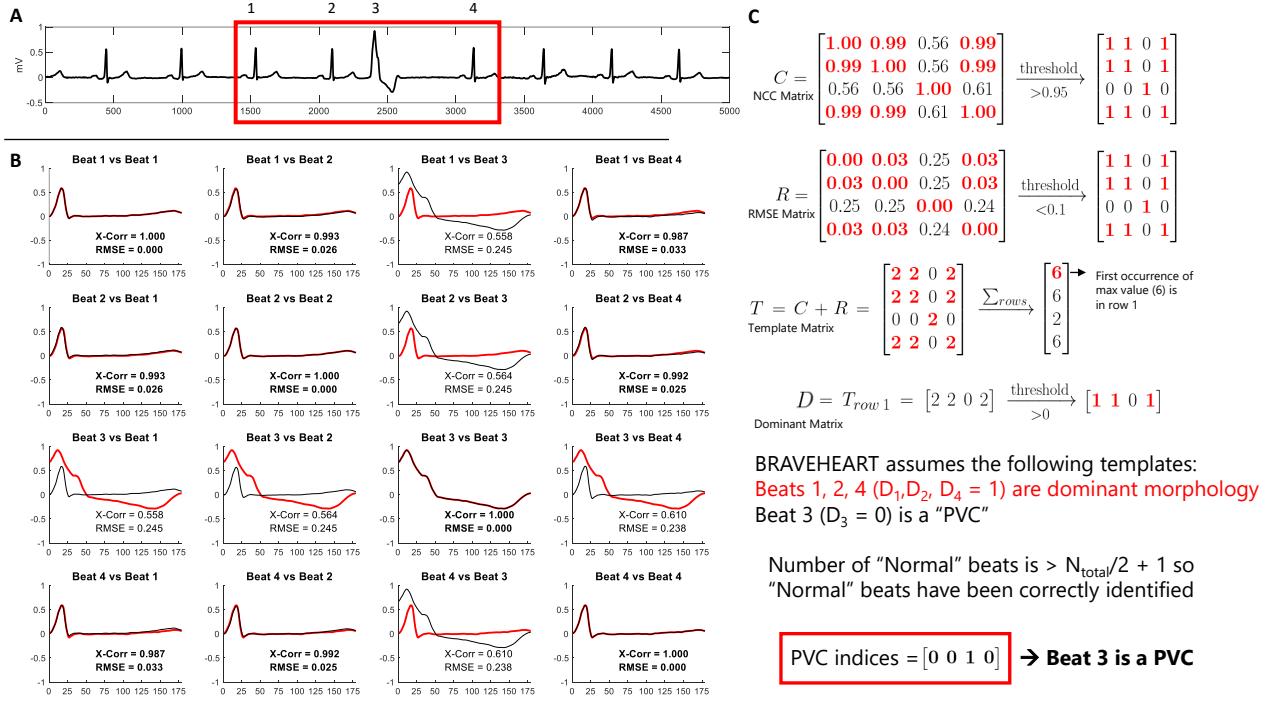


Figure 6: Example of how the BRAVEHEART PVC detection algorithm works on an ECG with a single PVC. **A:** A subset of 4 beats (3 sinus and 1 PVC) are analyzed for illustrative purposes, but the algorithm would have the same result if all 9 beats were analyzed. **B:** After trimming, each beat is aligned with the remaining beats, and normalized cross correlation (NCC) and root mean squared error (RMSE) are calculated. **C:** Values of NCC and RMSE are placed in matrices C and R , respectively, and then thresholds are applied as described in the text. The template matrix T is formed by adding $C + R$ and summing the rows. The dominant matrix D is formed by taking the first row of T that has the maximum row sum value, and then performing additional thresholding. BRAVEHEART starts by assuming that values of 1 in D are the dominant morphology and values of 0 in D are PVCs. In this case, since the number of beats with dominant morphology is $> n/2 + 1$, the algorithm is complete and the indices of PVCs are found by taking values of 0 in matrix D (beat 3).

Abbreviations: PVC - premature ventricular contraction.

2.9.1. Non-dominant Beat Detection Performance:

The performance of the BRAVEHEART PVC detector was assessed in a sample of 286 sequential ECGs obtained from persons with a history of PVCs, totaling 3,488 total beats and 508 PVCs/ventricularly paced/aberrant beats (14.6%). Not all ECGs contained PVCs. Ground truth was assigned by manual review of the ECGs and labeling beats as “dominant” or “PVC”. We excluded 26 beats where focal artifact/noise significantly distorted what was likely a normal beat to the degree that it was morphologically distinct enough that it would be reasonable to consider it a “PVC” and to remove it from the

analysis.

BRAVHEART’s PVC detection algorithm was then used to process this ECG dataset with various values of NCC and RMSE. Summary statistics, including sensitivity, specificity, positive predictive value, negative predictive value, F1 score, and overall accuracy were recorded. Performance was further assessed using bootstrapping using 1,000 replications to obtain confidence intervals for these measurements. PVC detection performance is presented in the Results Section (Section 3.1).

2.10. Outlier Detection/Removal:

After PVCs, paced beats, and aberrant beats are removed, remaining beats with artifact that preclude accurate fiducial point annotation are likewise detrimental to accurate median beat construction. Although most beats with enough artifact to significantly change their QRST morphology compared to the dominant beat morphology will be removed via PVC detection, some beats will retain the dominant QRST morphology but have enough artifact so that they cannot be aligned with other beats accurately. They may also introduce undesired noise into median beat construction. Noisy beats of the dominant morphology can also escape PVC detection if there is focal artifact towards the end of the T wave that makes annotation difficult or impossible because the PVC detection adjusts the ends of the QRST complexes to make them all equal length (see Supplemental Section 1.2).

To deal with this situation, BRAVEHEART has an algorithm to identify and remove outlier beats within the dominant QRST morphology. For each beat, the QR and RT intervals are computed. A modified Z-score (Z_i) for each of these quantities is then computed [25]. Details of modified Z-score calculation can be found in the “Outlier Detection Using Modified Z–Scores” section of the Supplemental Methods (Supplemental Section 1.3).

Nominally, any beat where the absolute value of Z_i for either the QR or RT interval is greater than 3.5 is marked as an outlier [25]. These outliers are nominally removed automatically. The modified Z-score threshold that reliably detects outliers depends on the total number of beats being compared and can be adjusted as needed; in cases of

bradycardia, it may be necessary to slightly increase the threshold to avoid tagging acceptable beats as outliers [25]. It is also possible to automatically adjust the modified Z-score threshold based on HR (and the number of beats in the ECG) for this purpose.

2.11. Median VCG Construction:

The n remaining beats, which should now be all of a single morphology and free of excessive noise/artifact are aligned to create a median beat which is subsequently used for all calculations and measurements. Although BRAVEHEART has the ability to align the individual QRST complexes based on their R peak locations, this method is, in general, less reliable as it is highly influenced by artifact and noise. To overcome this limitation and obtain more robust beat alignment, we calculate the “center of voltage” (CoV) of each of the remaining n QRS complexes in each lead in a manner analogous to calculating center of mass. For a given QRS complex, let S_i be sample i within it. The CoV of this QRS complex is then given by:

$$\text{CoV} = \frac{\sum_i i S_i}{\sum_i S_i} \quad (3)$$

For each of the n QRST complexes, the CoVs are then used as the fiducial point for alignment. The signal duration incorporated into the median beat prior to the CoV is the maximum Q-CoV interval of all beats + 40 ms and the signal duration incorporated into the median beat after the CoV is the maximum CoV-T interval of all beats + 60 ms. The window is expanded like this so that small errors in the prior heuristic annotation will not result in cutting off the true Q_{on} or T_{off} before being passed into the final median beat annotation algorithm (see Section 2.12 below). Once all beats are the same duration and aligned properly, the median beat is obtained by computing the median voltage value on these aligned samples (see Figure 7). Median X, Y, and Z leads are computed. The median VM lead is then calculated from the median X, Y, and Z leads as previously described in Equation 2.

2.12. Median Beat Annotation:

Accurate annotation of the median beat is critical for accurate measurement of ECG/VCG parameters. We found that the same heuristic methods for ECG annotation

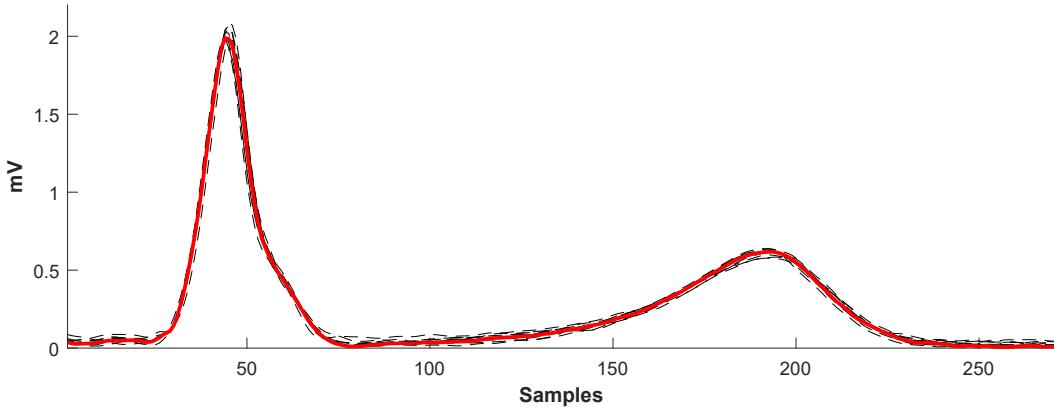


Figure 7: Example of median beat construction. Individual QRST complexes (black dashed lines) are aligned on their center of voltage, and the median of the aligned samples is calculated to obtain the median beat (red).

used in first pass annotation would often fail at annotating the median beat with enough accuracy for this purpose, especially in cases with atypical QRS complex morphology or when T waves were low amplitude or of abnormal morphology. We therefore created a custom bidirectional long short term memory neural network (NN) specifically for BRAVEHEART to accurately annotate fiducial points on VM median beat QRST complexes with a high degree of accuracy.

Details of NN architecture, training, and testing have previously been reported [26], and are also available in the Supplemental Methods Sections 1.4, 1.5, and 1.6. In brief, the NN takes in a variable length VM median beat signal and outputs probabilities of each sample of the signal being a “QRS complex”, “T wave”, or “other”. Transitions between “other” to “QRS complex” are taken as the location of Q_{on} , transitions between “QRS complex” to “T wave” are taken as the location of Q_{off} , and transitions between “T wave” to “other” are taken as the location of T_{off} . See Figure 8 for an illustration of this process. In rare cases where the NN output results in multiple locations where the definition for one of the fiducial points was satisfied, logic based on the physiologic properties and order of QRST signals attempts to identify the correct fiducial point location.

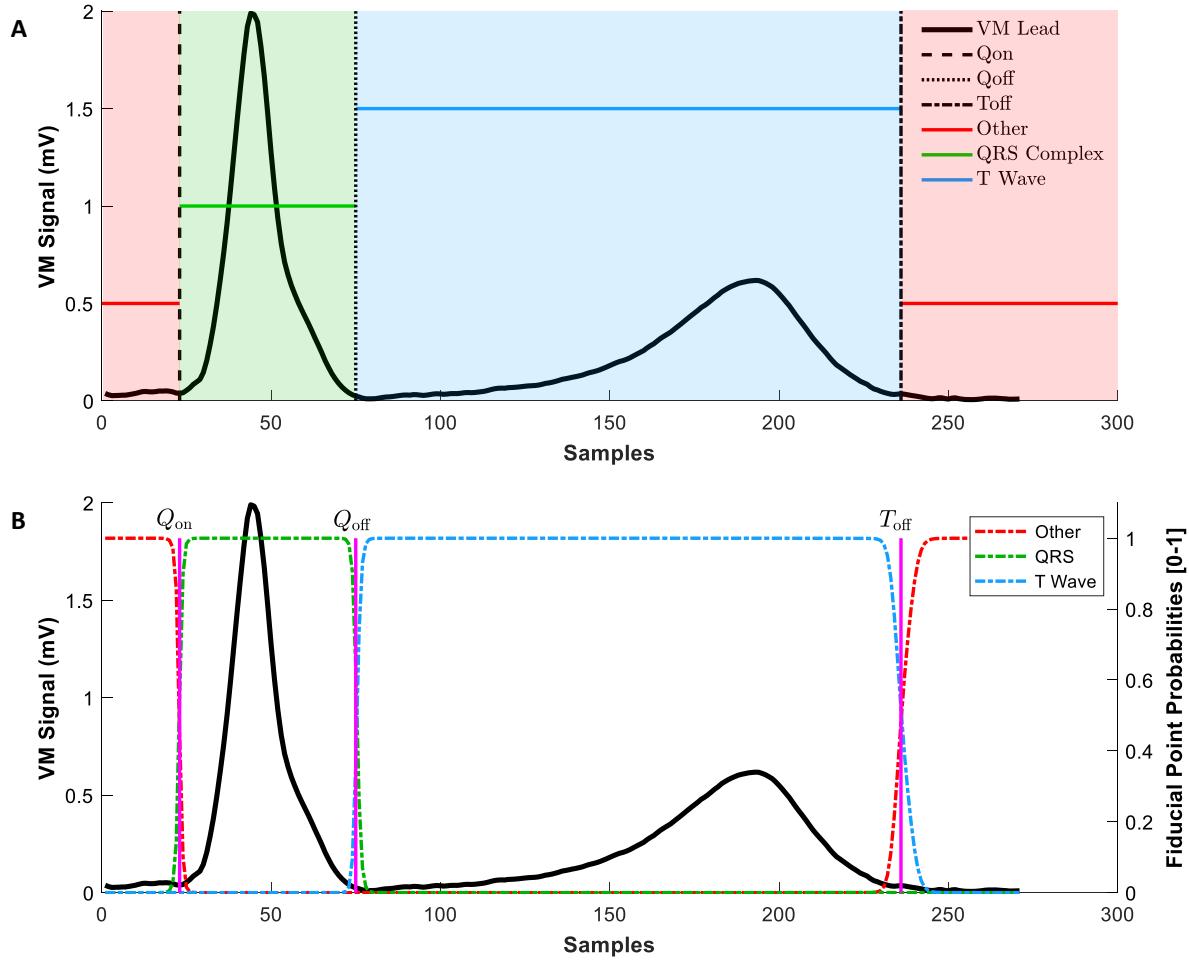


Figure 8: Signal processing and labeling for NN training/output. **A:** For NN training, 12-Lead ECGs were transformed to VCGs and the median VM beat was constructed (black signal). Q_{on} , Q_{off} , and T_{off} were annotated by a cardiac electrophysiologist as shown (black dashed, dotted, and dash-dotted lines, respectively). A categorical signal (red, green, and blue lines/shading) was constructed based on the location of Q_{on} , Q_{off} , and T_{off} . The QRS complex was labeled between Q_{on} and Q_{off} (green). The T wave was labeled between Q_{off} and T_{off} (blue). The remainder of the signal was labeled as “other” (red). The labeled categorical signal was used for NN training/testing. See the text for details. **B:** Example of NN output. The probabilities of “QRS complex” (green), “T wave” (blue), and “other” (red) labels are shown relative to the median VM beat signal and the predicted fiducial points which are assigned at transitions between fiducial point probabilities as described in the text.

Abbreviations: ECG - electrocardiogram, VCG - vectorcardiogram, VM - vector magnitude, NN - neural network, Q_{on} - QRS complex onset, Q_{off} - QRS complex end, T_{off} - T wave end.

2.12.1. Median Beat Neural Network Annotation Performance:

NN performance was assessed primarily based on calculating the mean and standard deviation of difference between NN predicted locations of Q_{on} , Q_{off} , and T_{off} and measured QRS duration and QT interval, and ground truth annotations performed by a board

certified cardiac electrophysiologist (JWW). Given that there were 3 output classes, micro-averaged F1 was reported. As this was a multi-class labeling problem, the recall (specificity), precision (positive predictive value), and F1 (the harmonic mean of precision and recall) – see Supplemental Statistical Methods (Supplement Section 1.8) are all equal, and therefore only F1 was reported.

The included NN was tested on 189 median beats which were randomly allocated for testing during NN training. Given that BRAVEHEART users care most about the performance of the included weights and biases rather than overall theoretical model performance which has previously been described [26], we also validated NN performance on a second independent dataset of 200 sequentially obtained and manually annotated median beats. Details of NN performance is described in the Results Section 3.2.

2.13. ECG/VCG Parameter Calculations:

BRAVEHEART has easily customizable modules that calculate a variety of ECG and VCG measurements on the median beat. Parameters include standard measurements such as HR, QRS duration, and QT interval, morphological measurements for each of the 16 leads such as R wave, S wave, and T wave magnitude (see Figure 9), and a variety of vectorcardiographic parameters including spatial QRS-T angle [5], TCRT [6], and peak and area vectors for the QRS complex, T wave, and entire QRST complex (spatial ventricular gradient [3]). VCG speed, loop coplanarity (how well the loops fit into a best fit plane), roundness (how circular vs oval), length, and area are also calculated. Further details and equations for calculated parameters are available in the “Equations” section of the Supplemental Methods and Supplemental Tables 1, 2, and and 3.

The software is designed to facilitate easy addition of new parameters as needed; new parameters and the calculations for these parameters can be added to one of the existing result classes, and the software automatically adjusts the output files to accommodate the new variables. Adding a new parameter requires minimal extra code beyond that used to compute the result (further details can be found in the BRAVEHEART User Guide which is available on the BRAVEHEART GitHub).

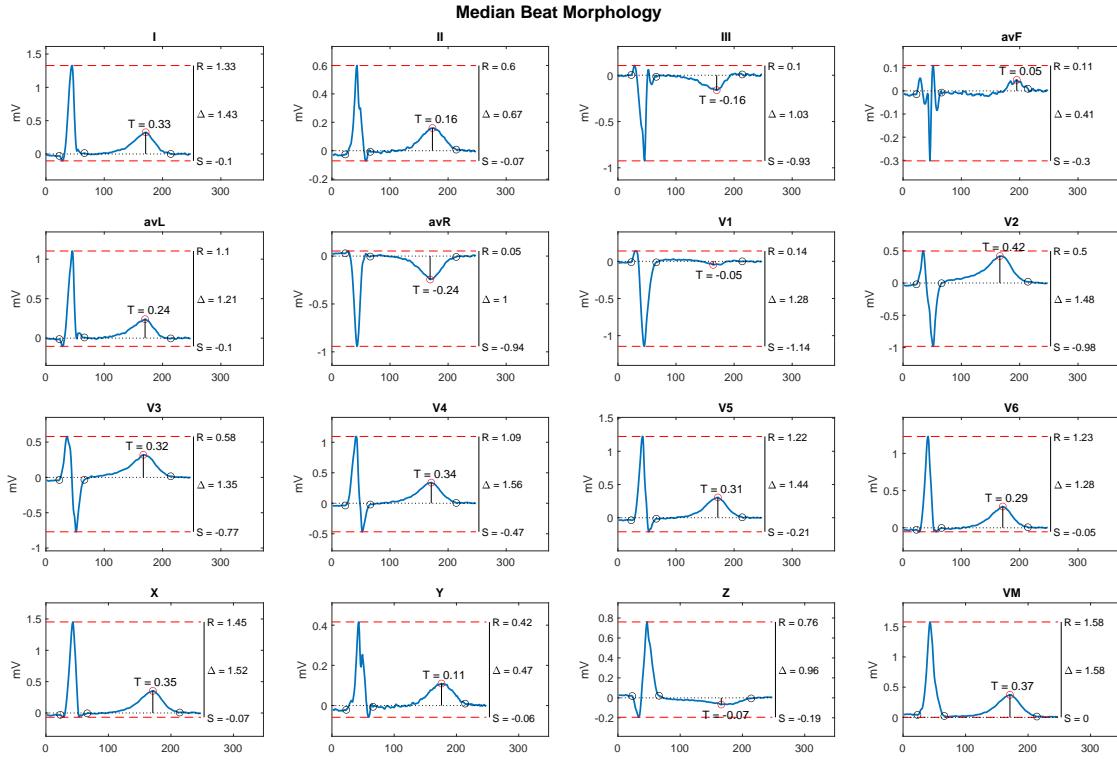


Figure 9: Lead morphology measurements including the magnitude of R, S, and T waves for all 16 leads are shown.

2.14. VCG Signal and Annotation Quality Checks:

BRAVEHEART was designed for batch processing of large numbers of ECGs with minimal need for human oversight and without the need to manually review every ECG that is processed for signal quality, beat labeling, or removal of PVCs/outliers. The program can also take advantage of MATLAB's parallel processing abilities which can significantly speed up processing of large ECG datasets. As ECG quality varies, especially when using clinical ECGs that were not collected specifically for research purposes, some lower quality ECG recordings are more likely to have errors in fiducial point annotation or median beat construction, usually due to significant artifact. In large studies it may not always be feasible to manually review every ECG that is processed, and we therefore designed ways of highlighting ECGs that are most likely to be problematic and require

either exclusion from the dataset due to overall poor signal quality, automatic reannotation with other annotation parameters, or in very rare cases, manual ECG processing.

Further details of the quality checking methodology can be found in the “Quality Labeling Methods” section of the Supplemental Methods (Supplemental Section 1.7).

Briefly, annotated median beats are automatically checked for possible processing errors or missannotation by checking a variety of metrics including the QT interval, QRS duration, T peak to QT ratio, T wave magnitude, average NCC between beats that make up the median beat, HR, number of beats included in the median beat construction, number of beats removed during PVC and outlier exclusion, presence of likely missing leads, and low neural network annotation confidence. If any of these values are outside of a nominal range (which can be easily adjusted via an external file as needed and based on the characteristics of the ECGs being processed), then the ECG is flagged for review. To streamline the review process, a separate file listing these flagged ECGs is generated, and figures from flagged ECGs can be automatically saved to a separate folder to facilitate review of flagged ECGs without needing to review the output of the entire dataset.

We also created a logistic regression based on a set of 481 sequentially processed ECGs/median beats which were manually labeled as “good” quality or “needs review” which can be used to predict the probability that the ECG being processed requires manual review (See Results Section 3.3, Supplemental Methods Section 1.7, Supplemental Results Section 2.1, Figure 10, and Supplemental Figure 4). ECGs that have a low probability of being “good” quality (based on an adjustable cutoff) can be flagged for review. This quality checking system is nominally designed to be specific, in that ECGs that have no flags are very unlikely to have any quality issues. Some ECGs that are flagged, however, may be perfectly fine after manual review or a minor adjustment. The thresholds for flagging ECGs can be adjusted by the user as needed if more specificity or more sensitivity is needed for a specific project (see Figure 10) and based on the types of ECGs being processed (e.g. are there frequent PVCs that will be removed or are all ECGs bradycardic), or the overall quality of the ECG recordings. This quality checking system can also be used to select ECGs with specific features (such as QRS duration > 150 ms or large amplitude T waves) from an unlabeled ECG dataset.

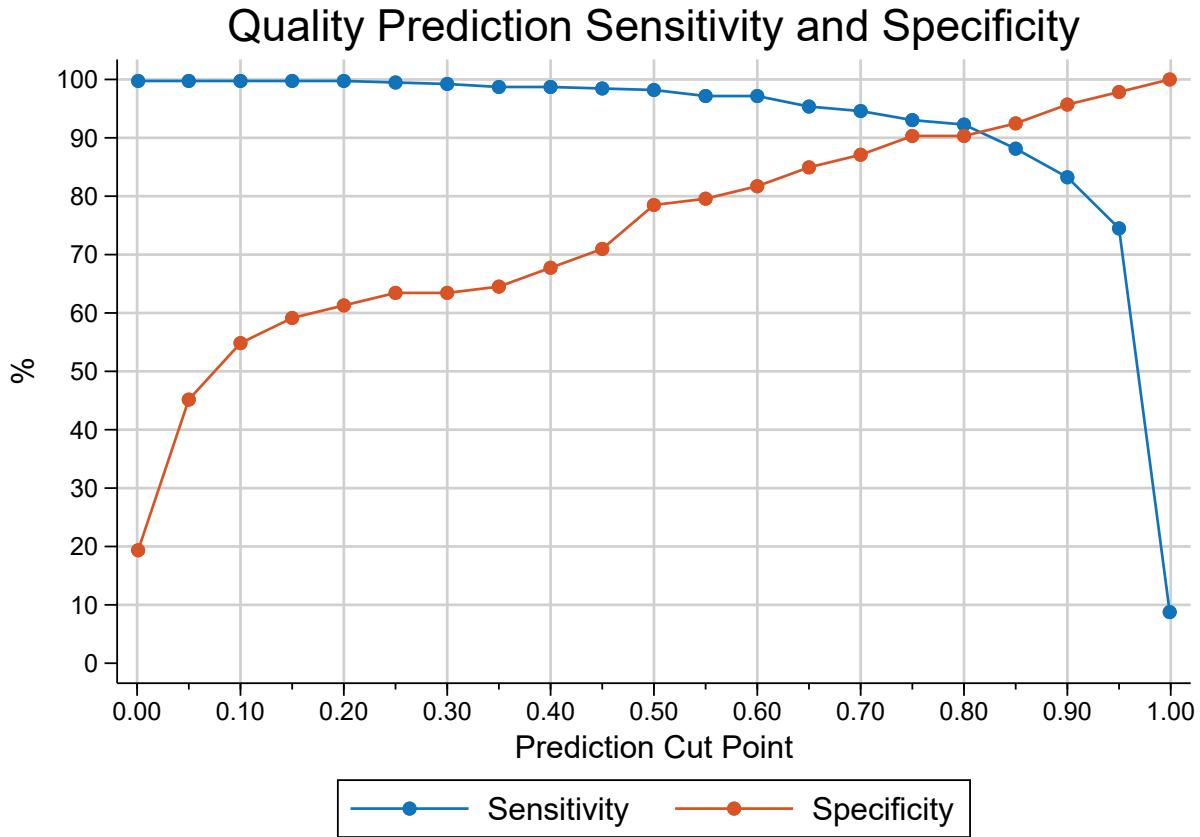


Figure 10: Sensitivity and specificity of the quality regression for predicting “good” quality ECGs. A cut point of 0.8 was set as nominal based on the relatively high and equal sensitivity and specificity at this point. No ECGs had a predicted probability of exactly 0 or exactly 1.00, so the point near “0.00” on the X-axis is for a cut point of 0.01, and the point near “1.00” on the X-axis is for a cut point of 0.99.

2.15. Graphical Display of VCG Median Beats and Export of Measurement Data:

The BRAVEHEART software pipeline automatically produces a figure illustrating and summarizing the annotation process for each ECG/VCG, and example summary figures are shown in Figure 11 and Supplementary Figure 5. These summary figures can be quickly inspected to determine the number of beats included/excluded, why beats were excluded (PVC detection, outlier detection, manually removed, or missed), the accuracy and quality of median beat construction (via average NCC for all of the beats that made up the X, Y, and Z median beats), and the location of the median beat fiducial points. A detailed graphical user interface (GUI) that allows more specific adjustments and multiple

additional visualizations is also available (see below).

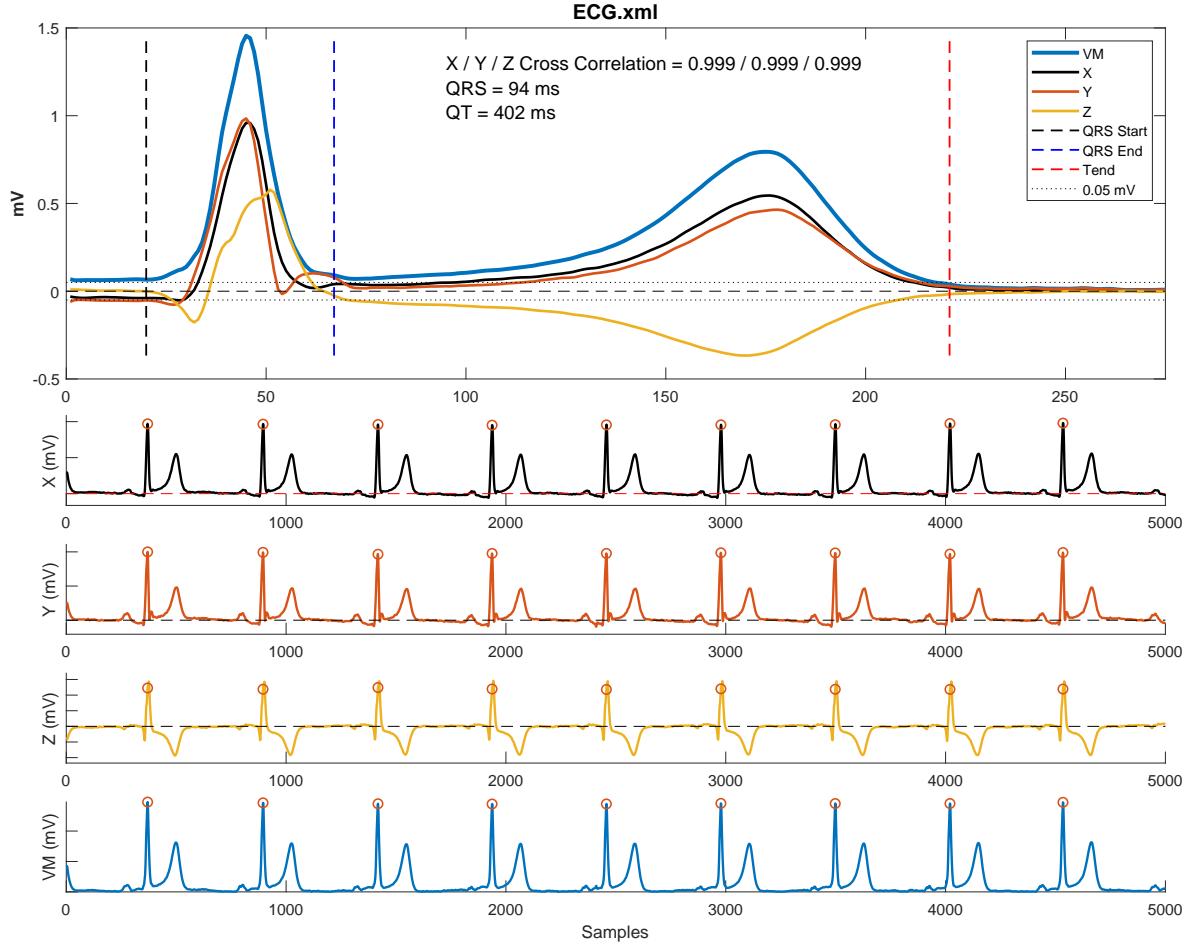


Figure 11: Example of summary figure output. The X, Y, Z, and VM median beats and fiducial point annotations are shown in the top panel, and the full VCG is shown in the bottom half of the figure. Beats included in the median beat have an orange circle at their R peak. Beats excluded from the median beat (none in this example) would be annotated with the reason for exclusion (PVC or outlier). Cross correlation represents the average normalized cross correlation between all pairs of beats (in the X, Y, and Z leads) that make up the median beat and represents the quality of median beat construction with values very close to 1 indicating excellent beat alignment.

Abbreviations: ECG - electrocardiogram, VCG - vectorcardiogram, VM - vector magnitude,

Measurement data for the processed ECG/VCG, which include the ECG filtering/processing settings used, are then exported to a file (.csv or .xlsx). ECGs processed in the same batch output their data to a single file for further analysis. The software can also save the ECG, VCG, and median beat signal data and ECG annotations to a separate file for additional processing/analysis outside of BRAVHEART.

The BRAVEHEART GUI (Figure 12) allows granular control of each step in ECG processing, including filtering, baseline correction, VCG construction, and beat removal, and can be quite helpful for troubleshooting ECGs which result in errors, or to better visualize each step in ECG/VCG processing. The GUI also allows visualization of different steps in ECG/VCG processing and additional figures, including standard 12-lead ECGs and VCGs, rotatable 3-dimensional VCG loops (Supplemental Figure 6), ECG/VCG lead morphology (Figure 9), and various figures describing the filtering and signal quality.

2.16. Software Formats:

BRAVEHEART is implemented in MATLAB and requires a version of at least R2022a. We have provided 2 packages that will suit different users. `braveheart_batch.m` is a command line driven program that is designed for rapid, batch processing of ECGs with no graphical interface. Parameters are set in an external file, a directory is chosen for source ECG files, and the program processes the ECGs. The summary figure (Figure 11 and Supplemental Figure 5), ECG, VCG, and median beat signal data stored as .mat files, and annotation files (.anno) which contain the parameters used for ECG processing and individual beat fiducial point annotations for each processed ECG can be saved if specified by the user. ECGs that are not successfully processed by the completely automatic method and nominal settings can have their parameters manually adjusted, and individual beats can be removed/edited without the need for a GUI by editing an annotation (.anno) file that is placed in the same directory as the corresponding ECG. If the program notes the presence of an .anno file with the same name as the ECG being processed, the annotation parameters and fiducial point locations (if included) are read from the .anno file for that specific ECG. This allows the user control over ECGs that require non-nominal settings without requiring the use of the GUI.

We also provide a detailed GUI (`braveheart_gui.m`) (Figure 12) that can perform both batch and individual ECG processing. The GUI is designed to give the user very fine control over each step of ECG processing including very granular control over beat annotation and removal. The GUI also allows additional visualizations of both ECG/VCG signal data and resulting measurements. A detailed overview of the GUI is available in the

User Guide which is available on the BRAVEHEART GitHub repository. For users who do not have access to MATLAB, we have provided executable files for both the command line and GUI versions of BRAVEHEART for both Windows and Mac operating systems.

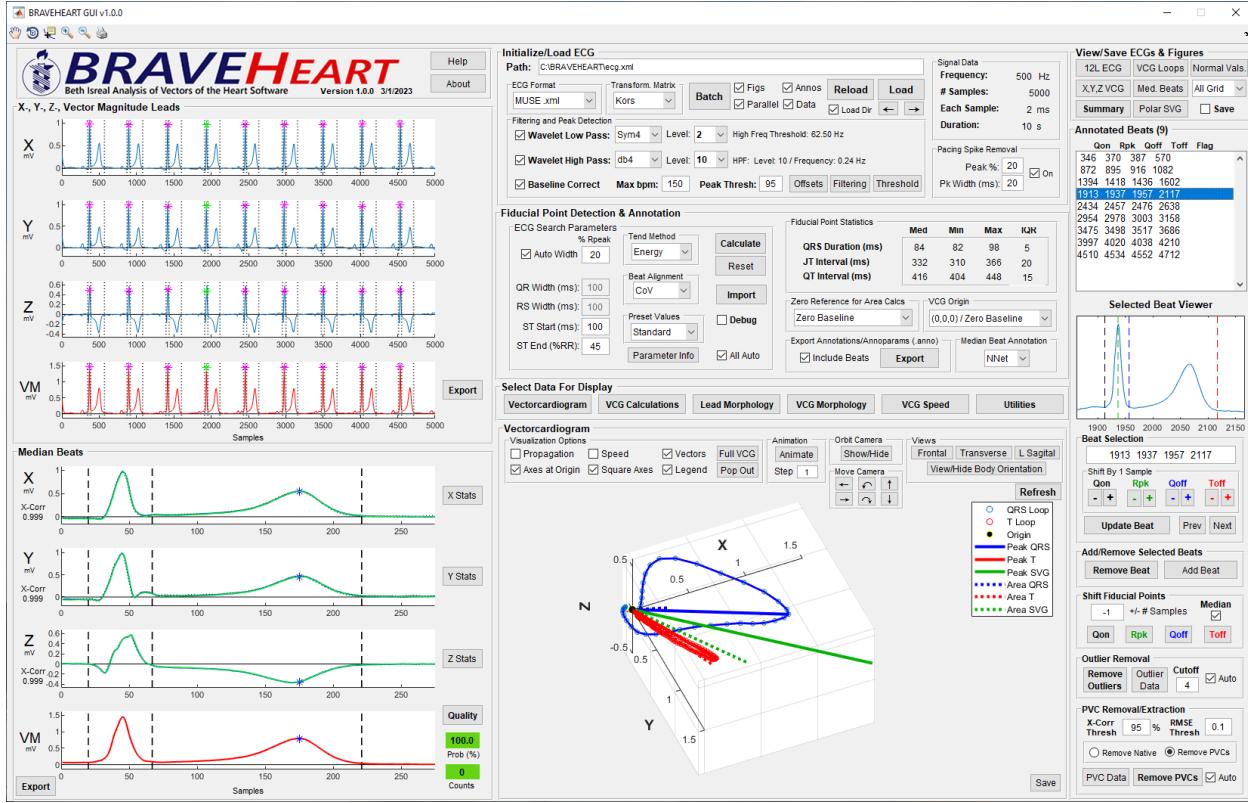


Figure 12: BRAVEHEART GUI. Further details on using the GUI can be found in the online User Guide.

The complete BRAVEHEART MATLAB source code, executable files for Windows and Mac, and a detailed user guide, are available for download from GitHub

<http://github.com/BIVectors/BRAVEHEART> under version 3 of the GPL license.

BRAVEHEART uses the MATLAB signal processing, wavelet, and deep learning toolboxes. The parallel computing toolbox is optional to improve the speed of processing large batches of ECGs in parallel.

2.17. Statistical Analysis:

All ECG processing and NN training were performed using MATLAB 2020b (Mathworks, Natick, MA, USA). Statistical analysis and data manipulation were performed using Stata 17 (StataCorp, College Station, TX, USA). The study was approved by the Institutional Review Board of Beth Israel Deaconess Medical Center.

3. Results

3.1. PVC Detection Performance:

Supplemental Table 4 shows performance of the BRAVEHEART non-dominant QRST morphology (PVC) detector using various cutoffs of NCC and RMSE based on the summary performance of 1,000 replications of bootstrapping. In general, the best performance was found using a NCC of $\geq 95\%$, and a RMSE = 0.1, which is why these values were chosen for nominal settings. Performance was also excellent when NCC was disabled, and decisions on similarity were based on RMSE alone. With NCC = 95% and RMSE = 0.1, sensitivity was 98.2%, specificity was 99.5%, and F1 was 97.7% for PVC detection. Although PVC detection performance was excellent using a RMSE cutoff of 0.1 with or without use of NCC in certain cases it can be advantageous to utilize both NCC and RMSE, as a beat with excellent matches by both NCC and RMSE may be a better template beat than if only matched by RMSE (See Supplemental Figure 12).

3.2. Neural Network Performance:

Supplemental Table 5 shows the features/diagnoses of ECGs used for NN training and testing during NN development. There were overall no significant differences in ECG features between the training and testing datasets with the exception of a borderline lower rate of left bundle branch block ECGs in the testing dataset (3.2 vs 7.1 %, $p=0.046$). Table 1, Figure 13 and Supplemental Figures 7, 9, and 8 show the accuracy of NN fiducial point annotations when compared to ground truth human annotations for the 189 ECGs set aside during NN training. There was excellent agreement between NN predicted and ground truth fiducial point locations (Q_{on} , Q_{off} , and T_{off}) and ECG intervals (QRS duration and QT interval) with mean differences $< \pm 2$ ms for all fiducial points/intervals, and all measurement error distributions were well within established acceptable limits for automated ECG annotation [27, 28, 29, 30]. Microaveraged F1 (μ F1 - See Supplemental Statistical Methods in Supplement Section 1.8) was 97.95%. The NN did not miss locating any fiducial points, and only 1 ECG was found to have more an extra fiducial point (1 T_{off}) detected which was appropriately dealt with using logic that accounts for the order of

fiducial points. Supplemental Figure 7 shows intraclass correlation coefficients for all fiducial points and intervals were very close to 1.0. Supplemental Figure 8 shows the percent of annotations that were within different intervals of accuracy; more than 99% of Q_{on} and Q_{off} annotations and more than 98% of T_{off} annotations were within 20 ms of ground truth. Bland-Altman plots for NN performance are available in Supplemental Figure 9.

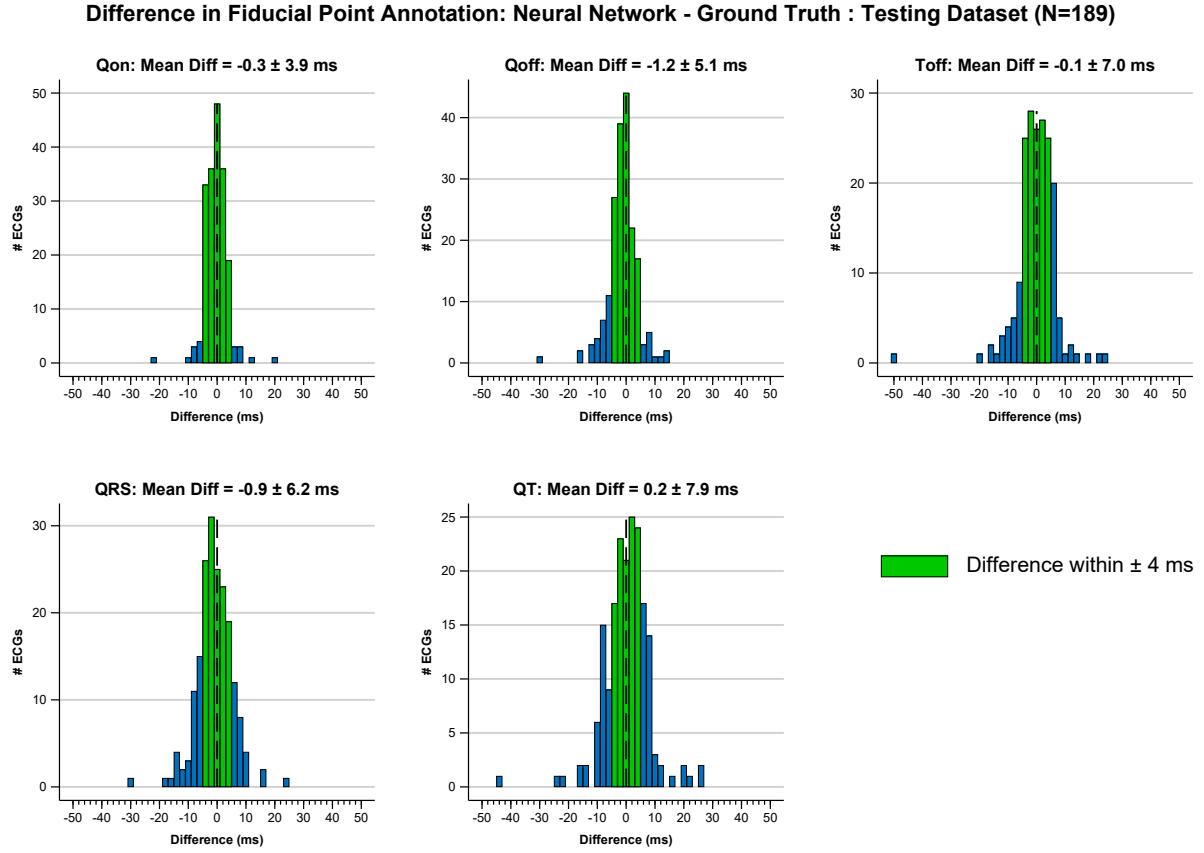


Figure 13: Mean and standard deviation and histograms for results of NN testing for the initial 189 ECG testing dataset. Bins where the difference between NN predictions and ground truth fiducial points were $\leq \pm 4$ ms are shown in green.

Supplemental Table 6 shows the features/diagnoses of ECGs included in the second, independent testing dataset of 200 sequentially obtained ECGs. Table 2 shows that the NN performed similarly on this second dataset, with mean differences $< \pm 2$ ms for all fiducial points/intervals. More than 99% of all fiducial points and intervals were within 20 ms of ground truth. $\mu F1$ was 98.03%. The NN did not miss locating any fiducial points, and only

Table 1: Results from NN testing using the initial 189 ECG testing dataset.

Result	Q_{on}	Q_{off}	T_{off}	QRS duration	QT interval
Mean ± SD (ms)	-0.3 ± 3.9	-1.2 ± 5.1	-0.1 ± 7.0	-0.9 ± 6.2	0.2 ± 7.9
Missed Points	0	0	0	N/A	N/A
Extra Points	0	0	1	N/A	N/A
% Annotations ± 10 ms	98.4	95.2	92.6	93.7	92.1
% Annotations ± 20 ms	99.5	99.5	98.4	98.9	96.8

Table 2: Results from NN testing using the second 200 ECG testing dataset.

Result	Q_{on}	Q_{off}	T_{off}	QRS duration	QT interval
Mean ± SD (ms)	-0.1 ± 3.3	-1.7 ± 4.0	1.2 ± 5.8	-1.6 ± 5.4	1.3 ± 6.7
Missed Points	0	0	0	N/A	N/A
Extra Points	0	0	1	N/A	N/A
% Annotations ± 10 ms	99.5	98.5	93.5	94.0	89.5
% Annotations ± 20 ms	100.0	100.0	99.0	100.0	99.5

1 ECG was found to have an extra fiducial point (1 T_{off}) detected which was appropriately dealt with using logic that accounts for the order of fiducial points. Figures summarizing the performance of the second dataset are shown in Supplemental Figures 10 and 11.

3.3. Quality Labeling Results:

Figure 10 and Supplemental Figure 4 show the results for identifying an ECG/median beat as “good quality” or “needs review” using the previously discussed logistic regression (See Supplemental Sections 1.7 and 2.1). The cutoff used to distinguish “good quality” and “needs review” ECGs can be adjusted as needed for a specific project to increase/decrease sensitivity or specificity. In general, a cutoff of ~80% tends to have a good mix of sensitivity and specificity, with both values >90%. Cutoff values >90% have higher specificity at the expense of lower sensitivity, and cutoff values <70% have higher sensitivity at the expense of lower specificity. The ROC curve for the quality regression is shown in Supplemental Figure 4 and had an area under the ROC of 0.973 indicating excellent agreement between prediction and manual labeling. Further details can be found in the Supplemental Methods/Results (Supplemental Sections 1.7 and 2.1).

4. Discussion

We present BRAVEHEART, an easy to use and easily customizable research software package for ECG and VCG analysis. BRAVEHEART satisfies an important need in the development of ECG/VCG research by providing open-source software that performs all signal processing and analyses itself without the need for separate programs to process or annotate ECG signals. The software accepts a wide variety of ECG formats, and is easily customizable with minimal extra programming. We also have provided executable versions of the software for those without access to MATLAB. Extensive documentation is available in a user guide to help users familiarize themselves with the software. Although other ECG/VCG processing software does exist [8, 9, 12, 13, 10, 11], these other software packages either are not publically available, not open-source, or provide partial functionality without allowing complete processing from ECG to VCG and final measurements. We are aware of the importance of being able to modify the software and add additional measurements based on specific research needs, and have taken care to make it easy to add new measurements with minimal effort and coding (see the online user guide).

Pre-release versions of BRAVHEART have been used in research investigating associations between VCG parameters (primarily the SVG) and drug induced ventricular arrhythmias [31], antiarrhythmic drug administration [20], post pulmonary embolism risk stratification [32], and chemotherapy-induced cardiotoxicity [33], and the current release version of BRAVEHEART is being used for ECG/VCG analysis in multiple ongoing studies. We hope that members of the community will help improve BRAVEHEART over time by suggesting or adding new or improved features.

5. Limitations/Future Directions

BRAVEHEART currently does not support P wave annotation. We plan to add this as a feature in the near future.

6. Conclusion

BRAVEHEART is an open-source, easily customizable ECG/VCG analysis software package that reproducibly processes digital 12-lead ECG files without the need for external software libraries or signal pre-processing. Source code, executables, and a detailed user guide are available at <http://github.com/BIVectors/BRAVEHEART> where it is distributed under the General Public License (GPL) version 3.

References

- [1] J. Schläpfer and H. J. Wellens. Computer-Interpreted Electrocardiograms: Benefits and Limitations. *J Am Coll Cardiol*, 70(9):1183–1192, Aug 2017.
- [2] J. Vondrak and M. Penhakert. Review of Processing Pathological Vectorcardiographic Records for the Detection of Heart Disease. *Front. Physiol*, March 21, 2022.
- [3] J. W. Waks and L. G. Tereshchenko. Global electrical heterogeneity: A review of the spatial ventricular gradient. *J Electrocardiol*, 49(6):824–830, 2016.
- [4] L. G. Tereshchenko, A. Cheng, B. J. Fetters, B. Butcher, J. E. Marine, D. D. Spragg, S. Sinha, D. Dalal, H. Calkins, G. F. Tomaselli, and R. D. Berger. A new electrocardiogram marker to identify patients at low risk for ventricular tachyarrhythmias: sum magnitude of the absolute QRST integral. *J Electrocardiol*, 44(2):208–216, 2011.
- [5] A. Oehler, T. Feldman, C. A. Henrikson, and L. G. Tereshchenko. QRS-T angle: a review. *Ann Noninvasive Electrocardiol*, 19(6):534–542, Nov 2014.
- [6] B. Acar, G. Yi, K. Hnatkova, and M. Malik. Spatial, temporal and waveform direction characteristics of 12-lead T-wave morphology. *Med Biol Eng Comput*, 37(5):574–584, Sep 1999.
- [7] G. Sedaghat, E. Ghafoori, J. W. Waks, M. M. Kabir, A. Shvilkin, M. E. Josephson, and L. G. Tereshchenko. Quantitative Assessment of Vectorcardiographic Loop Morphology. *J Electrocardiol*, 49(2):154–163, 2016.
- [8] Lars Johannessen, Jose Vicente, Loriano Galeotti, and David G. Strauss. ECGlib: Library for processing electrocardiograms. *Computing in Cardiology*, 40:951–954, 2013.
- [9] Jose Vicente, Lars Johannessen, Loriano Galeotti, and David G. Strauss. ECGlab: User friendly ECG/VCG analysis tool for research environments. *Computing in Cardiology*, 40:775–778, 2013.
- [10] H. H.M. Draisma, C. A. Swenne, H. Van De Vooren, A. C. Maan, B. Hooft Van Huysduynen, E. E. Van Der Wall, and M. J. Schalij. LEADS: An interactive research oriented ECG/VCG analysis system. *Computers in Cardiology*, 32(June 2014):515–518, 2005.
- [11] Sum-Che Man, Arie C Maan, Ernst E van der Wall, Martin J Schalij, and Cees A Swenne. Beats: An interactive research oriented ecg analysis system. In *2010 Computing in Cardiology*, pages 1007–1010, 2010.
- [12] P. Hamilton. Open source ecg analysis. In *Computers in Cardiology*, pages 101–104, 2002.

- [13] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23):E215–220, Jun 2000.
- [14] M. Kabir, E. A. Perez-Alday, K. Haq, J. Waks, and L. Tereshchenko. Measurement of Global Electrical Heterogeneity. *Physionet*, <https://physionet.org/content/geh/1.0.0/>, 2018.
- [15] National Academies of Sciences, Engineering, and Medicine. *Reproducibility and Replicability in Science*. The National Academies Press, Washington, DC, 2019.
- [16] N. V. Thakor, J. G. Webster, and W. J. Tompkins. Estimation of QRS complex power spectra for design of a QRS filter. *IEEE Trans Biomed Eng*, 31(11):702–706, Nov 1984.
- [17] J. A. Kors, G. van Herpen, A. C. Sittig, and J. H. van Bemmel. Reconstruction of the Frank vectorcardiogram from standard electrocardiographic leads: diagnostic comparison of different methods. *Eur Heart J*, 11(12):1083–1092, Dec 1990.
- [18] G. E. Dower, H. B. Machado, and J. A. Osborne. On deriving the electrocardiogram from vectorradiographic leads. *Clin Cardiol*, 3(2):87–95, Apr 1980.
- [19] Paul Kligfield, Leonard S. Gettes, James J. Bailey, Rory Childers, Barbara J. Deal, E. William Hancock, Gerard van Herpen, Jan A. Kors, Peter Macfarlane, David M. Mirvis, Olle Pahlm, Pentti Rautaharju, and Galen S. Wagner. Recommendations for the standardization and interpretation of the electrocardiogram. *Circulation*, 115(10):1306–1324, 2007.
- [20] H. F. Stabenau, C. Shen, L. G. Tereshchenko, and J. W. Waks. Changes in global electrical heterogeneity associated with dofetilide, quinidine, ranolazine, and verapamil. *Heart Rhythm*, 17(3):460–467, 03 2020.
- [21] E. A. Perez-Alday, Y. Li-Pershing, A. Bender, C. Hamilton, J. A. Thomas, K. Johnson, T. L. Lee, R. Gonzales, A. Li, K. Newton, and L. G. Tereshchenko. Importance of the heart vector origin point definition for an ECG analysis: The Atherosclerosis Risk in Communities (ARIC) study. *Comput Biol Med*, 104:127–138, Jan 2019.
- [22] J. L. Willems, C. Zywietz, P. Arnaud, J. H. van Bemmel, R. Degani, and P. W. Macfarlane. Influence of noise on wave boundary recognition by ECG measurement programs. Recommendations for preprocessing. *Comput Biomed Res*, 20(6):543–562, Dec 1987.
- [23] P. G. Postema and A. A. Wilde. The measurement of the QT interval. *Curr Cardiol Rev*, 10(3):287–294, Aug 2014.

- [24] L. Johannessen, J. Vicente, M. Hosseini, and D. G. Strauss. Automated Algorithm for J-Tpeak and Tpeak-Tend Assessment of Drug-Induced Proarrhythmia Risk. *PLoS One*, 11(12):e0166925, 2016.
- [25] B. Iglewicz and D. Hoaglin. *The ASQC Basic References in Quality Control: Statistical Techniques*, chapter Volume 16: How to Detect and Handle Outliers. 1993.
- [26] H. F. Stabenau, C. P. Bridge, and J. W. Waks. ECGAug: A novel method of generating augmented annotated electrocardiogram QRST complexes and rhythm strips. *Comput Biol Med*, 134:104408, 07 2021.
- [27] Recommendations for measurement standards in quantitative electrocardiography. The CSE Working Party. *Eur Heart J*, 6(10):815–825, Oct 1985.
- [28] S. H. Zhou, E. D. Helfenbein, J. M. Lindauer, R. E. Gregg, and D. Q. Feild. Philips QT interval measurement algorithms for diagnostic, ambulatory, and patient monitoring ECG applications. *Ann Noninvasive Electrocardiol*, 14 Suppl 1:3–8, Jan 2009.
- [29] J. L. Willems, P. Arnaud, J. H. van Bemmel, P. J. Bourdillon, C. Brohet, S. Dalla Volta, J. D. Andersen, R. Degani, B. Denis, and M. Demeester. Assessment of the performance of electrocardiographic computer programs with the use of a reference data base. *Circulation*, 71(3):523–534, Mar 1985.
- [30] C. Zywiertz and D. Celikag. Testing results and derivation of minimum performance criteria for computerized ecg-analysis. In *[1991] Proceedings Computers in Cardiology*, pages 97–100, 1991.
- [31] H. F. Stabenau, C. Shen, P. Zimetbaum, A. E. Buxton, L. G. Tereshchenko, and J. W. Waks. Global electrical heterogeneity associated with drug-induced torsades de pointes. *Heart Rhythm*, 18(1):57–62, 01 2021.
- [32] H. F. Stabenau, M. Marcus, J. D. Matos, I. McCormick, D. Litmanovich, W. J. Manning, B. J. Carroll, and J. W. Waks. The spatial ventricular gradient is associated with adverse outcomes in acute pulmonary embolism. *Ann Noninvasive Electrocardiol*, page e13041, Jan 2023.
- [33] Andrea Nathalie Rosas Diaz, Hans Friedrich Stabenau, Gabriel Pajares Hurtado, Sarah Warack, Jonathan W. Waks, and Aarti Asnani. The spatial ventricular gradient is an independent predictor of anthracycline-associated cardiotoxicity. *JACC: Advances*, 2(2):100269, 2023.

Supplement to BRAVEHEART: Open-source software for automated electrocardiographic and vectorcardiographic analysis

Hans Friedrich Stabenau, MD, PhD and Jonathan W. Waks, MD
Harvard-Thorndike Electrophysiology Institute
Beth Israel Deaconess Medical Center
Harvard Medical School, Boston, MA

1. Supplemental Methods

1.1. Heuristic Annotation Details:

Once the QRS search window is established, for each QRS complex, the start (Q_{on}) and end (Q_{off}) are determined with the following method. First, the algorithm searches to determine if there is an additional peak in the VM lead before the dominant R wave peak within a short distance (i.e. a Q wave – all peaks are positive in the VM lead). Two methods are then used to determine the start of the peak. Method A finds the maximum slope of the R peak and walks forward/back until it finds a point where the absolute value of the derivative is 2% of the maximum slope (or the smallest value found if not < 2%). Method B looks for the first local minimum and labels this as the fiducial point. The point closest to dominant VM R peak from methods A and B is chosen as the fiducial point.

After Q_{on} and Q_{off} are located, a blanking window (nominally 100 ms) is set after the location of Q_{off} during which the T wave end (T_{off}) cannot be detected. At the end of this blanking window, the T_{off} search window starts and is extended forward nominally by 45% of the mean RR interval. The T_{off} search window is set as a % of the mean RR interval rather than an absolute value in ms to minimize the need to change parameter values as the QT interval is related to the RR interval/HR. The location of T_{off} is then detected by the chosen method as noted in the main manuscript. Fiducial point search windows are illustrated in Figure 4 in the main manuscript.

1.2. PVC Detection Algorithm Details:

First, The n segmented QRST complexes that are used to create the median beat are trimmed around their center of voltage (CoV - see Equation 3 in the main manuscript) to

the shortest QR and RT intervals from all n beats, so that all segmented beats are the same length. This is required to perform cross correlation, and since the QRS complex is the most important feature to perform QRST beat comparison, trimming short lengths of signal from the start or end of the beat is not detrimental. Trimming the beats also reduces the influence of long isoelectric segments where artifact or signal offsets can cause poor matching.

Figure 6, and Supplemental Figures 2 and 3 provide a visual overview and explanation of the PVC detection algorithm. Each of the n beats is compared to itself and the remaining $n - 1$ beats using NCC and RMSE normalized to the max-min of the signals being compared. This results in a $n \times n$ matrix (C) of NCC values and a second $n \times n$ matrix (R) of RMSE values. NCC values above a set threshold (nominally 95%) and RMSE values below a set threshold (nominally 0.1) are assigned a value of 1, while beat-to-beat comparisons that do not meet these criteria are assigned a value of 0. After thresholding, matrices C and R are therefore both $n \times n$ matrices with values of either 0 or 1; beat pairs that have a value of 1 are considered to have similar morphologies based on meeting either NCC or RMSE criteria, while beat pairs that have a value of 0 are considered to have dissimilar morphologies. The matrices C and R are then added together to form a new $n \times n$ “template matrix” (T) which has values of 0, 1, or 2 based on whether the beat pair was considered similar by neither NCC or RMSE, either NCC or RMSE, or both NCC and RMSE, respectively.

The n rows of template matrix T are summed, so that there is now a $n \times 1$ matrix where beats that are correlated with more of the remaining $n - 1$ beats have higher values compared to beats that are correlated with fewer other beats. The first maximum value of this $n \times 1$ matrix is set as the “reference” beat (beat n_{ref}), and the corresponding row of matrix T is saved as a corresponding $1 \times n$ “dominant morphology” matrix D where the values of 0, 1, or 2 indicate the strength of similarity of the reference beat n_{ref} to the remaining $n - 1$ beats. Another thresholding procedure sets values of matrix D which are ≥ 1 equal to 1, and does not change values of 0. Matrix D therefore is now a $1 \times n$ matrix with values D_i equal to 0 if the reference beat had no correlation by either NCC or RMSE with the i th beat, and D_i equal to 1 if the reference beat had correlation by NCC

and/or RMSE with the i th beat.

The software starts by assuming that the reference beat, n_{ref} , by virtue of it being correlated with the largest number of other beats, represents the “dominant”, non-PVC QRST morphology rather than PVC/paced/aberrant QRST morphology. If the number of “dominant” beats (the number of 1s in D) is $> n/2 + 1$, the software assumes that the reference beat QRST morphology is truly the “dominant” / “normal” QRST morphology for the ECG as it is present in the majority of beats. Beats with a dominant matrix (D_i) value of 1 are therefore assigned to be dominant beats, while beats with a D_i value of 0 are assigned to be PVCs.

If $\leq n/2 + 1$ beats are “dominant” based on matrix D , there is a possibility that the reference beat is actually a PVC because there may be similar numbers of normal beats and PVCs (such as occurs with ventricular bigeminy), and/or the PVCs may be more similar with each other than the normal beats are with each other due to noise, differences in alignment, or other factors. In this case, to determine which template is truly the “dominant” QRST morphology, the software compares the QRS durations and QT intervals for the best matches of the dominant and non-dominant beat morphologies. Beat n_{ref} is chosen as the template for the “dominant” QRST template, and a similar selection procedure is performed on “PVC” beats to find the beat that has the highest correlation with all other PVC beats. If the QRS interval of the “dominant” template beat is $< 90\%$ of the QRS interval of the “PVC” template beat, the reference beat is defined as “dominant” and values of 0 in matrix D are identified as PVCs. If the QRS interval of the “dominant” template beat is not $< 90\%$ of the QRS interval of the “PVC” template beat, the software does the same comparison with the QT interval, and if the QT interval of the “dominant” template beat is $< 90\%$ of the QT interval of the “PVC” template beat the reference beat is defined as “normal” and values of 0 in matrix D are identified as PVCs. Otherwise the software decides that it initially selected the “dominant” QRST morphology incorrectly, and values of 1 in matrix D are identified as PVCs.

1.3. Outlier Detection Using Modified Z-Scores:

For each beat, the QR and RT intervals are computed. A modified Z-score (Z_i) for each of these quantities is then computed based on the median absolute deviation (MAD) or mean absolute deviation (MeAD) [1]:

$$\begin{aligned} \text{MAD} &= \text{median}(|x_i - \tilde{x}|) \\ \text{MeAD} &= \text{mean}(|x_i - \bar{x}|) \end{aligned} \tag{1}$$

where x_i is the value for the i th beat, \tilde{x} is the median value for all i beats, and \bar{x} is the mean value for all i beats.

For each QR or RT interval:

$$\begin{aligned} Z_i &= 0.6745 * (x_i - \tilde{x})/\text{MAD} && \text{if MAD} \neq 0 \\ Z_i &= 0.7979 * (x_i - \tilde{x})/\text{MeAD} && \text{if MAD} = 0 \end{aligned} \tag{2}$$

If Z_i is > 3.5 nominally [1] (this cutoff point can be adjusted) the i th beat is marked as an outlier and can be automatically removed.

1.4. Neural Network Training Data Source/Preparation:

To train the vector magnitude (VM) median beat annotator neural network (NN), we searched the Beth Israel Deaconess Medical Center ECG database (MUSE, General Electric, Boston, MA) and obtained 1,262 sequential, 10-second 12-lead ECGs sampled at 500 Hz with a $4.88 \mu\text{V}/\text{unit}$ resolution which were of diagnostic quality. ECGs with extreme noise, missing leads, or other technical issues were excluded. Using a pre-release version of BRAVEHEART, 12-lead ECGs were first filtered with both low- (62.5 Hz) and high-pass (0.24 Hz) wavelet filters using the Symlets 4 and Daubechies 4 wavelets, respectively, to remove noise and baseline wander. ECGs were transformed into orthogonal X, Y, Z leads using the Kors transformation matrix [2]. The median VM lead was constructed from the median X, Y, and Z leads as described in the main manuscript methods. Determination of ECG fiducial points is most unambiguously performed on the VM lead [3] and use of the VM lead allows assessment of “global” fiducial points and

intervals on a single lead [4]. Use of the VM lead also simplifies computation needed for fiducial point detection because the VM lead is always positive, and VM lead morphology is simpler/more conserved than the variety of possible signal morphologies seen in the 12 standard ECG leads.

Median VM beats were constructed by aligning individual beats based on their center of voltage (CoV) as described in the main manuscript. QRS onset (Q_{on}), QRS offset (Q_{off}), and T wave offset (T_{off}) were then manually annotated by an experienced, board certified cardiac electrophysiologist (JWW) using electronic calipers, and the annotated median beat signals and annotations were exported for later use. We included normal and abnormal ECGs, including paced ECGs. For paced QRS complexes, Q_{on} was defined as occurring just after the completion of the high-frequency pacing spike. P waves were not annotated.

Each sample in the VM waveform was associated with a categorical signal containing one of three labels; “QRS complex”, “T wave,” or “other” based on the annotated fiducial points. “QRS complex” was defined as the signal between Q_{on} and Q_{off} , “T wave” was defined as the signal between $Q_{off} + 1$ sample and T_{off} , and all other parts of the VM waveform were defined as “other” (see Figure 8 in the main manuscript).

1.5. Neural Network Architecture and Training:

We trained a bidirectional Long-Short Term Memory (biLSTM) NN [5] for sequence-to-sequence classification using Matlab version 2020b. Details of the chosen NN architecture and its performance have been previously reported [6]. The 1,262 annotated median VM beats from 1,262 unique patients were split into training (75%, n=947), validation (10%, n=126), and testing (15%, n=189) datasets. The training dataset beats were then augmented using previously described and validated median beat ECG augmentation software (ECGAug) [6] by recombining different QRS and T wave segments with further physiologically appropriate distortion while maintaining the locations of manually annotated fiducial points. Augmentation resulted in a final training dataset of 8,479 annotated VM median beats. Prior to training, the VM median beat signals were further processed by removing frequencies outside of 0.5-40 Hz (the power spectrum of the QRST complex)[7] using the Fourier synchrosqueezed transform with a Kaiser-Bessel

window of length 128 and $\beta=0.5$, and then standardized by subtracting the mean and diving by the standard deviation.

For NN training, the 20 real and imaginary parts of the transformed signal were used as input into the biLSTM. The input layer was fed into a biLSTM layer with 80 neurons which connected to a fully-connected layer with 3 outputs and then a softmax activation function to assign probabilities for the 3 possible labels (“QRS complex”, “T wave”, or “other”) at each sample of the VM median beat signal. Hyperparameters were manually tuned to find optimal NN performance based on prior experiments [6]. The final/best NN was trained with a batch size of 32 for 5,000 iterations using the Adam optimizer [8]. The initial learning rate was 0.003 with a 80% decrease in learning rate every 400 iterations. We empirically found that more layers and/or more neurons resulted in over-fitting of the data, while fewer neurons resulted in poorer annotation performance. Forcing training to proceed above 5,000 iterations tended to result in over-fitting [6]. Training curves for the final NN are shown in Supplemental Figure 13.

Just as human fiducial point annotations were used to create a categorical signal for NN training (Figure 8A in the main manuscript), the NN output probabilities of each of the 3 labels (“QRS complex”, “T wave”, or “other”) were then used to assign NN predicted fiducial points based on the locations of transition between segments of the QRST complex (Figure 8B in the main manuscript). Logic was used to convert these probabilities into fiducial point locations. In general, Q_{on} was assigned as the sample where there was a transition in the predicted categorical signal from “other” to “QRS complex”, Q_{off} was assigned as the sample where there was a transition in the predicted categorical signal from “QRS complex” to “T wave”, and T_{off} was assigned as the sample where the predicted categorical signal transitioned from “T Wave” to “other”. For signals that are not sampled at 500 Hz, the signal is down/upsampled to 500 Hz before being passed into the NN, and the calculated fiducial points are then adjusted based on the original frequency.

In rare cases where there were multiple points where the definition for a fiducial points was satisfied (usually ECGs with significant artifact or very low T wave amplitude), an algorithm which takes the relative locations of Q_{on} , Q_{off} , and T_{off} was used to choose the optimal location for the fiducial point; for example, the algorithm does not allow

assignment of fiducial points out of order, and will not assign a T_{off} before a Q_{off} . These ECGs are flagged for review as part of BRAVEHEART post processing quality checking (see Methods of the main manuscript for details).

1.6. Neural Network Performance Assessment:

Since we were primarily interested in fiducial points which occur at transitions between labels in the QRST complex, overall classification accuracy for the entire VM signal was not felt to be a good measure of NN performance. The majority of the VM median beat signal would always be annotated correctly regardless of NN performance, but the predicted fiducial points could differ significantly from ground truth even with high overall classification accuracy for the entire signal; for example, if a median beat signal was 500 samples, Q_{on} could be off by 10 samples (20 ms at 500 Hz), which would indicate overall poor performance, but the entire signal could still be annotated with 99% accuracy.

NN performance was therefore assessed by comparing the predicted location of Q_{on} , Q_{off} , and T_{off} and the clinically relevant intervals (QRS duration [$Q_{\text{off}} - Q_{\text{on}}$] and QT interval [$T_{\text{off}} - Q_{\text{on}}$]) to the ground truth expert manual annotations/intervals in the testing dataset. Unlike many other classification problems where there is an unambiguous ground truth for use in NN training and testing, fiducial point annotation is associated with inherent uncertainty, and even trained experts can sometimes disagree as to the true location of a specific fiducial point (especially the T wave end) [9]. Thresholds for adequate performance were based on consensus statements for electronic ECG annotation [10, 11, 12, 13].

The mean and standard deviation of the difference between the NN predicted and ground truth fiducial point locations and intervals were calculated. Missing transitions were defined as when a Q_{on} , Q_{off} , or T_{off} fiducial point was not found in a median VM QRST signal, and extra transitions were defined as when multiple potential Q_{on} , Q_{off} , or T_{off} fiducial points were present in a single median VM QRST signal. As this was a multi-class labeling problem with imbalanced classes, the precision (positive predictive value), recall (specificity), and F1 (the harmonic mean of precision and recall) are all equal, and only micro-averaged F1 was therefore reported (See Supplemental Statistical Methods below).

We have previously described detailed training results using the current biLSTM NN architecture [6], and given that BRAVEHART relies heavily on using the specific weights and biases which are included in the software, for the purpose of the software release we validated NN performance on a second, independent testing dataset using 200 sequential ECGs obtained a single day which were free of significant artifact or missing leads. NN predicted fiducial points and intervals were compared to ground truth annotations for this dataset in the same way as noted above for the original testing dataset.

1.7. Quality Labeling Methods:

A set of 481 sequential ECGs obtained on a single day and which were not used for any other parts of BRAVEHEART development were processed by BRAVEHEART using nominal settings and without regard to signal quality. Processed ECGs/VM median beats were manually reviewed and labeled as “good quality” or “needs review”. Multivariable logit regression was performed, initially including the minimum NCC for the X, Y, and Z leads (NCC), median VM lead voltage in the 30 ms after the end of the T wave (baseline), maximum VM lead T wave voltage (TMag), VM lead T peak-T end ratio (TPTE), number of removed beats (either via PVC or outlier detection), and measures of high-frequency and low-frequency noise. After adjustment, only NCC, baseline, TMag, and TPTE were independently associated with the quality of the final processed ECG. The final logit coefficients were used to calculate predicted probabilities (range 0–1) of “good quality” ECGs after BRAVEHEART processing, and receiver-operating curves (ROC) and sensitivity and specificity were calculated at various cut points.

1.8. Supplemental Statistical Methods:

Microaveraged (μ) Statistics help adjust for the imbalances present between different classes in the VM median beat annotation data:

$$\mu\text{Precision} = \frac{\sum_{k=1}^3 \text{TP}_k}{\sum_{k=1}^3 \text{TP}_k + \text{FP}_k} \quad (3)$$

$$\mu\text{Recall} = \frac{\sum_{k=1}^3 \text{TP}_k}{\sum_{k=1}^3 \text{TP}_k + \text{FN}_k} \quad (4)$$

$$\mu F1 = 2 \times \frac{\mu Precision \times \mu Recall}{\mu Precision + \mu Recall} \quad (5)$$

where TP are true positives, FP are false positives, and FN are false negatives, and for $k = 3$ different categorical labels are “QRS Complex”, “T Wave”, or “other”.

2. Supplemental Results

2.1. Quality Labeling Results:

The final logit model was given as:

$$p = -22.56 + (18.21 * NCC) + (-87.59 * \text{baseline}) + (9.06 * \text{TMag}) + (11.19 * \text{TPTE}) \quad (6)$$

and therefore the probability of a specific processed ECG being “good” is given as:

$$P(\text{good}) = e^p / (1 + e^p) \quad (7)$$

The ROC curve for the quality regression is shown in Supplemental Figure 4 and had an area under the ROC of 0.973 indicating excellent agreement between prediction and manual labeling. Sensitivity and specificity at various cut points of $P(\text{good})$ are shown in Figure 10 in the main manuscript. As can be seen, at around a cut point of 0.8 there was a good combination of sensitivity and specificity, with both values $> 90\%$. Cut points > 0.9 tended to have higher specificity and lower sensitivity, while values < 0.7 tended to have lower specificity and higher sensitivity. Cut points can be adjusted in BRAVEHEART based on the dataset being analyzed and the relative sensitivity or specificity that is desired in the dataset being analyzed, which will depend on the overall quality of the ECG tracings being analyzed.

3. Equations

Vector Magnitude:

Vector magnitude (VM) is defined as the Euclidean norm of the VCG:

$$\text{VM} = \sqrt{X^2 + Y^2 + Z^2} \quad (8)$$

Area Vectors:

Area vectors are obtained by taking the area under the relevant segment of the median VCG ($\mathbf{V}(t) = [X(t), Y(t), Z(t)]$) QRST complex using the trapezoidal rule:

$$\mathbf{QRS}_{\text{area}} = \int_{Q_{\text{on}}}^{Q_{\text{off}}} \mathbf{V}(t) dt = \left[\int_{Q_{\text{on}}}^{Q_{\text{off}}} X(t) dt, \int_{Q_{\text{on}}}^{Q_{\text{off}}} Y(t) dt, \int_{Q_{\text{on}}}^{Q_{\text{off}}} Z(t) dt \right] \quad (9)$$

$$\mathbf{T}_{\text{area}} = \int_{Q_{\text{off}}}^{T_{\text{off}}} \mathbf{V}(t) dt = \left[\int_{Q_{\text{off}}}^{T_{\text{off}}} X(t) dt, \int_{Q_{\text{off}}}^{T_{\text{off}}} Y(t) dt, \int_{Q_{\text{off}}}^{T_{\text{off}}} Z(t) dt \right] \quad (10)$$

Peak Vectors:

Peak vectors are obtained by taking the value of relevant segment of $\mathbf{V}(t)$ at the time point which is of maximum distance from the origin:

$$\mathbf{QRS}_{\text{peak}} = [X(t_{Q_{\max}}), Y(t_{Q_{\max}}), Z(t_{Q_{\max}})] \quad (11)$$

$$\mathbf{T}_{\text{peak}} = [X(t_{T_{\max}}), Y(t_{T_{\max}}), Z(t_{T_{\max}})] \quad (12)$$

where $t_{Q_{\max}}$ is the time of max distance of the QRS loop from the origin, and $t_{T_{\max}}$ is the time of max distance of the T loop from the origin. These times correspond to the maximum values of the QRS complex and T wave in the VM lead, respectively.

Spatial Ventricular Gradient:

The Spatial Ventricular Gradient (SVG) is the vector created by the QRST integrals in X , Y , and Z :

$$\mathbf{SVG} = \mathbf{QRS}_{\text{area}} + \mathbf{T}_{\text{area}} = \int_{Q_{\text{on}}}^{T_{\text{off}}} \mathbf{V}(t) dt = \left[\int_{Q_{\text{on}}}^{T_{\text{off}}} X(t) dt, \int_{Q_{\text{on}}}^{T_{\text{off}}} Y(t) dt, \int_{Q_{\text{on}}}^{T_{\text{off}}} Z(t) dt \right] \quad (13)$$

Azimuth Angle:

Azimuth is defined as the angle in the transverse (XZ) plane with negative angles pointing anterior, and positive angles pointing posterior (see Supplemental Figure 14). Azimuth can take values from 0 to ± 180 degrees, with 0 degrees pointing to towards the left, and ± 180 degrees pointing towards the right.

$$\text{azimuth} = \arctan\left(\frac{Z}{X}\right) \quad (14)$$

Elevation Angle:

Elevation is defined as the angle in the frontal (XY) plane (See Supplemental Figure 14). Values range from 0 to 180 degrees, with 0 degrees pointing towards the feet and 180 degrees pointing towards the head.

$$\text{elevation} = \arccos\left(\frac{Y}{\text{VM}}\right) \quad (15)$$

Absolute Integrals:

The sum absolute integral (SAI) is defined as the area under the absolute value of the QRST complex:

$$\text{SAI}_i = \int_{Q_{\text{on}}}^{T_{\text{off}}} |V_i(t)| dt \quad \text{for } i = X, Y, Z, \text{ or VM} \quad (16)$$

SAI QRST is defined as:

$$\text{SAI QRST} = \text{SAI}_x + \text{SAI}_y + \text{SAI}_z = \int_{Q_{\text{on}}}^{T_{\text{off}}} |X(t)| dt + \int_{Q_{\text{on}}}^{T_{\text{off}}} |Y(t)| dt + \int_{Q_{\text{on}}}^{T_{\text{off}}} |Z(t)| dt \quad (17)$$

QRST Angles:

The spatial QRST angle is the 3-dimensional angle between QRS and T vectors:

$$\text{QRST Angle} = \arccos\left(\frac{\mathbf{QRS} \cdot \mathbf{T}}{|\mathbf{QRS}| |\mathbf{T}|}\right) \quad (18)$$

where the peak QRST angle uses $\mathbf{QRS}_{\text{peak}}$ and \mathbf{T}_{peak} , and the area QRST angle uses $\mathbf{QRS}_{\text{area}}$ and \mathbf{T}_{area} .

Total Cosine R to T (TCRT):

TCRT was calculated as previously described using singular value decomposition [14].

VCG Loop Length:

Loop length is calculated as the sum of distances covered as the VCG loop increments by each sample. For example, QRS loop length is calculated as:

$$\sum_{i=Q_{\text{on}}}^{Q_{\text{off}-1}} \sqrt{(X_{i+1} - X_i)^2 + (Y_{i+1} - Y_i)^2 + (Z_{i+1} - Z_i)^2} \quad (19)$$

and T loop length is calculated as:

$$\sum_{i=Q_{\text{off}}}^{T_{\text{off}-1}} \sqrt{(X_{i+1} - X_i)^2 + (Y_{i+1} - Y_i)^2 + (Z_{i+1} - Z_i)^2} \quad (20)$$

VCG Loop Speed:

The instantaneous speed of the QRS or T loops (v_i) in mV/ms for a VCG with frequency f is calculated as the distance traveled in a sample of time:

$$v_i = \sqrt{(X_{i+1} - X_i)^2 + (Y_{i+1} - Y_i)^2 + (Z_{i+1} - Z_i)^2} / \Delta t \quad \text{where } \Delta t = 1000/f \quad (21)$$

Left Ventricular Hypertrophy (LVH):

LVH metrics are calculated as previously described: The Cornell Voltage is calculated as the sum of the S wave in lead V3 + the R wave in lead aVL [15]. The Sokolow-Lyon LVH criteria is calculated as the sum of the S wave in lead V1 and the maximum of the R wave in either lead V5 or V6 [16].

VCG Loop Morphology and Singular Value Decomposition:

The singular value decomposition (SVD) is used to find the best fit planes for the QRS and T loops separately. Let $X' = X - c_x$, $Y' = Y - c_y$, and $Z' = Z - c_z$ where the centroid $\mathbf{c} = (c_x, c_y, c_z)$ is the mean QRS or T vector. Then let the matrix \mathbf{M}' have the

centroid-subtracted leads as columns. The SVD of \mathbf{M}' is:

$$\mathbf{M}' = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (22)$$

After SVD, unit vectors that span the best fit plane for the VCG loop are given by the columns of \mathbf{V} .

The 3 singular values along the diagonal of matrix \mathbf{S} , when squared, give the proportion of mean-squared error or variance (σ^2) in the direction of the 3 corresponding basis vectors in matrix \mathbf{V} :

$$S = \begin{bmatrix} S_1 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & S_3 \end{bmatrix} \quad \sigma^2 = \begin{bmatrix} S_1^2 & 0 & 0 \\ 0 & S_2^2 & 0 \\ 0 & 0 & S_3^2 \end{bmatrix} \quad (23)$$

The columns are arranged such that $S_1 > S_2 > S_3$. Assuming that the VCG loop is approximately planar, the third column of matrix \mathbf{V} is normal to the best fit plane; let this vector be denoted by \mathbf{n} . Note that if the the VCG loop is perfectly coplanar, then $S_3 = 0$. The major and minor axes of the VCG loop are then given by the first and second columns of \mathbf{V} , respectively.

The degree of VCG loop coplanarity is assessed in 2 ways. First, the mean standard error (MSE) in the direction normal to the best fit plane (S_3^2) is obtained from squaring the 3rd singular value as noted above. If $S_3^2 = 0$ then all points in the loop are coplanar. Coplanarity is also assessed by calculating RMSE for points in the VCG loop compared to the points projected onto the best fit plane:

For a given point in the QRS or T loops $\mathbf{s} = [s_x, s_y, s_z]$ let $\mathbf{s}' = \mathbf{s} - \mathbf{c}$. Then the projection s'_{proj} relative to the centroid \mathbf{c} onto the best-fit plane is found by subtracting out the component of the vector in the direction of the normal vector \mathbf{n} defined above:

$$\mathbf{s}'_{\text{proj}} = \mathbf{s}' - (\mathbf{s}' \cdot \mathbf{n})\mathbf{n} \quad (24)$$

The contribution to the RMSE is then the square root of the mean distance between

corresponding points \mathbf{s} and \mathbf{s}_{proj} . Using this metric, if all points are coplanar, RMSE = 0.

VCG Loop Dihedral Angle:

VCG loop dihedral angle is defined as the 3-dimensional angle between the unit normal vectors that define the QRS loop best fit plane (N_{QRS}) and the T loop best fit plane (N_{T}). By convention, the dihedral angle is an acute angle:

$$\text{Dihedral Angle} = \arccos(|N_{\text{QRS}} \cdot N_{\text{T}}|) \quad (25)$$

VCG Loop Roundness:

Assuming that the MSE in the direction normal to the best fit plane should be relatively small compared to the MSE in the plane itself, the “roundness” (R) of the VCG loop is defined as the ratio of the largest to second largest singular values:

$$R = S_1/S_2 \quad (26)$$

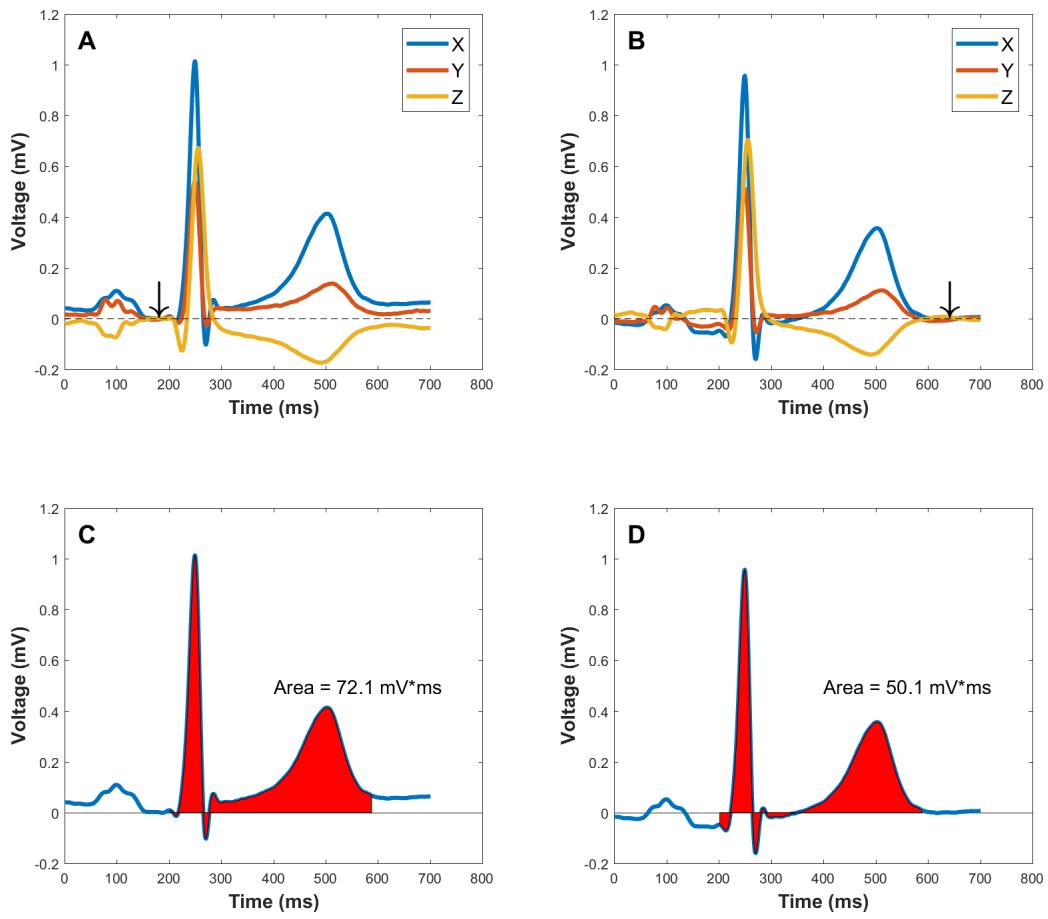
If $R = 1$ the VCG loop is a perfect circle, and as the value of R increases above 1 the loop is more oval or oblong.

VCG Loop Perimeter:

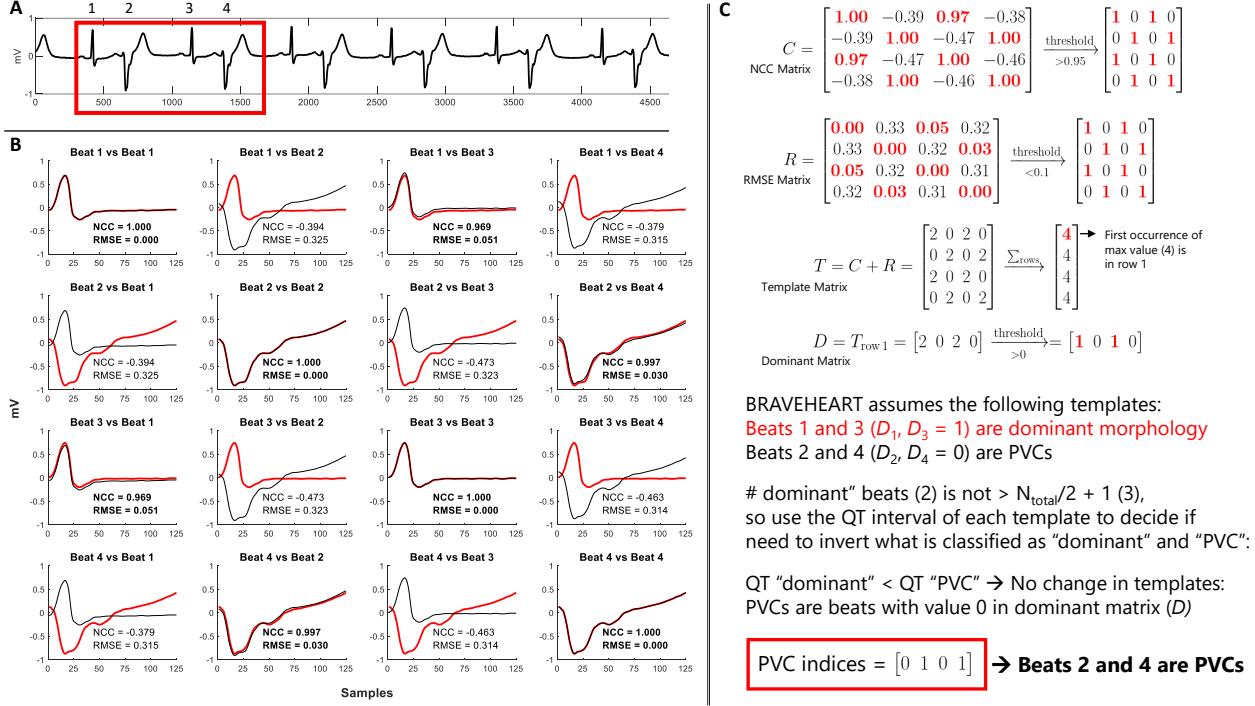
QRS and T loop perimeter are calculated as the length of the QRS or T loop projected into the best fit plane with the set of points defined as M_{proj} , and is analogous to the QRS and T loop length which is the length of the loop without projection (set of points defined as M). The perimeter of M_{proj} is calculated using Matlab polyshapes.

VCG Loop Area:

QRS and T loop area are calculated as the area of the QRS or T loop projected into the best fit plane with the set of points defined as M_{proj} . The area of M_{proj} is calculated using Matlab polyshapes.

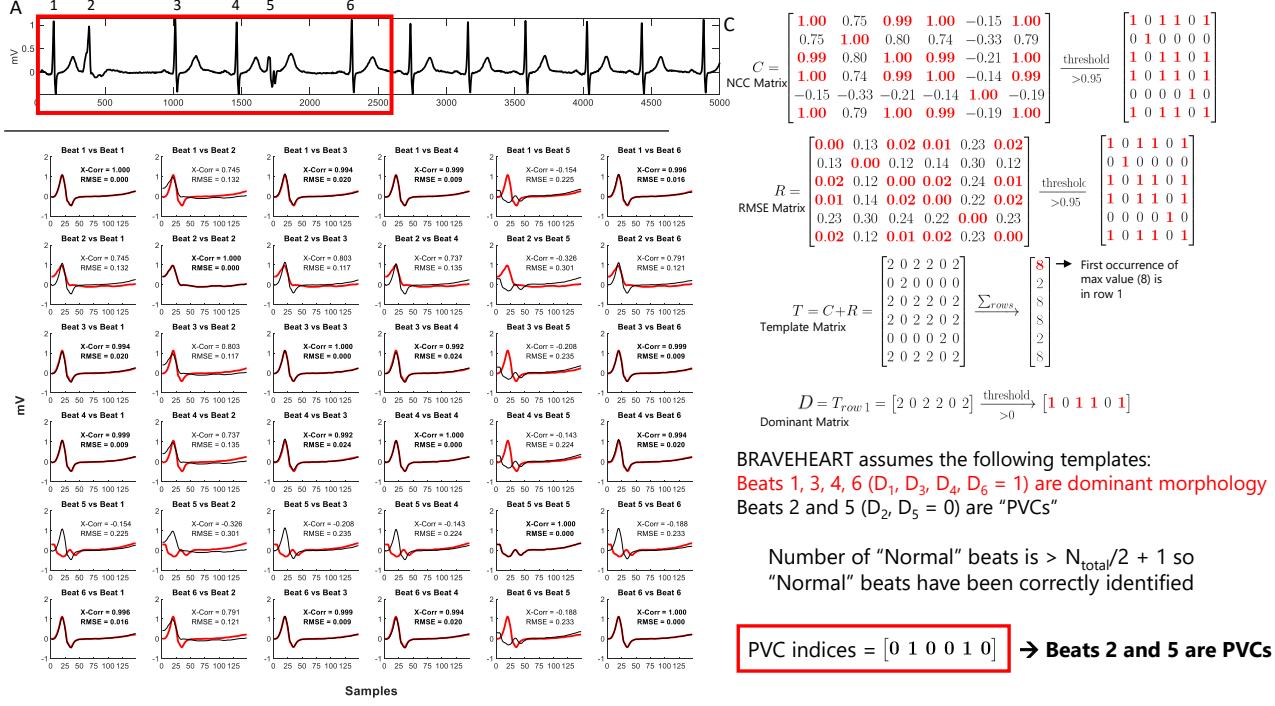


Supplemental Figure 1: Significant differences in QRST area calculations due to different definitions of the vectorcardiographic origin point. **A:** Median VCG beats (X, Y, Z) without any baseline processing where the onset of the QRS complex (arrow) is defined as zero voltage. **B:** After baseline correction using the method described in the text, the TP segment, a physiologically isoelectric interval, is now set as the zero voltage (arrow). **C** and **D:** Physiological baseline correction results in a 30% decrease in area under the X median beat. Reproduced with permission from [17].



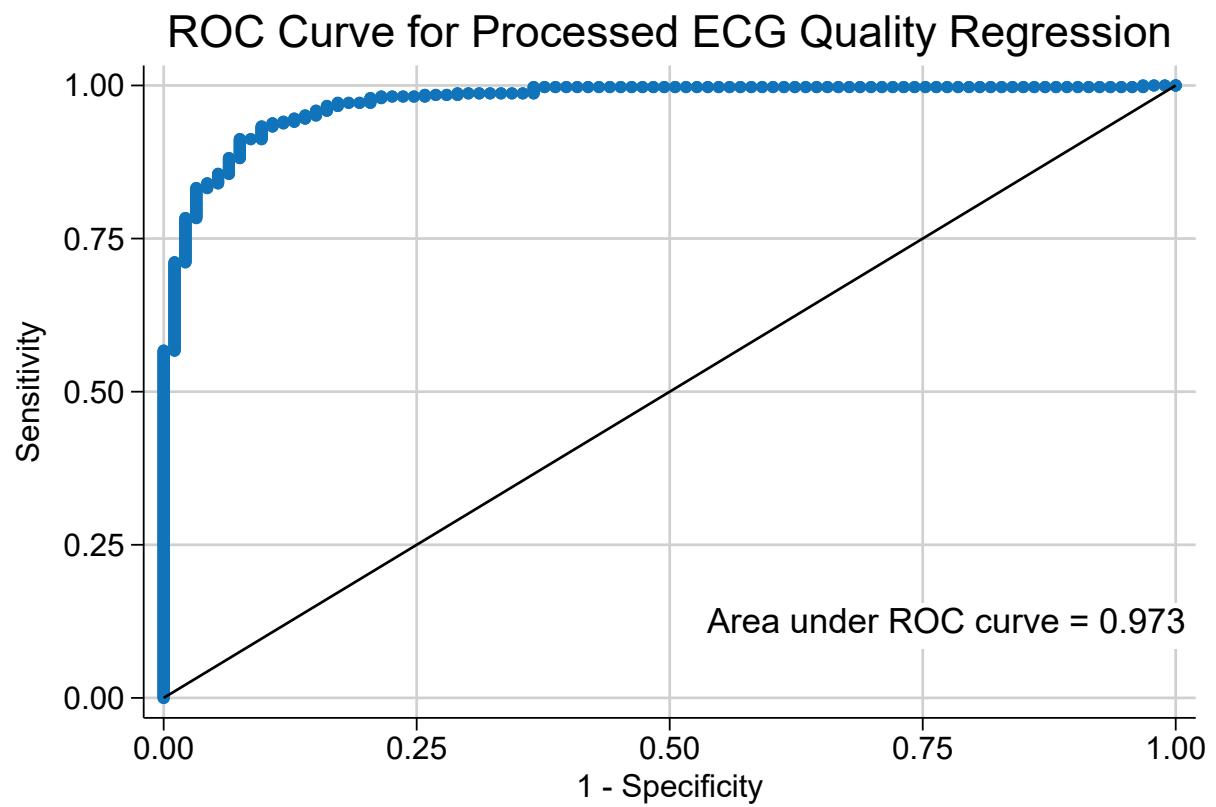
Supplemental Figure 2: Example of how the BRAVEHEART PVC detection algorithm works on an ECG with bigeminy and equal numbers of sinus beats and PVCs. **A:** A subset of 4 beats (2 sinus and 2 PVCs) are analyzed for illustrative purposes, but the algorithm would have the same result if all 12 beats were analyzed. **B:** After trimming, each beat is aligned with the remaining beats, and normalized cross correlation (NCC) and root mean squared error (RMSE) are calculated. **C:** Values of NCC and RMSE are placed in matrices C and R , respectively, and then thresholds are applied as described in the text. The template matrix T is formed by adding $C + R$ and summing the rows. The dominant matrix D is formed by taking the first row of T that has the maximum value of T , and then performing additional thresholding. BRAVEHEART starts by assuming that values of 1 in D are the dominant morphology and values of 0 in D are PVCs. In this case, the number of beats with dominant morphology is not $> n/2 + 1$, so the algorithm has to look at the QRS and QT interval of each morphology to decide if the “dominant” morphology is sinus or a PVC. In this case the QRS duration of the “dominant” morphology is $<90\%$ the QRS duration of the “PVC” morphology, so the template associated with the “dominant” morphology (narrow beats) is the true dominant morphology. The indices of PVCs are found by taking values of 0 in matrix D (beats 2 and 4).

Abbreviations: PVC - premature ventricular contraction

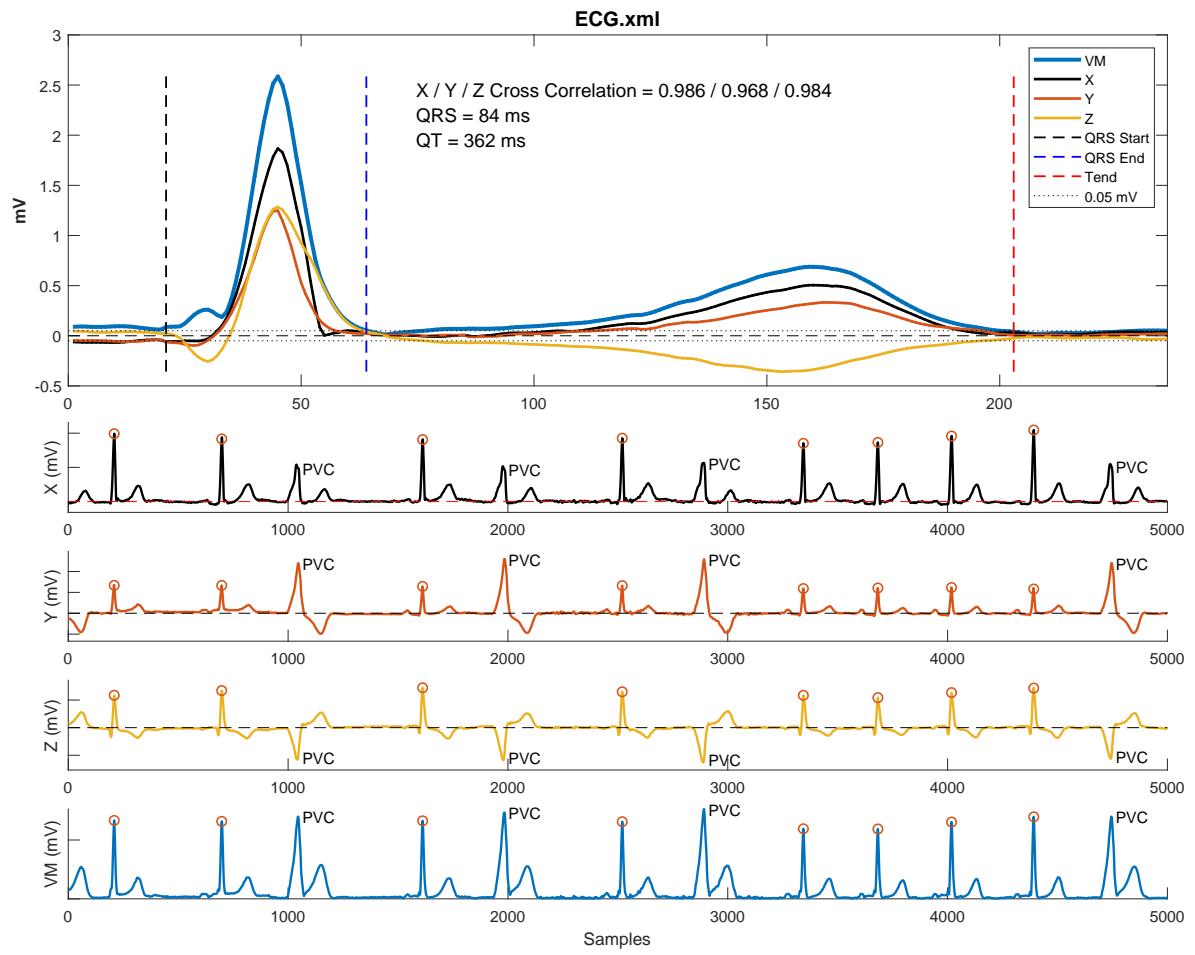


Supplemental Figure 3: Example of how the BRAVEHEART PVC detection algorithm works on an ECG with multiple PVC morphologies. **A:** A subset of 6 beats (4 sinus and 2 PVCs of different morphology) are analyzed for illustrative purposes, but the algorithm would have the same result if all 12 beats were analyzed. **B:** After trimming, each beat is aligned with the remaining beats, and normalized cross correlation (NCC) and root mean squared error (RMSE) are calculated. **C:** Values of NCC and RMSE are placed in matrices C and R , respectively, and then thresholds are applied as described in the text. The template matrix T is formed by adding $C + R$ and summing the rows. The dominant matrix D is formed by taking the first row of T that has the maximum value of T , and then performing additional thresholding. BRAVEHEART starts by assuming that values of 1 in D are the dominant morphology and values of 0 in D are PVCs. In this case, since the number of beats with dominant morphology is $> n/2 + 1$, the algorithm is complete and the indices of PVCs are found by taking values of 0 in matrix D (beats 2 and 5).

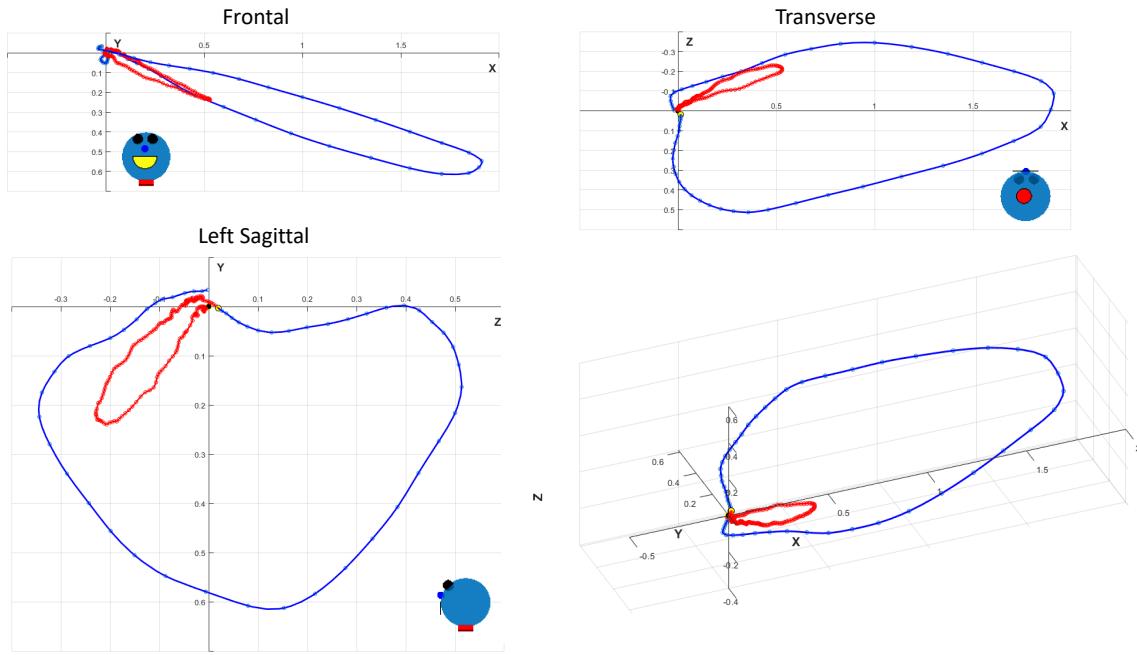
Abbreviations: PVC - premature ventricular contraction



Supplemental Figure 4: Receiver Operating Curve for quality regression. The area under the receiver operating curve of 0.973 indicates that the final model has excellent predictive ability.

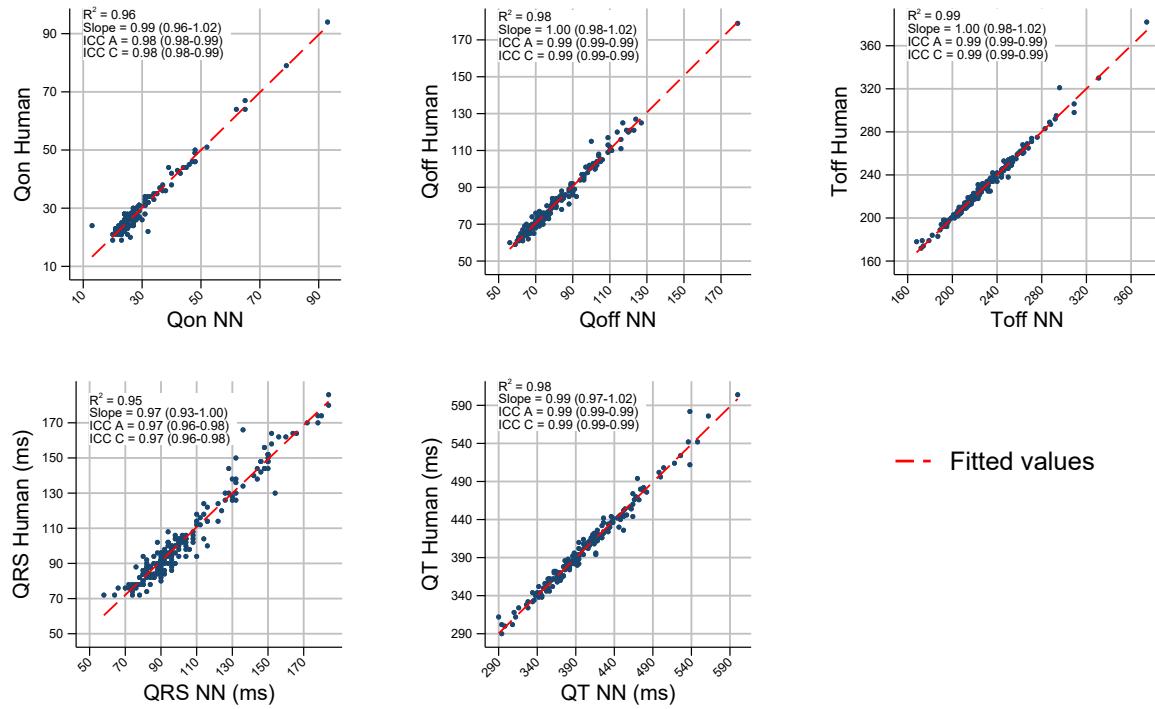


Supplemental Figure 5: Example of summary figure output. The X, Y, Z, and VM median beats and their annotations are show in the top panel, and the full VCG is shown in the bottom half of the figure. Beats included in the median beat have an orange circle at their R peak. In this example, beats without an orange circle are labeled as “PVC” to indicate they were detected as PVCs and removed from the median beat analysis. Cross correlation represents the average normalized cross correlation between all pairs of beats that make up the median beat and represents the quality of median beat construction with values very close to 1 indicating excellent beat alignment.



Supplemental Figure 6: Example of VCG generated from an ECG. The QRS loop is blue and the T wave loop is red.

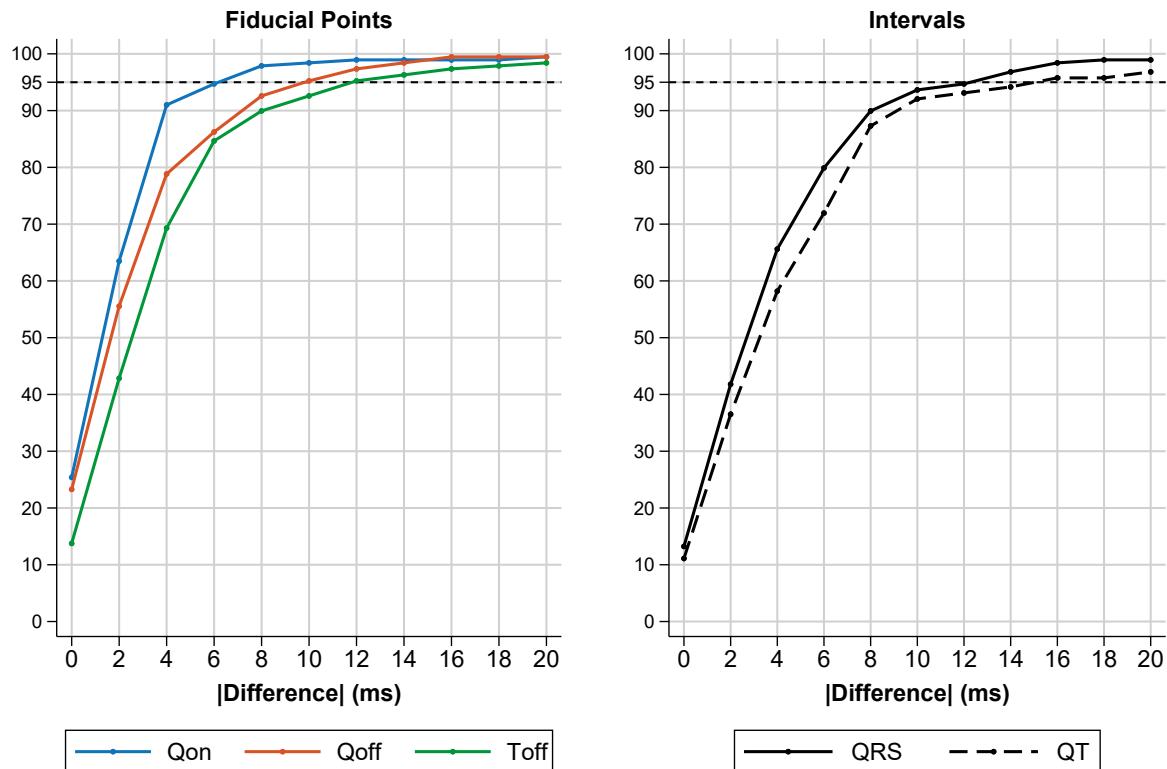
Ground Truth vs. Neural Network Annotations : Testing Dataset (N=189)



Supplemental Figure 7: Correlations between NN predictions and ground truth fiducial points for the initial 189 ECG testing dataset. Note that relationship between NN predicted and ground truth values of all fiducial points are very liner and ICC values are all very close to 1.

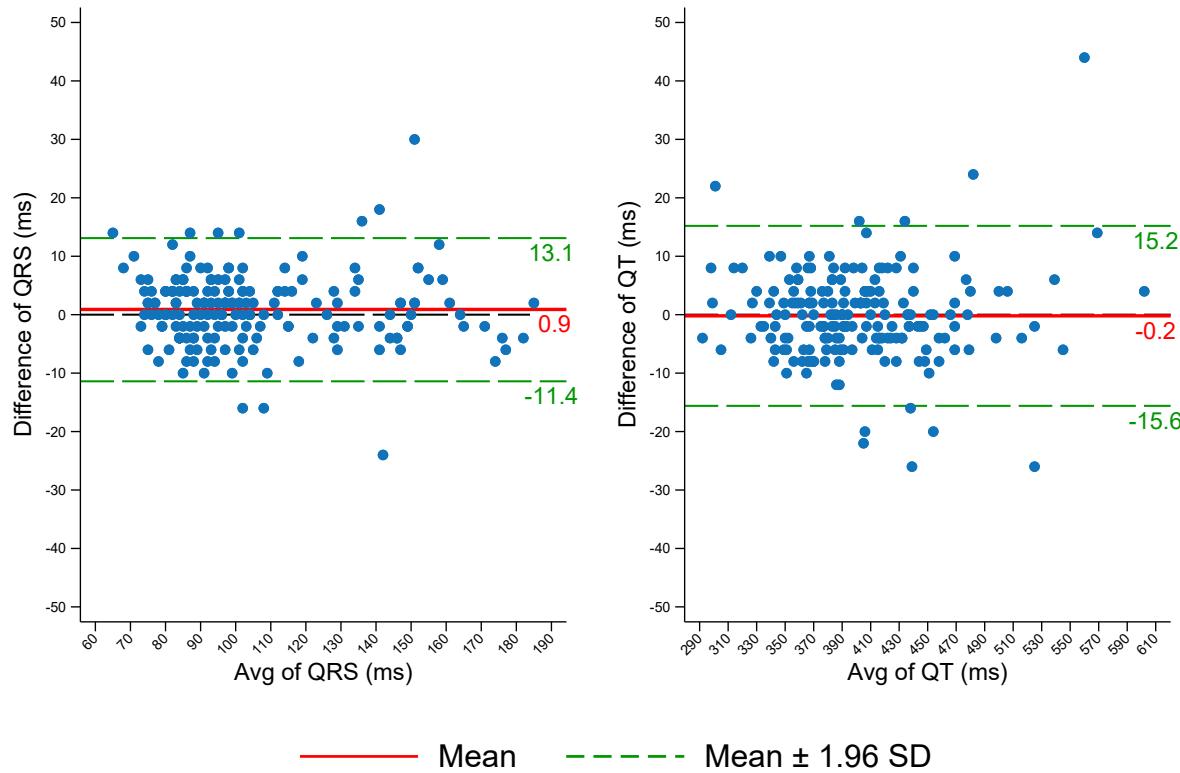
Abbreviations: ICC A - absolute agreement intraclass correlation coefficient, ICC C - consistency of agreement intraclass correlation coefficient.

% of Neural Network Annotations Within Specific Accuracy



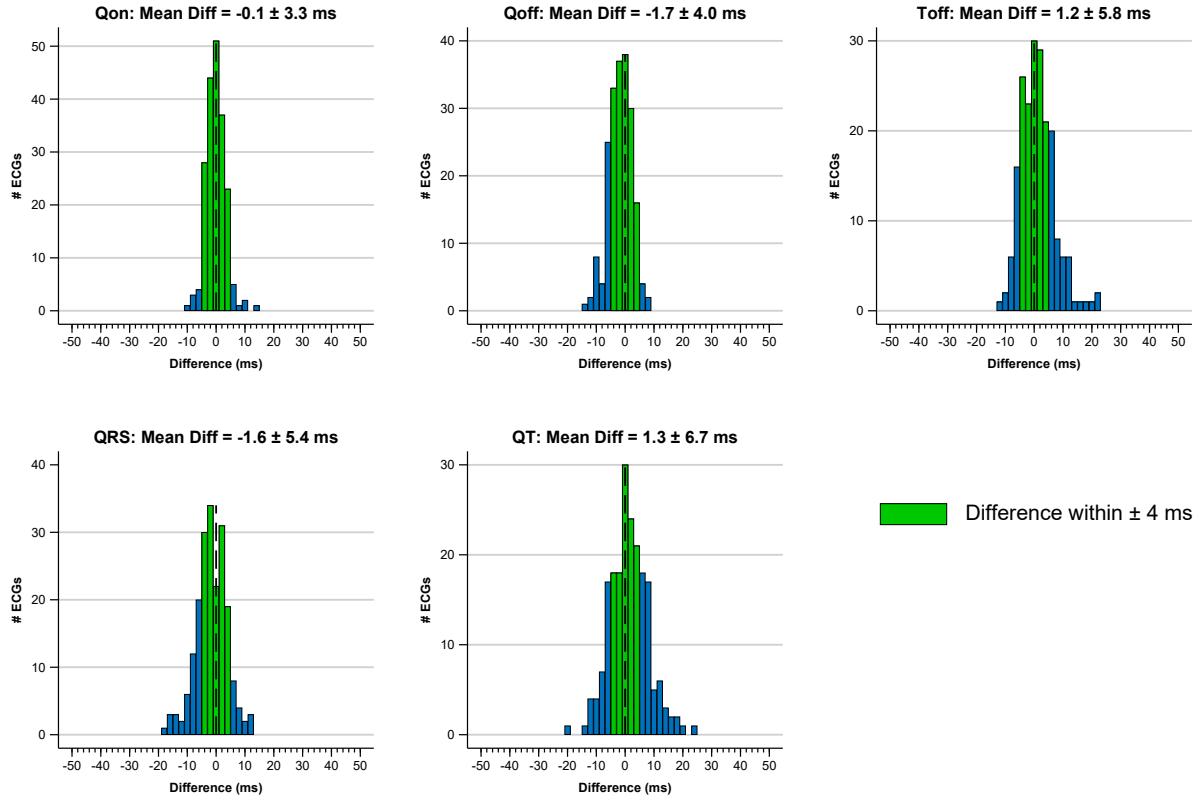
Supplemental Figure 8: Percentage of NN detected fiducial points/intervals that were within various errors (in ms) compared to ground truth for the initial 189 ECG testing dataset.

Bland-Altman Plot: Ground Truth vs. Neural Network



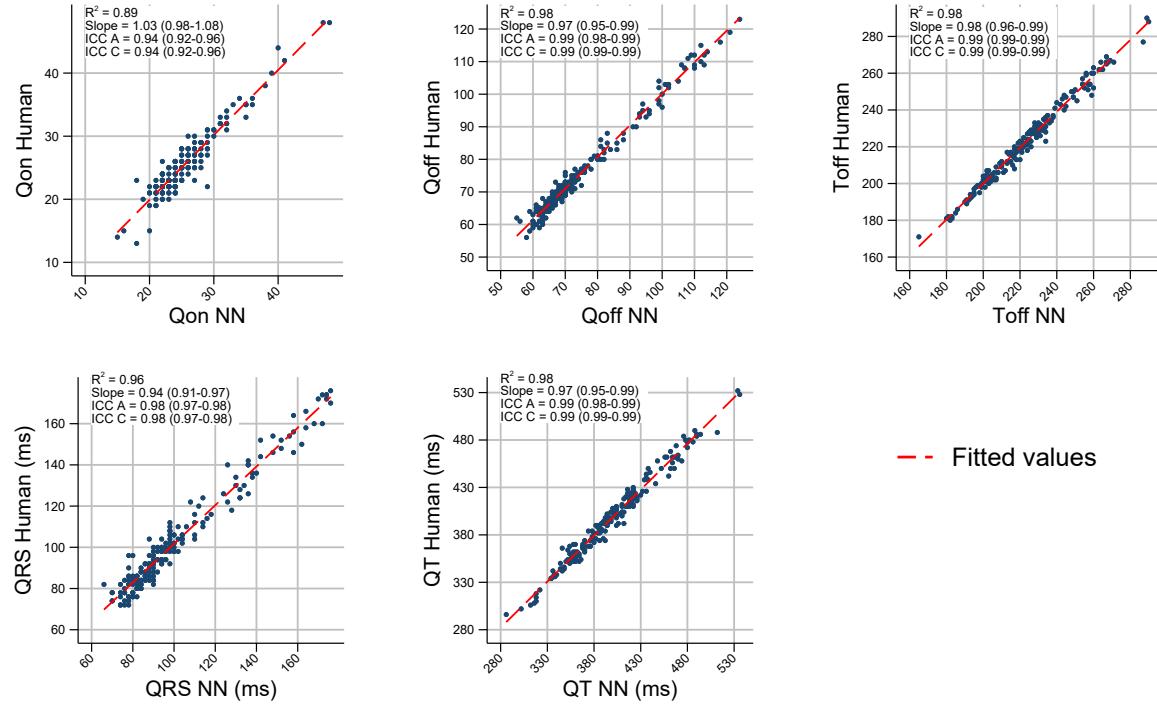
Supplemental Figure 9: Bland-Altman plot for QRS duration and QT interval for the initial 189 ECG testing dataset.

Difference in Fiducial Point Annotation: Neural Network - Ground Truth : Testing Dataset 2 (N=200)



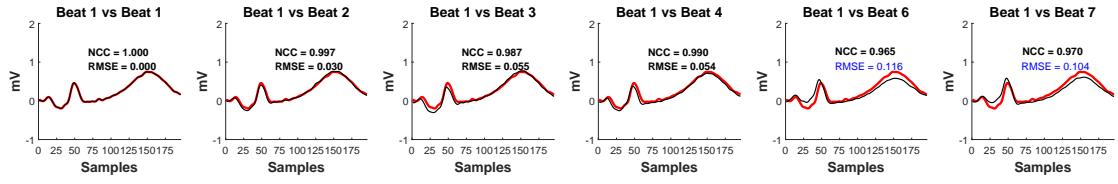
Supplemental Figure 10: Mean and standard deviation and histograms for results of NN testing for the second 200 ECG testing dataset. Bins where the difference between NN predictions and ground truth fiducial points were $\leq \pm 4$ ms are shown in green.

Ground Truth vs. Neural Network Annotations : Testing Dataset 2 (N=200)

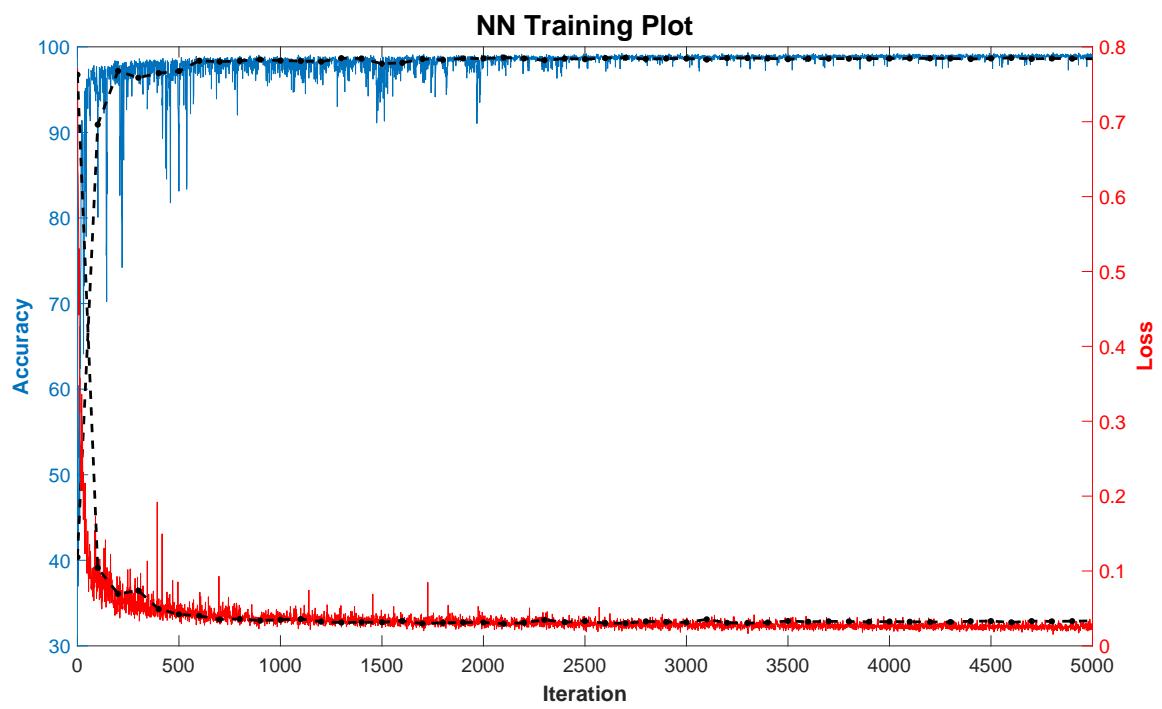


Supplemental Figure 11: Correlations between NN predictions and ground truth fiducial points for the second 200 ECG testing dataset. Note that relationship between NN predicted and ground truth values of all fiducial points are very liner and ICC values are all very close to 1.

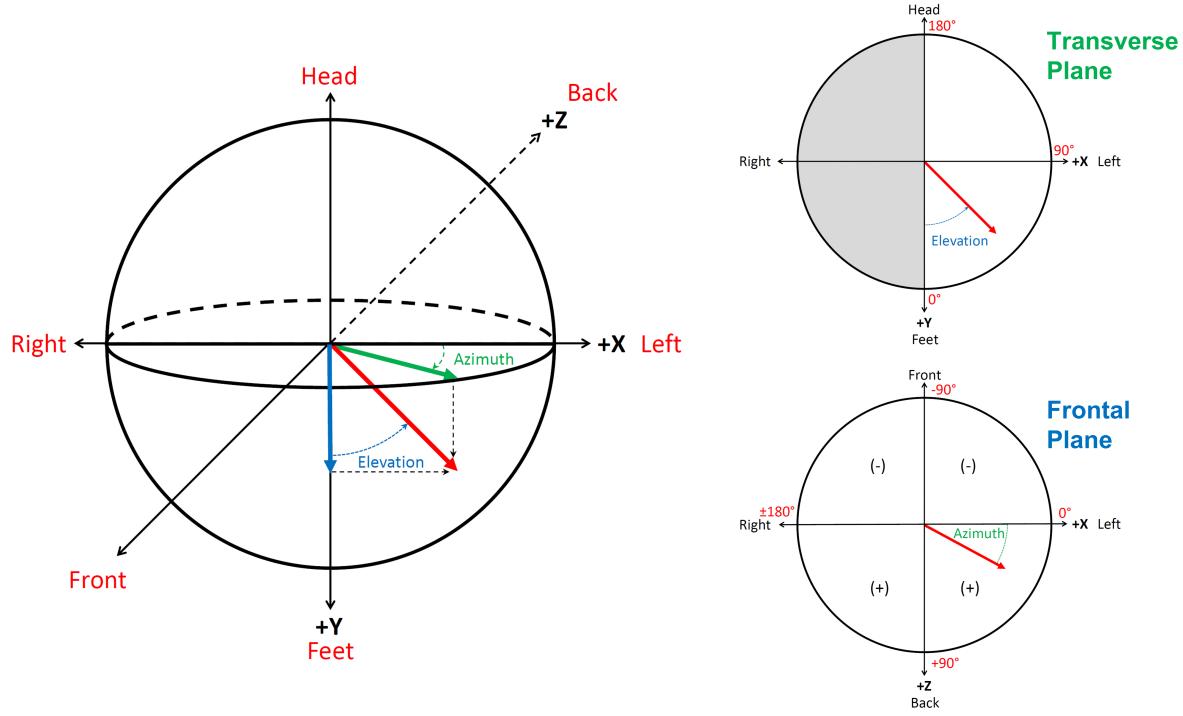
Abbreviations: ICC A - absolute agreement intraclass correlation coefficient, ICC C - consistency of agreement intraclass correlation coefficient.



Supplemental Figure 12: Example where use of both NCC and RMSE helps prevent over-calling beats with slightly different morphology due to artifact/respiration as PVCs. In this case the RMSE and NCC thresholds were set to the nominal values of 0.1 and 95%, respectively. Beats 6 and 7 have RMSE just above the threshold, but have good NCC and should not be detected as PVCs.



Supplemental Figure 13: Neural network training curves for the neural network included with BRAVEHEART for median bean annotation. The blue line represents training accuracy and the red line represents training loss. The dashed lines represent validation training/loss.



Supplemental Figure 14: Definitions of azimuth and elevation. A vector is shown in red. Elevation is the angle in the XY plane with range 0 to 180° with 0° defined as pointing towards the feet (+Y), and 180° pointing towards the head (-Y). Azimuth is the angle in the XZ plane with range 0 to $\pm 180^\circ$. Azimuth 0° points towards the left (+X), while $\pm 180^\circ$ points towards the left (-X), with positive angles oriented posteriorly and negative angles oriented anteriorly.

Supplemental Table 1: Output Variables 1 of 3

Variable	Description	Units
Standard Intervals		
qrs_int	QRS duration	ms
qt_int	QT interval	ms
QRS Area Vector		
XQ_area	Area under median X QRS complex	mV·ms
YQ_area	Area under median Y QRS complex	mV·ms
ZQ_area	Area under median Z QRS complex	mV·ms
q_area_mag	Magnitude of QRS area vector ($[XQ_area, YQ_area, ZQ_area]$)	mV·ms
q_area_az	Azimuth of QRS area vector	deg
q_area_el	Elevation of QRS area vector	deg
T Area Vector		
XT_area	Area under median X T wave	mV·ms
YT_area	Area under median Y T wave	mV·ms
ZT_area	Area under median Z T wave	mV·ms
t_area_mag	Magnitude of T-wave area vector [$XT_area, YT_area, ZT_area$]	mV·ms
t_area_az	Azimuth of T-wave area vector	deg
t_area_el	Elevation of T-wave area vector	deg
SVG		
svg_x	X component of SVG = $XQ_area + XT_area$	mV·ms
svg_y	Y component of SVG = $YQ_area + YT_area$	mV·ms
svg_z	Z component of SVG = $ZQ_area + ZT_area$	mV·ms
svg_area_mag	Magnitude of the SVG vector [svg_x, svg_z, svg_z]	mV·ms
svg_area_az	Azimuth of the SVG vector	deg
svg_area_el	Elevation of the SVG vector	deg
SAI QRST		
sai_x	Area under the absolute value of the median X QRST complex	mV·ms
sai_y	Area under the absolute value of the median Y QRST complex	mV·ms
sai_z	Area under the absolute value of the median Z QRST complex	mV·ms
sai_qrst	SAI QRST = $sai_x + sai_y + sai_z$	mV·ms
sai_vm	Area under the absolute value of the median VM QRST complex	mV·ms
QRS Peak Vector		
XQ_peak	Value of median X QRS complex at time of maximum distance from origin	mV
YQ_peak	Value of median Y QRS complex at time of maximum distance from origin	mV
ZQ_peak	Value of median Z QRS complex at time of maximum distance from origin	mV
q_peak_mag	Magnitude of peak QRS vector	mV
q_peak_az	Azimuth of peak QRS vector	mV
q_peak_el	Elevation of peak QRS vector	mV
T Peak Vector		
XT_peak	Value of median X T wave at time of maximum distance from origin	mV
YT_peak	Value of median Y T wave at time of maximum distance from origin	mV
ZT_peak	Value of median Z T wave at time of maximum distance from origin	mV
t_peak_mag	Magnitude of peak T wave vector	mV
t_peak_az	Azimuth of peak T wave vector	mV
t_peak_el	Elevation of peak T wave vector	mV
QRST Angles		
qrst_angle_area	Area QRST angle: 3D angle between area QRS and area T wave vectors	deg
qrst_angle_peak	Peak QRST angle: 3D angle between peak QRS and peak T wave vectors	deg
qrst_angle_peak_frontal	Projection of area QRST angle into frontal plane	deg
qrst_angle_area_frontal	Projection of peak QRST angle into frontal plane	deg
TCRT	Total Cosine R to T (TCRT)	mV
TCRT_angle	TCRT Angle = $\text{acos}(\text{TCRT})$	deg

Supplemental Table 2: Output Variables 2 of 3

Measurement	Description	Units
LVH		
<code>cornell_lvh_mv</code>	Cornell LVH voltage	mV
<code>sokolow_lvh_mv</code>	Sokolow-Lyon LVH voltage	mV
VCG Loop Length		
<code>vcg_length_qrs</code>	Length of QRS VCG loop	mV
<code>vcg_length_t</code>	Length of T wave VCG loop	mV
<code>vcg_length_qrst</code>	Length of QRST VCG loop = <code>vcg_length_qrs + vcg_length_t</code>	mV
VCG Loop Speed		
<code>speed_max</code>	Maximum speed across the entire VCG loop	mV/ms
<code>speed_min</code>	Minimum speed across the entire VCG loop	mV/ms
<code>speed_med</code>	Median speed across the entire VCG loop	mV/ms
<code>time_speed_max</code>	Time after QRS onset of maximum VCG speed	ms
<code>time_speed_min</code>	Time after QRS onset of minimum VCG speed	ms
<code>speed_qrs_max</code>	Maximum speed across the QRS VCG loop	mV/ms
<code>speed_qrs_min</code>	Minimum speed across the QRS VCG loop	mV/ms
<code>speed_qrs_med</code>	Median speed across the QRS VCG loop	mV/ms
<code>time_speed_qrs_max</code>	Time after QRS onset of maximum QRS speed	ms
<code>time_speed_qrs_min</code>	Time after QRS onset of minimum QRS speed	ms
<code>speed_t_max</code>	Maximum speed across the T wave loop	mV/ms
<code>speed_t_min</code>	Minimum speed across the T wave loop	mV/ms
<code>speed_t_med</code>	Median speed across the T wave loop	mV/ms
<code>time_speed_t_max</code>	Time after QRS onset of maximum T-wave speed	ms
<code>time_speed_t_min</code>	Time after QRS onset of minimum T-wave speed	ms
T Wave Morphology		
<code>vm_tpeak_time</code>	Time after QRS onset of peak of median VM Twave	ms
<code>vm_tpeak_tend_abs_diff</code>	Time difference between T wave peak and T wave end in median VM lead	ms
<code>vm_tpeak_tend_ratio</code>	Ratio between time of T wave peak and time of T wave end in median VM lead	-
Lead Morphology		
<code>[lead]_r_wave</code>	[lead] is any of the 16 leads (I, II, III, avL, avR, avF, V1-V6, X, Y, Z, VM)	mV
<code>[lead]_s_wave</code>	Magnitude of R wave on median beat of [lead]	mV
<code>[lead]_rs_wave</code>	Magnitude of S wave on median beat of [lead]	mV
<code>[lead]_rs_ratio</code>	Magnitude of entire QRS complex = <code>[lead]_r_wave + abs([lead]_s_wave</code>	-
<code>[lead]_sr_ratio</code>	Ratio of R wave to magnitude of entire QRS complex = <code>[lead]_r_wave / [lead]_rs_wave</code>	-
<code>[lead]_t_max</code>	Ratio of S wave to magnitude of entire QRS complex = <code>[lead]_s_wave / [lead]_rs_wave</code>	-
<code>[lead]_t_max_loc</code>	Maximum magnitude of T wave in [lead]	mV
	Timing of T wave maximum (after QRS onset) in [lead]	ms
VCG Loop Morphology		
<code>qrsloop_residual</code>	SVD variance from fitting QRS loop to a plane = qrs_S3^2 (0 = perfect fit)	-
<code>qrsloop_rmse</code>	RMSE for fit of QRS loop to best fit plane (0 = perfect fit)	mV
<code>qrsloop_roundness</code>	QRS loop roundness. 1 = perfect circle, larger values are increasingly elliptical	-
<code>qrsloop_area</code>	Area of QRS loop	mV
<code>qrsloop_perimeter</code>	Length of QRS loop projected into best fit plane	mV^2
<code>tloop_residual</code>	SVD variance from fitting T loop to a plane = t_S3^2 (0 = perfect fit)	-
<code>tloop_rmse</code>	RMSE for fit of T loop to best fit plane (0 = perfect fit)	mV
<code>tloop_roundness</code>	T loop roundness. 1 = perfect circle, larger values are increasingly elliptical	-
<code>tloop_area</code>	Area of T loop	mV
<code>tloop_perimeter</code>	Length of T loop projected into best fit plane	mV^2
<code>qrs_loop_normal</code>	Unit vector normal to best fit QRS loop plane	-
<code>t_loop_normal</code>	Unit vector normal to best fit T loop plane	-
<code>qrst_dihedral_ang</code>	Dihedral angle between best fit QRS loop and T loop planes	deg

Supplemental Table 3: Output Variables 3 of 3

Measurement	Description	Units
qrs.S1	1st singular value of QRS loop	—
qrs.S2	2nd singular value of QRS loop	—
qrs.S3	3rd singular value of QRS loop	—
t.S1	1st singular value of T loop	—
t.S2	2nd singular value of T loop	—
t.S3	3rd singular value of T loop	—
qrs_var_s1_total	% of total variance made up by 1st QRS singular value	%
qrs_var_s2_total	% of total variance made up by 2nd QRS singular value	%
qrs_var_s3_total	% of total variance made up by 3rd QRS singular value	%
t_var_s1_total	% of total variance made up by 1st T singular value	%
t_var_s2_total	% of total variance made up by 2nd T singular value	%
t_var_s3_total	% of total variance made up by 3rd T singular value	%

Supplemental Table 4: Results of assessing PVC detection efficacy at different cutoffs for NCC and RMSE.

NCC	RMSE	F1	Acc	NPV	PPV	Spec	Sens	% Missed PVC
80	0.05	0.959 (0.931-0.979)	0.988 (0.980-0.994)	0.989 (0.983-0.994)	0.983 (0.958-0.998)	0.997 (0.993-1.000)	0.935 (0.901-0.966)	6.5 (3-4-9.9)
85	0.05	0.967 (0.940-0.986)	0.991 (0.983-0.996)	0.993 (0.988-0.997)	0.993 (0.948-0.992)	0.974 (0.991-0.999)	0.996 (0.928-0.983)	3.9 (1.7-7.2)
90	0.05	0.962 (0.938-0.981)	0.989 (0.981-0.994)	0.995 (0.990-0.999)	0.995 (0.925-0.977)	0.952 (0.987-0.996)	0.992 (0.944-0.995)	2.8 (0.5-5.6)
95	0.05	0.892 (0.855-0.923)	0.965 (0.952-0.976)	0.997 (0.991-1.000)	0.818 (0.769-0.865)	0.980 (0.951-0.973)	0.980 (0.950-1.000)	2.0 (0-5.0)
99	0.05	0.735 (0.686-0.780)	0.901 (0.879-0.920)	0.990 (0.981-0.996)	0.602 (0.545-0.654)	0.893 (0.872-0.912)	0.945 (0.902-0.978)	5.5 (2.2-9.8)
Off	0.05	0.758 (0.710-0.804)	0.911 (0.892-0.930)	0.992 (0.984-0.998)	0.628 (0.573-0.687)	0.904 (0.884-0.923)	0.955 (0.916-0.988)	4.5 (1.2-8.4)
80	0.1	0.963 (0.938-0.982)	0.989 (0.982-0.995)	0.990 (0.985-0.995)	0.984 (0.960-0.998)	0.997 (0.993-1.000)	0.943 (0.912-0.970)	5.7 (3-0.8-8)
85	0.1	0.969 (0.943-0.987)	0.991 (0.983-0.996)	0.993 (0.988-0.997)	0.980 (0.954-0.996)	0.997 (0.992-0.999)	0.959 (0.931-0.982)	4.1 (1.8-6.9)
90	0.1	0.976 (0.949-0.993)	0.993 (0.985-0.998)	0.996 (0.990-0.999)	0.978 (0.951-0.996)	0.996 (0.992-0.999)	0.974 (0.944-0.994)	2.6 (0.6-5.6)
95	0.1	0.977 (0.952-0.994)	0.993 (0.986-0.998)	0.997 (0.992-1.000)	0.971 (0.945-0.992)	0.995 (0.990-0.999)	0.982 (0.955-1.000)	1.8 (0-0-4.5)
99	0.1	0.975 (0.949-0.993)	0.993 (0.985-0.998)	0.998 (0.994-1.000)	0.962 (0.932-0.986)	0.993 (0.988-0.998)	0.988 (0.963-1.000)	1.2 (0-0-3.7)
Off	0.1	0.978 (0.951-0.994)	0.993 (0.985-0.998)	0.998 (0.993-1.000)	0.969 (0.942-0.990)	0.995 (0.990-0.998)	0.986 (0.959-1.000)	1.4 (0-0-4.1)
80	1.5	0.935 (0.903-0.962)	0.982 (0.973-0.989)	0.982 (0.974-0.988)	0.983 (0.957-1.000)	0.997 (0.993-1.000)	0.892 (0.849-0.932)	10.8 (6.8-15.1)
85	1.5	0.951 (0.921-0.974)	0.986 (0.978-0.993)	0.987 (0.979-0.992)	0.983 (0.956-1.000)	0.997 (0.993-1.000)	0.921 (0.881-0.954)	7.9 (4.6-11.9)
90	1.5	0.963 (0.934-0.984)	0.989 (0.981-0.995)	0.990 (0.984-0.995)	0.984 (0.958-1.000)	0.997 (0.993-1.000)	0.943 (0.907-0.973)	5.7 (2.7-9.3)
95	1.5	0.964 (0.934-0.984)	0.990 (0.981-0.995)	0.991 (0.984-0.996)	0.984 (0.958-1.000)	0.997 (0.993-1.000)	0.945 (0.906-0.973)	5.5 (2.7-9.4)
99	1.5	0.966 (0.940-0.986)	0.990 (0.983-0.996)	0.991 (0.985-0.997)	0.984 (0.958-1.000)	0.997 (0.993-1.000)	0.949 (0.915-0.979)	5.1 (2.1-8.5)
Off	1.5	0.964 (0.935-0.985)	0.990 (0.981-0.996)	0.991 (0.984-0.996)	0.984 (0.958-1.000)	0.997 (0.993-1.000)	0.945 (0.908-0.977)	5.5 (2.3-9.2)
80	Off	0.954 (0.924-0.975)	0.987 (0.979-0.993)	0.988 (0.981-0.993)	0.983 (0.957-0.998)	0.997 (0.993-1.000)	0.927 (0.888-0.958)	7.3 (4.2-11.2)
85	Off	0.968 (0.939-0.986)	0.991 (0.982-0.996)	0.993 (0.987-0.997)	0.980 (0.953-0.996)	0.997 (0.992-0.999)	0.957 (0.924-0.980)	4.3 (2.0-7.6)
90	Off	0.972 (0.947-0.988)	0.992 (0.984-0.997)	0.996 (0.991-0.999)	0.967 (0.943-0.987)	0.994 (0.990-0.998)	0.976 (0.948-0.996)	2.4 (0-4-5.2)
95	Off	0.902 (0.868-0.930)	0.969 (0.958-0.978)	0.996 (0.990-0.999)	0.839 (0.792-0.880)	0.968 (0.958-0.977)	0.974 (0.943-0.996)	2.6 (0-4-5.7)
99	Off	0.512 (0.465-0.558)	0.754 (0.723-0.783)	0.975 (0.962-0.985)	0.360 (0.318-0.404)	0.731 (0.697-0.763)	0.888 (0.833-0.933)	11.2 (6.7-16.7)

ECG Diagnosis	Training N (%)	Testing N (%)	P Value
Sinus rhythm	766 (80.9)	154 (81.5)	0.85
Atrial fibrillation/flutter	85 (9.0)	19 (10.1)	0.64
Atrial pacing	27 (2.9)	5 (2.6)	0.88
Ventricular pacing	94 (9.9)	15 (7.9)	0.4
Left bundle branch block	67 (7.1)	6 (3.2)	0.046
Right bundle branch block	72 (7.6)	18 (9.5)	0.37
Nonspecific IVCD	54 (5.7)	9 (4.8)	0.61
Hemiblock	48 (5.1)	11 (5.8)	0.67
Bifascicular block	27 (2.9)	4 (2.1)	0.57
Left ventricular hypertrophy	124 (13.1)	20 (10.6)	0.34
Myocardial infarction	20 (2.1)	4 (2.1)	1.0

Supplemental Table 5: ECG features used for neural network training (N = 947) and testing (N = 189). Overall features were well matched with the exception of a borderline lower rate of left bundle branch block in the testing dataset (p=0.46).

ECG Diagnosis	N (%)
Sinus rhythm	174 (87.0)
Atrial fibrillation/flutter	16 (8.0)
Atrial pacing	1 (0.5)
Ventricular pacing	12 (6.0)
Left bundle branch block	7 (3.5)
Right bundle branch block	15 (7.5)
Nonspecific IVCD	7 (3.5)
Hemiblock	8 (4.01)
Bifascicular block	2 (1.0)
Left ventricular hypertrophy	20 (10.0)
Myocardial infarction	1 (0.5)

Supplemental Table 6: Diagnoses for the second 200 ECG neural network testing dataset.

References

- [1] B. Iglewicz and D. Hoaglin. *The ASQC Basic References in Quality Control: Statistical Techniques*, chapter Volume 16: How to Detect and Handle Outliers. 1993.
- [2] J. A. Kors, G. van Herpen, A. C. Sittig, and J. H. van Bemmel. Reconstruction of the Frank vectorcardiogram from standard electrocardiographic leads: diagnostic comparison of different methods. *Eur Heart J*, 11(12):1083–1092, Dec 1990.
- [3] S. Man, A. C. Maan, M. J. Schalij, and C. A. Swenne. prognostic information derived from the 12-lead electrocardiogram: Historical review and clinical perspective. *J Electrocardiol*, 48(4):463–475, 2015.
- [4] Paul Kligfield, Leonard S. Gettes, James J. Bailey, Rory Childers, Barbara J. Deal, E. William Hancock, Gerard van Herpen, Jan A. Kors, Peter Macfarlane, David M. Mirvis, Olle Pahlm, Pentti Rautaharju, and Galen S. Wagner. Recommendations for the standardization and interpretation of the electrocardiogram. *Circulation*, 115(10):1306–1324, 2007.
- [5] Y. Yu, X. Si, C. Hu, and J. Zhang. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Comput*, 31(7):1235–1270, 07 2019.
- [6] H. F. Stabenau, C. P. Bridge, and J. W. Waks. ECGAug: A novel method of generating augmented annotated electrocardiogram QRST complexes and rhythm strips. *Comput Biol Med*, 134:104408, 07 2021.
- [7] N. V. Thakor, J. G. Webster, and W. J. Tompkins. Estimation of QRS complex power spectra for design of a QRS filter. *IEEE Trans Biomed Eng*, 31(11):702–706, Nov 1984.
- [8] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014.
- [9] P. G. Postema and A. A. Wilde. The measurement of the QT interval. *Curr Cardiol Rev*, 10(3):287–294, Aug 2014.
- [10] Recommendations for measurement standards in quantitative electrocardiography. The CSE Working Party. *Eur Heart J*, 6(10):815–825, Oct 1985.
- [11] S. H. Zhou, E. D. Helfenbein, J. M. Lindauer, R. E. Gregg, and D. Q. Feild. Philips QT interval measurement algorithms for diagnostic, ambulatory, and patient monitoring ECG applications. *Ann Noninvasive Electrocardiol*, 14 Suppl 1:3–8, Jan 2009.
- [12] J. L. Willems, P. Arnaud, J. H. van Bemmel, P. J. Bourdillon, C. Brohet, S. Dalla Volta, J. D. Andersen, R. Degani, B. Denis, and M. Demeester. Assessment of the performance of

- electrocardiographic computer programs with the use of a reference data base. *Circulation*, 71(3):523–534, Mar 1985.
- [13] C. Zywietz and D. Celikag. Testing results and derivation of minimum performance criteria for computerized ecg-analysis. In *[1991] Proceedings Computers in Cardiology*, pages 97–100, 1991.
- [14] B. Acar, G. Yi, K. Hnatkova, and M. Malik. Spatial, temporal and wavefront direction characteristics of 12-lead T-wave morphology. *Med Biol Eng Comput*, 37(5):574–584, Sep 1999.
- [15] Peter M. Okin, Mary J. Roman, Richard B. Devereux, and Paul Kligfield. Electrocardiographic identification of increased left ventricular mass by simple voltage-duration products. *Journal of the American College of Cardiology*, 25(2):417–423, 1995.
- [16] Maurice Sokolow and Thomas P. Lyon. The ventricular complex in left ventricular hypertrophy as obtained by unipolar precordial and limb leads. *American Heart Journal*, 37(2):161–186, 1949.
- [17] H. F. Stabenau, C. Shen, L. G. Tereshchenko, and J. W. Waks. Changes in global electrical heterogeneity associated with dofetilide, quinidine, ranolazine, and verapamil. *Heart Rhythm*, 17(3):460–467, 03 2020.