# A Novel Model Based on Spatial and Morphological Domains to Predict the Origin of Premature Ventricular Contraction

Kaiyue He[1†], Jian Sun[2†], Yiwen Wang[1], Gaoyan Zhong[1] and Cuiwei Yang[1,3,4*]

[1] Department of Electronic Engineering, Fudan University, Shanghai, China, [2] Department of Cardiology, School of Medicine, Xinhua Hospital, Shanghai Jiao Tong University, Shanghai, China, [3] Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention of Shanghai, Shanghai Medical College of Fudan University, Shanghai, China, [4] Shanghai Engineering Research Center of Cardiac Electrophysiology, Shanghai, China

Pace mapping is commonly used to locate the origin of ventricular arrhythmias, especially premature ventricular contraction (PVC). However, this technique relies on clinicians' ability to rapidly interpret ECG data. To avoid time-consuming interpretation of ECG morphology, some automated algorithms or computational models have been explored to guide the ablation. Inspired by these studies, we propose a novel model based on spatial and morphological domains. The purpose of this study is to assess this model and compare it with three existing models. The data are available from the Experimental Data and Geometric Analysis Repository database in which three *in vivo* PVC patients are included. To measure the hit rate (A hit occurs when the predicted site is within 15 mm of the target) of different algorithms, 47 target sites are tested. Moreover, to evaluate the efficiency of different models in narrowing down the target range, 54 targets are verified. As a result, the proposed algorithm achieves the most hits (37/47) and fewest misses (9/47), and it narrows down the target range most, from $27.62 \pm 3.47$ mm to $10.72 \pm 9.58$ mm among 54 target sites. It is expected to be applied in the real-time prediction of the origin of ventricular activation to guide the clinician toward the target site.

Keywords: pace mapping, ventricular arrhythmias, ablation, automated algorithm, origin of PVC

## INTRODUCTION

Premature ventricular contraction (PVC) is one of the most common ventricular arrhythmias encountered in clinical practice, occurring in 1–4% of the general population (Kostis et al., 1981). Frequent and repetitive PVCs can increase the risk of arrhythmia-induced syncope, ventricular dysfunction, and sudden death (Ahn, 2013). Hitherto, catheter ablation has become an important therapy in the management of ventricular arrhythmias (Al-Khatib et al., 2018). In the last decade (from 2000 to 2012), the annual ventricular tachycardia (VT) ablation volumes have quadrupled (Hsia and Xiong, 2019). Also, multiple studies have shown that catheter ablation can be more effective in reducing arrhythmia recurrence than anti-arrhythmic drugs (Sapp et al., 2016).

It is of therapeutic importance to localize the origin of abnormal ventricular activation before catheter ablation. The localization can be done by several approaches. Activation mapping is the

most direct technique which can be applied in patients with frequent PVCs (Adams et al., 2012). Yet, it requires time-consuming intracardiac mapping by moving the catheter to different sites of the ventricles, and it can only be performed in a small number of patients who can endure a sustained VT during the whole mapping operation. Since the origin of PVC largely determines the QRS morphology of 12-lead ECG (Josephson et al., 1982), an alternative technique, known as pace mapping, can be applied in more patients by physically stimulating multiple ventricle sites until finding the site where pacing reproduces the morphology of spontaneous PVC (Kobayashi, 2018). However, this practice relies heavily on rapid and accurate manual interpretation of ECG.

In order to automatically analyze the information of pacing sites and progressively guide the clinician to the origin of PVC, several methods have been developed. One method is to train a universal model from a cohort of patients based on machine learning methods (Sapp et al., 2017; Zhou et al., 2019). However, due to anatomical and physiological variations in patients, there is a limited accuracy when a universal model is applied to a new patient. An alternative strategy is to build patient-specific prediction models. To our knowledge, some studies have used the image-based simulated ECG data to train a customized prediction model for each patient (Potse et al., 2000; Yang et al., 2018) and the domain adaptation method has newly been applied to modify the prediction model with clinical data to account for the potential errors in the simulation data (Alawad and Wang, 2019).

In addition to the image-based simulation method, some simpler but less computational models based on information of multiple pacing sites have also been investigated. Lately, the QRS integrals (QRS-Ints) of 12-lead ECG have been used to predict the 3D coordinate of the PVC origin directly (Sapp et al., 2017; Zhou et al., 2018). Besides, the relationship between distance and change in 12-lead ECG morphology has also been inspected to assist in the localization of PVC origin (Li et al., 2017, 2018; Odille et al., 2019; Dharmaprani et al., 2020).

Inspired by previous studies, in this paper, a novel model only based on the information of pacing sites is proposed and compared with three existing models [QRS-Int Model (Sapp et al., 2017), dis-E12 Model (Li et al., 2017), and dis-corr Model (Dharmaprani et al., 2020)]. We evaluated these models in three patients with PVC and found that the proposed model was slightly superior to the other three models. This method is very suitable for the location of PVC origins in non-organic heart disease.

## MATERIALS AND METHODS

### Data
The data used throughout this study is obtained from the Experimental Data and Geometric Analysis Repository (EDGAR) database (Aras et al., 2015). The data were collected during endocardial pacing from three PVC patients. The patients were consented for an add-on experimental procedure involving ventricular pacing, performed according to a protocol approved

by the ethical committee of Charles University Hospital, Prague, Czechia (Erem et al., 2014). For each patient, there is a mean of $25 \pm 6$ distinct sites of endocardial pacing with known coordinates. For each pacing site, a mean of $28 \pm 8$ ECG beats are available and a representative beat is calculated by averaging these beats. The equation is as follows:

$$V_i = \frac{1}{N} \sum_{n=1}^{N} V_i^{(n)} \tag{1}$$

where $V_i$ and $V_i^{(n)}$ are the $i$th-lead ECG signals of representative beat and beat n, respectively.

## Models
### QRS-Int Model
The QRS-Int values were proposed by Sapp et al. (2017) as predictor variables to fit the geometric coordinate system of the heart. A statistical estimate of the coordinates $\hat{x}$, $\hat{y}$, and $\hat{z}$ for any pacing site can be obtained by fitting the multiple linear regression equation with intercept. The equation is as follows:

$$\begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{y}} \\ \hat{\mathbf{z}} \end{bmatrix} = \begin{bmatrix} \hat{\alpha_0} & \hat{\alpha_1} & \cdots & \hat{\alpha_k} \\ \hat{\beta_0} & \hat{\beta_1} & \cdots & \hat{\beta_k} \\ \hat{\gamma_0} & \hat{\gamma_1} & \cdots & \hat{\gamma_k} \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ \mathbf{I_1} \\ \cdots \\ \mathbf{I_k} \end{bmatrix} \tag{2}$$

where $\hat{\alpha_i}$, $\hat{\beta_i}$, and $\hat{\gamma_i}$ are estimated regression coefficients, and $I_i$ represents the QRS-Int. To minimize the training set of required pacing sites, three optimal predictors (the initial 120-ms QRS-Int of leads III, V2, and V6) were found by exhaustive search (Sapp et al., 2017). Then, the least-square method was used to solve 4 equations ($k = 3$) to obtain the patient-specific QRS integral model (QIM). We used at least 5 pacing points to avoid matrix singularity. The best regression coefficients can be calculated by least-square regression (Sapp et al., 2017). Once the regression coefficients best fitted for the training-set data are found, they can be used for prediction of the unknown site. Here, the initial 120-ms QRS-Int values are extracted manually from the representative beat of each pacing site.

### dis-E12 Model
The E12 value proposed by Anthony et al. (Li et al., 2017) can be used to quantify the difference of 12-lead ECG between 2 pacing sites.

$$E_{12} = \sum_{i=1}^{12} \sqrt{\frac{1}{N} \sum_{j=1}^{N} (V_{i,j} - V_{i,j}^r)^2}$$

$$dis = \sqrt{(x - x^r)^2 + (y - y^r)^2 + (z - z^r)^2} \tag{3}$$

$$dis = k_1 \cdot E_{12}$$

$$\widehat{dis}^i = k_1 \cdot E_{12}(V^{ith-known}, V^{unknown})$$

where V and $V^r$ represent 2 pacing beats being compared, which are 150-ms waveforms centered on the maximum of the 12-lead composite signal (Li et al., 2017), and $\mathbf{V_{i,j}}$ and $\mathbf{V_{i,j}^r}$ represent the

voltages of one moment of the ECG. $N$ is the length of the ECG signal. $i$ ranges from 1 to 12, representing the index of 12 leads. $dis$ represents the Euclidean distance between pacing sites. Similarly, the patient-specific dis-E12 model (DEM) can be solved by origin-constrained least-square linear regression. After that, the E12 value between the unknown site and each known site is calculated and then used to estimate the corresponding distance $\widehat{dis^i}$ for i $1, 2, \cdots, m$. Finally, by minimizing the following cost function $J$, a statistical estimate of coordinates $\hat{x}$, $\hat{y}$, and $\hat{z}$ for the unknown site can be found.

$$J = \sum_{i=1}^{m}(\sqrt{(\hat{x} - x_i)^2 + (\hat{y} - y_i)^2 + (\hat{z} - z_i)^2} - \widehat{dis^i})^2 \quad (4)$$

## dis-Corr Model

The correlation coefficient (Corr) proposed by Dharmaprani et al. (2020) can be used to quantify the similarity of ECG morphology between 2 pacing sites.

$$r(X, Y) = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}}$$

$$Corr = \frac{1}{12}\sum_{i=1}^{12} r(V_i, V_i^r)$$

$$dis = \sqrt{(x - x^r)^2 + (y - y^r)^2 + (z - z^r)^2} \quad (5)$$

$$dis = k_2 \cdot (Corr - 1)$$

$$\widehat{dis^i} = k_2 \cdot (1 - Corr(V^{ith-known}, V^{unknown}))$$

where $r(X, Y)$ represents the Pearson correlation coefficient between time series X and Y. Moreover, the Corr value is the average result of 12 leads. Similarly, the patient specific dis-Corr model (DCM) can be solved by constrained least-square linear regression, and the Corr value between the unknown site and each known site can also be calculated and transformed into the estimated distance $\widehat{dis^i}$ for i $1, 2, \cdots, m$. Finally, by minimizing the cost function $J$ presented in Eq. (4), a statistical estimate of coordinates $\hat{x}$, $\hat{y}$, and $\hat{z}$ for the unknown site can be found.

## dp-dw Model

In this study, we observed a phenomenon that there are some connections between waveform morphology and physical position. **Figures 1**, **2** show two examples based on simulation data and real data, respectively. The simulation data were generated by an isotropic ventricular simulation model with electric conduction rate of 0.7 m /s (Schulze et al., 2015). As can be observed from **Figure 1**, points 1, 2, 3, 4, and 5 are almost on the same line, while points 3, 6, and 7 are almost on another line. **Figure 1B** shows the splicing signal of 12-lead ECG in accordance with positions, and **Figure 1C** shows the waveform difference between each pair of positions. It seems that the waveform differences on the same line are more similar, while those on different lines are less similar. For example,

s31, s32, s34, and s35 are similar with each other and so are s36 and s37, but s31 and s37 are less similar. Then the real data are extracted from the first patient. As can be seen from **Figure 2**, points LVP11, LVP1, and LVP20 are almost on the same line, while LVP18 and LVP4 are almost on another line. Besides, the two lines are nearly parallel. Similarly, we observed the similarity of waveform differences on the same line. We also observed the similarity of waveform differences between parallel lines.
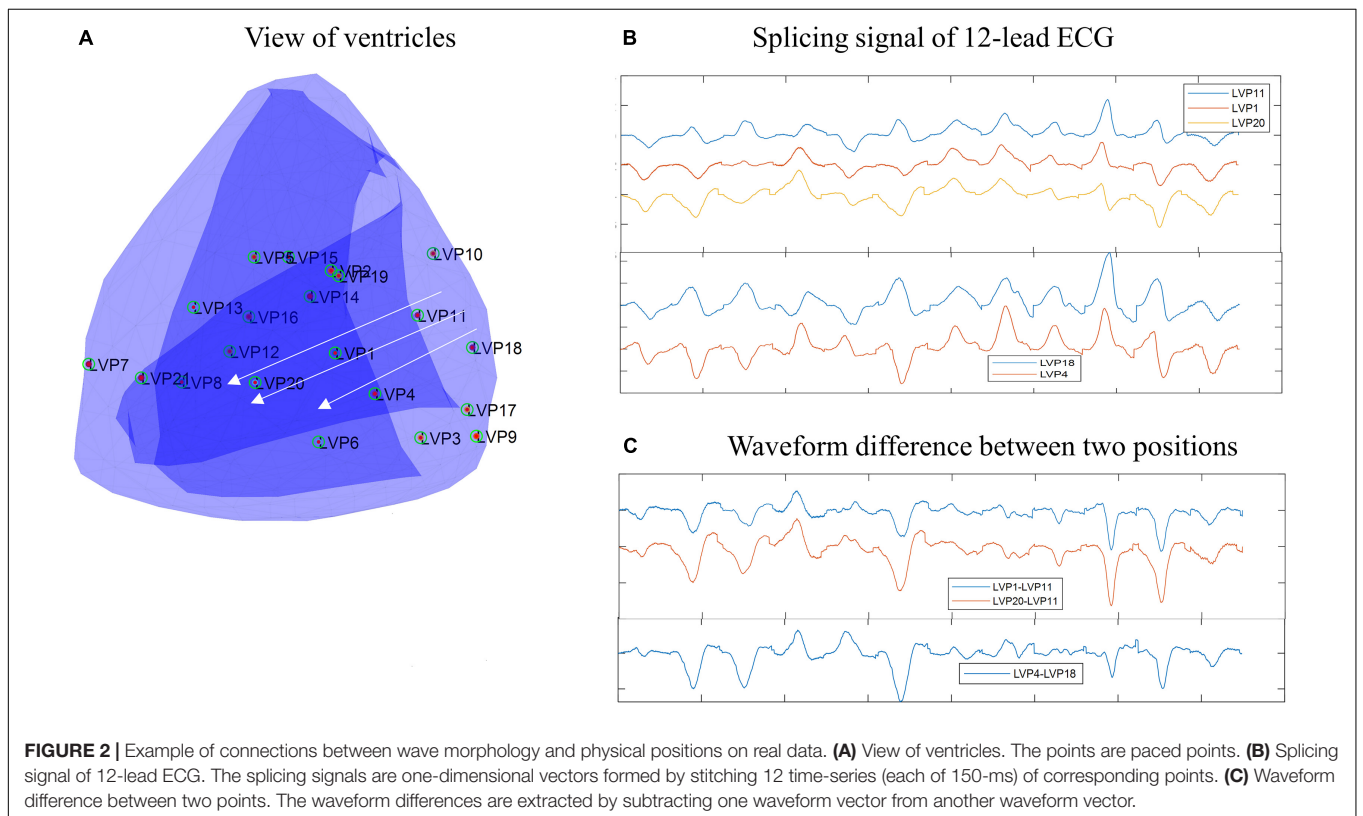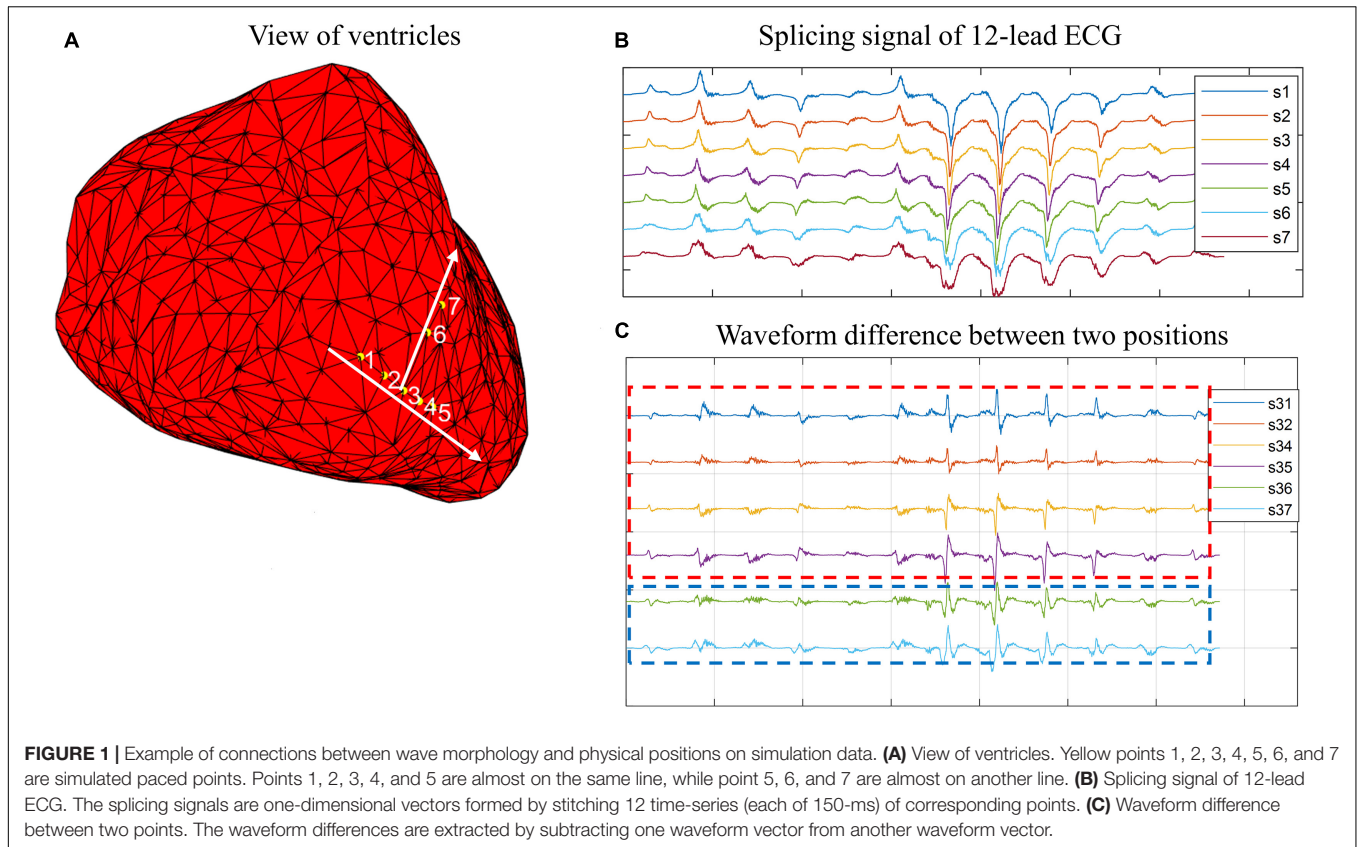
The above phenomenon may be explained by the theory of electrocardiographic dipoles. During ventricular depolarization, electric dipoles can be formed between depolarized and non-depolarized regions, and the integrated vector of all dipoles can be recorded by 12 leads from different positions and directions. The recorded voltage on each lead at one moment is related to the distances between the recording electrode and the electric dipoles, and it is also related to the cosine angles formed by the orientation of the lead axis and the directions of myocardial depolarization. When a ventricular premature occurs, the depolarization wave spreads from the earliest excitation point to all sides, and the directions of electric dipoles are the same as the directions of myocardial depolarization. When the earliest excitation point moves along a certain direction, the electric dipoles will change most in the same direction, which may lead to more obvious waveform changes in leads parallel to the direction and less obvious waveform changes in leads perpendicular to the direction. Therefore, there might be a certain relationship between the waveform changes of 12-lead ECG and the position changes of the earliest excitation point.
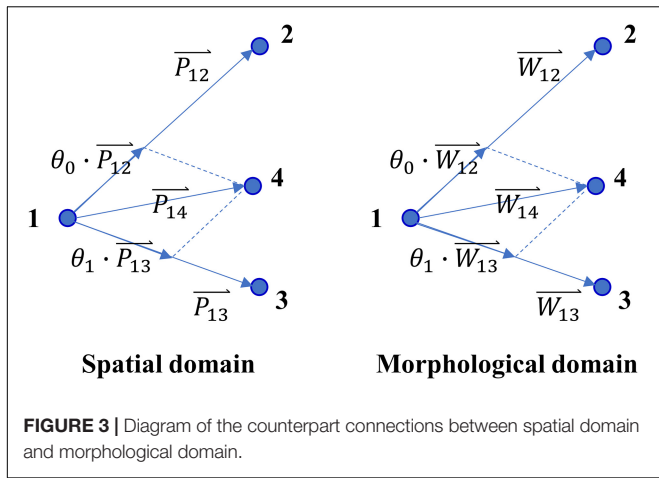
Based on these observations, we proposed a novel prediction model based on the assumption that there are some counterpart connections between the spatial domain and the morphological domain. As **Figure 3** shows, $\vec{P}_{ij}$ and $\vec{W}_{ij}$ represent the vector of position difference (dp) and the vector of waveform difference (dw) between point i and point j, respectively. Here, $\vec{W}_i$ represents a one-dimensional vector formed by stitching 12 time-series (each of 150-ms) of pacing site i together, and $\vec{W}_{ij}$ is obtained by subtracting $\vec{W}_i$ from $\vec{W}_j$. Supposing Eq. (6) holds in the spatial domain, and then Eq. (7) holds in the morphological domain, and vice versa.

$$\vec{P}_{14} = \theta_0 \cdot \vec{P}_{12} + \theta_1 \cdot \vec{P}_{13} \quad (6)$$

$$\vec{W}_{14} = \theta_0 \cdot \vec{W}_{12} + \theta_1 \cdot \vec{W}_{13} \quad (7)$$

**Figure 4** illustrates the establishment and prediction process of the dp-dw model (DDM). As shown in the figure, points 1, 2, and 3 are the known sites whose 12-lead ECG signals and physical locations are known by us, while point O is the unknown site whose 12-lead ECG information is known by us, but its physical location needs to be estimated by algorithm. Of course, the actual physical location of point O is known, but we pretend not to know that. Moreover, we use the waveform difference between point

**FIGURE 1 |** Example of connections between wave morphology and physical positions on simulation data. **(A)** View of ventricles. Yellow points 1, 2, 3, 4, 5, 6, and 7 are simulated paced points. Points 1, 2, 3, 4, and 5 are almost on the same line, while point 5, 6, and 7 are almost on another line. **(B)** Splicing signal of 12-lead ECG. The splicing signals are one-dimensional vectors formed by stitching 12 time-series (each of 150-ms) of corresponding points. **(C)** Waveform difference between two points. The waveform differences are extracted by subtracting one waveform vector from another waveform vector.



**FIGURE 2 |** Example of connections between wave morphology and physical positions on real data. **(A)** View of ventricles. The points are paced points. **(B)** Splicing signal of 12-lead ECG. The splicing signals are one-dimensional vectors formed by stitching 12 time-series (each of 150-ms) of corresponding points. **(C)** Waveform difference between two points. The waveform differences are extracted by subtracting one waveform vector from another waveform vector.

**FIGURE 3** | Diagram of the counterpart connections between spatial domain and morphological domain.

O and its adjacent points to estimate its location. The vectors of dp and dw between each pair of known sites are calculated as $[\triangle P]$ $\left[\overrightarrow{P_{12}}, \overrightarrow{P_{13}}\right]$ and $[\triangle W]$ $\left[\overrightarrow{W_{12}}, \overrightarrow{W_{13}}\right]$, and the vector of dw between the unknown site and each known site is calculated as $\left[\triangle W^*\right]$ $\left[\overrightarrow{W_{1O}}, \overrightarrow{W_{2O}}, \overrightarrow{W_{3O}}\right]$. Analogously, the transfer matrix $[\theta]$ mapping $[\triangle W]$ to $\left[\widehat{\triangle W^*}\right]$ can be found by least-square regression [Eq. (8)]. Then, by applying the same transfer matrix $[\theta]$ to the spatial domain, the estimated vector of dp between the unknown site and each known site can be calculated as shown in Eq. (9). Finally, a statistical estimate of coordinates $\hat{x}$, $\hat{y}$, and $\hat{z}$ for the unknown site can be calculated as Eq. (10).

$$[\widehat{\triangle W^*}] = [\triangle W] \times [\theta]$$

$$[\theta] = ([\triangle W]^T \times [\triangle W])^{-1} \times [\triangle W]^T \times [\triangle W^*] \quad (8)$$

$$[\widehat{\triangle P^*}] = [\triangle P] \times [\theta]$$

$$[\widehat{\triangle P^*}] = [\overrightarrow{P_{1O}}, \widehat{\underset{2O}{\overrightarrow{P}}}, \widehat{\underset{3O}{\overrightarrow{P}}}] \quad (9)$$

$$\begin{bmatrix} \widehat{x} \\ \widehat{y} \\ \widehat{z} \end{bmatrix} = \frac{1}{3} \sum_{i=1}^{3} \left( P_i + \widehat{\underset{iO}{\overrightarrow{P}}} \right) \quad (10)$$

**Figure 5** shows an example of using DDM to predict the PVC origin. In **Figure 5A**, the red point represents the unknown site, the green points represent the known sites, and the yellow point represents the predicted position. **Figure 5B** shows the waveform differences (dws) between the known sites, and **Figure 5D** shows the dws between the known sites and the unknown site. By using least-square regression, the estimated dws between the known sites and the unknown site can be transformed from the dws between the known sites, as **Figure 5C** shows. Finally, by applying the same transfer matrix $[\theta]$ in the spatial domain, the position differences (dps) between the known sites and the unknown site could be calculated. And by executing Eq.(10), the predicted position was obtained.

## Emulation of Clinical Protocols
### Target Site Selection

Two target site selection schemes were adopted for different purposes. First, in order to evaluate the hit rate (a hit occurs when the predicted site is within 15 mm of the target), the target site is defined as the site with at least 5 adjacent sites which are greater than 15 mm and less than 35 mm away from it, and a total of 47 pacing sites meet the conditions. Secondly, in order to evaluate the efficiency of different models in narrowing down the target range, the target site is defined having at least 5 adjacent sites within the range of 35 mm of it, and a total of 54 pacing sites meet the requirement. Once a target site is selected, its adjacent sites that meet the corresponding definition serve as potential modeling sites. In this study, the modeling sites are those whose physical locations and corresponding 12-lead ECG signals are known, while the target sites are those whose 12-lead ECG signals are known, but their physical locations need to be estimated by the algorithm.



**FIGURE 4** | Establishment and prediction process of DDM.

**FIGURE 5 |** An example of using DDM to predict the PVC origin. **(A)** View of ventricles. The green points represent known sites, the red point represents the unknown site, and the yellow site represents the predicted site. **(B)** Vector of waveform differences between known sites. For example, W12 is extracted by subtracting the waveform of point 1 from the waveform of point 2. **(C)** Mapping process. By least-square regression, the estimated waveform differences between the known sites and the unknown sites can be transformed from the waveform differences between the known sites. Then by applying the same transfer matrix in the spatial domain, the position differences between the known sites and the unknown site can be calculated. **(D)** Vector of waveform difference between known sites and unknown site. The ground-truth waveform differences are shown with solid lines, while the estimated waveform differences are shown with dotted lines.



**FIGURE 6 |** Flowchart of modeling and prediction.

## Modeling and Prediction

**Figure 6** shows the flowchart of the process of modeling and prediction. It can be divided into the following six steps.

Step 1: Initialize modeling sites. The 3 or 5 farthest unused potential modeling sites (3 for DEM, DCM, and DDM, and 5 for QIM) from the target site are selected as initial modeling sites, and they will be removed from the list of unused potential modeling sites. For example, supposing point 0 is selected as the target site, and its adjacent points 1, 2, 3, 4, 5, and 6 match the definition of potential modeling sites. Therefore, the initial list of unused potential modeling sites is (Kostis et al., 1981; Adams et al., 2012; Ahn, 2013; Sapp et al., 2016; Al-Khatib et al.,

2018; Hsia and Xiong, 2019). We first pick out three farthest modeling sites to predict the coordinates of point 0. Assuming points 1, 2, and 3 are selected, then we will remove them from the list of potential modeling sites that have never been used, so as not to select the duplicate modeling sites next time. Hence, the list will be updated to Adams et al. (2012), Sapp et al. (2016), Hsia and Xiong (2019).

Step 2: Train the models mentioned above to predict the target site.

Step 3: Termination judgment. If there are no unused potential modeling sites left, terminate. If the predicted site hits the target, terminate.

Step 4: Evaluate predicted site. If the predicted site is outside 35 mm of the target site, it is not credible, skip to step 6.

Step 5: Pick out a nearest site from the unused potential modeling sites if it is within 15 mm of the predicted site, then remove it from the list of unused potential modeling sites and skip to step 2. Otherwise, there is no unused potential modeling site that can replace the predicted site, turn to the next step.

Step 6: Pick out a site that is farthest from the geometric center of current modeling sites to obtain as much spatial information as possible, and remove it from the list of unused potential modeling sites and skip to step 2.

## RESULTS

### Hits and Misses

A total of 47 target sites are used to evaluate the hit rate of four models. As **Figure 7A** shows, the proposed DDM presents with the most hits (37/47), the fewest misses (9/47), and one early termination. Here an early termination means that the reduction of estimated error is interrupted by lack of potential modeling sites. Then, slightly inferior to DDM, DEM performs with 35 hits, 11 misses, and one early termination. Finally, inferior to DDM and DEM, QIM and DCM achieve with 31 hits, 14 misses, one early termination and 29 hits, 18 misses, and no early termination, respectively. In addition, when the number of modeling sites is 5, DDM has much more hits than other models.

### Estimated Error

**Figure 7B** presents the trend of estimated errors of four models with the increase of modeling sites. It must be noted that for each target site, the estimated error remains unchanged after minimization. As can be observed from the figure, with the increase of modeling sites, the estimated errors of four models tend to decrease, especially when the number of modeling sites is less than 8 when most of the samples remained non-minimization (see **Figure 7A**). In terms of decline velocity of estimated error, DCM and DDM perform better than QIM and DEM when number of modeling sites is less than 5. Also, in terms of final estimated error, DEM and DDM perform better than QIM and DCM.

### Reduced Distance

The reduced distance is equal to the minimum distance between the modeling sites and the target site minus the estimated error, and a positive reduced distance indicates a reduction in the unknown range of the target by modeling and prediction. **Figure 7C** shows the reduced distances of four models with the increase of modeling sites. For each number of modeling sites, samples that have reached the minimum estimated error are not counted. As the figure shows, for different numbers of modeling sites, the mid-values of reduced distances of four models are almost positive, indicating that four models tend to reduce the unknown range of the target. When the number of modeling sites is 5, QIM has the largest mid-value of reduced distances and the corresponding hit rate also rises rapidly (see **Figure 7A**). However, due to the cumulative reduction of distances in the previous two rounds, DDM still has the highest hit rate.
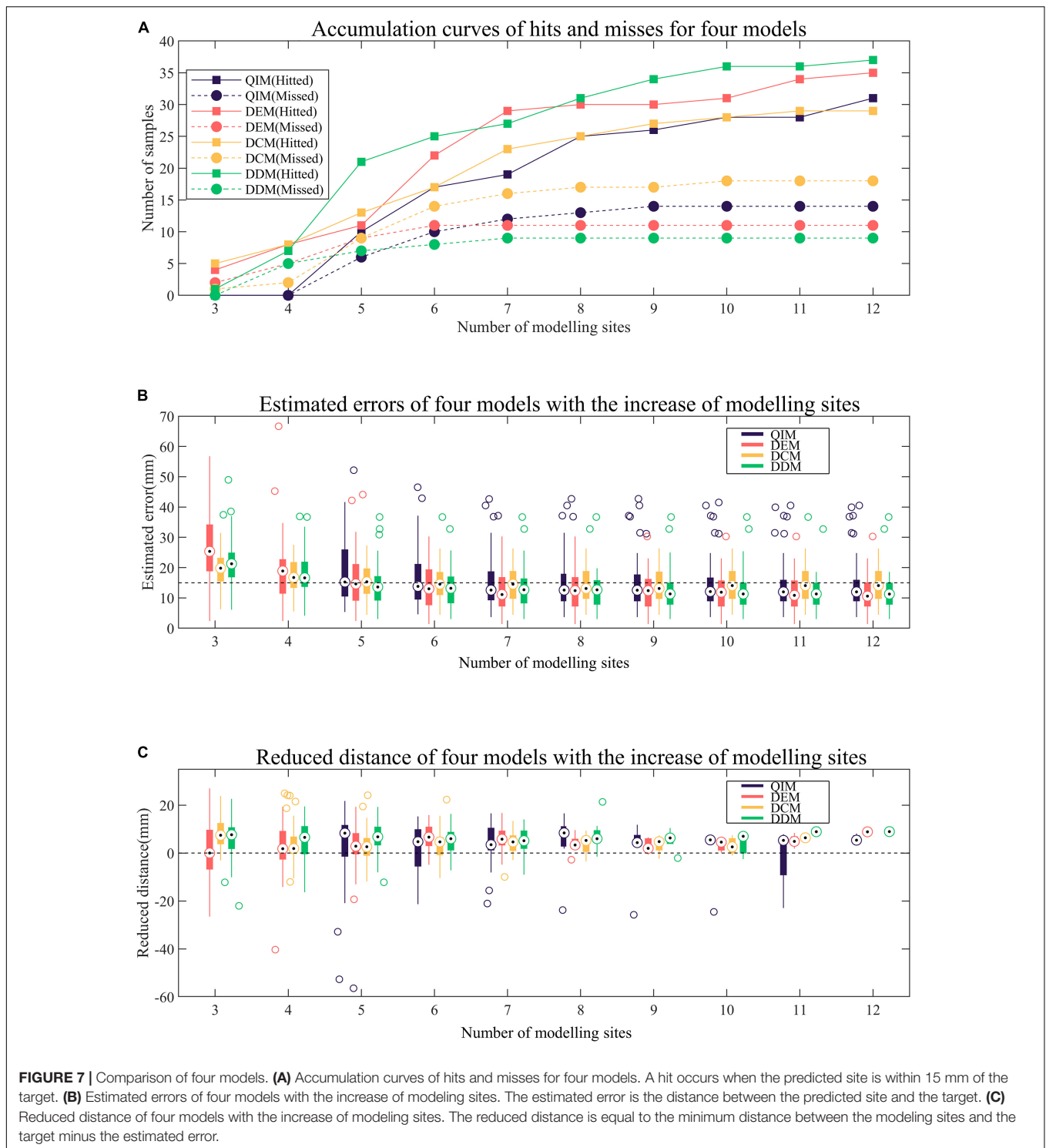
### Target Range

A total of 54 target sites are used to evaluate the efficiency of different models in narrowing down the target range. **Figure 8** shows two examples of the predicted sites of four models with the increase of modeling sites. The initial target range (marked in orange in **Figure 8**) is defined by the maximum radius of adjacent sites, and the final target range is defined by the minimum estimated error. As the figure shows, the first target site has an initial radius of 32.92 mm; after modeling and prediction, the radius is reduced to 8.51, 3.34, 9.86, and 5.19 mm with 4 models, respectively. Similarly, the radius of the second target is reduced from 33.13 to 8.80, 9.34, 10.98, and 7.39 mm with 4 models, respectively.

**Table 1** lists the statistical results of 54 target sites. Among the four models, DDM narrows down the target range most, from $27.62 \pm 3.47$ mm to $10.72 \pm 6.22$ mm, and DEM uses the fewest modeling sites ($5.98 \pm 2.49$) to minimize the target range. In addition, t tests show that the estimation errors of DCM and DDM have a significant difference ($P = 0.046$), and the numbers of modeling sites of QIM and DEM, QIM and DDM have a significant difference ($P = 0.007$, $P = 0.045$), indicating that DCM has the worst estimated error and QIM used the most modeling sites.

## DISCUSSION

This work proposed a novel model for the localization of PVC target sites based on the mapping between the spatial domain and the morphological domain. In our study, the pacing sites are not as adjacent as those generated by clinical pace mapping, so we selected modeling sites from a larger range to predict the target site step by step. For inexperienced doctors, the results obtained by our method may provide a reference location, so that they can simply determine the most likely ablation site as soon as possible and shorten the mapping procedure. We compared our model with three existing models and found that the proposed model was slightly superior to other models by achieving the most hits, the smallest estimated errors, and the biggest reduced distances. Especially when the number of modeling sites is small, the advantages of our model are more obvious. By observation of **Figure 7**, it can be found that the proposed DDM tends to have more hits,

**FIGURE 7 |** Comparison of four models. **(A)** Accumulation curves of hits and misses for four models. A hit occurs when the predicted site is within 15 mm of the target. **(B)** Estimated errors of four models with the increase of modeling sites. The estimated error is the distance between the predicted site and the target. **(C)** Reduced distance of four models with the increase of modeling sites. The reduced distance is equal to the minimum distance between the modeling sites and the target minus the estimated error.

smaller estimated errors, and bigger reduced distances than the other methods when the number of modeling sites is less than 6.

Then, as can be observed from **Table 1**, considering the minimum estimated error, DDM, and DEM perform better than QIM and DCM. Compared with DDM and DEM, QIM only uses

the information of three-lead ECG, which may account for its less satisfying result. Though DCM also uses full information of 12-lead ECG, according to reference (Li et al., 2017), in contrast to Corr, E12 theoretically has no upper limit and, therefore, can provide better quantification of the morphology difference than Corr.
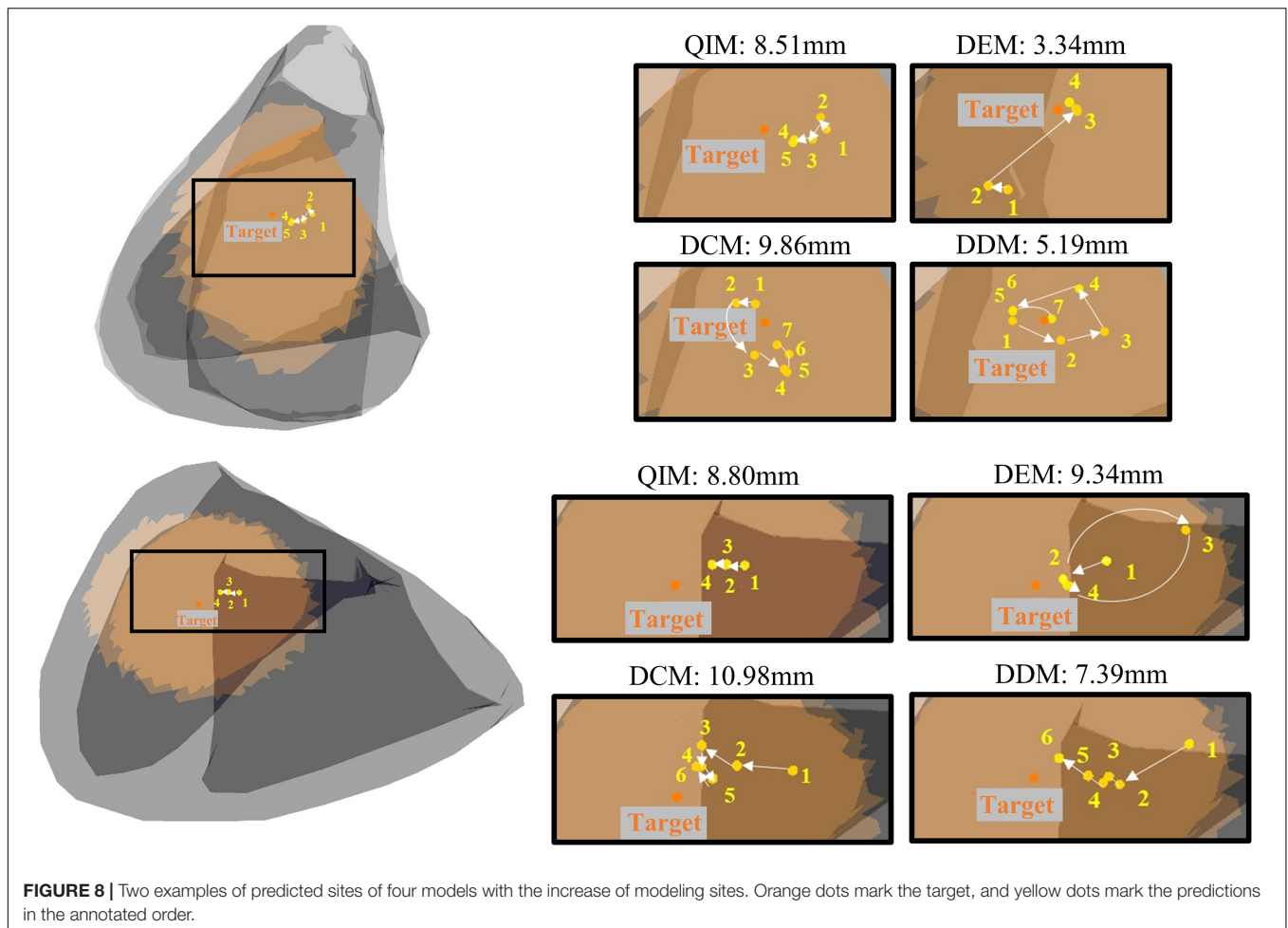
**FIGURE 8 |** Two examples of predicted sites of four models with the increase of modeling sites. Orange dots mark the target, and yellow dots mark the predictions in the annotated order.

In addition, there is a certain relationship between DDM and DEM. In essence, DEM is to establish a scalar model through the relationship of the module length between dp and dw, while DDM directly uses the relationship between dp and dw to build a vector model. When the number of modeling sites is less than 5, the prediction effect of DEM is worse than that of DDM, which is likely due to the lack of direction information. However, when the number of modeling sites increases, the

**TABLE 1 |** Comparison of four models in narrowing down the target range.

| Models | Radius of neighboring sites (mm) | Estimated error (mm) | Number of modeling sites used |
|---|---|---|---|
| QIM | Mean: 27.62 | Mean: 12.41 | Mean: 7.28[*2, *3] |
|  | Std: 3.47 | Std: 8.05 | Std: 2.43 |
|  | Mid: 27.53 | Mid: 10.05 | Mid: 7 |
| DEM | Mean: 27.62 | Mean: 11.08 | Mean: **5.98[*2]** |
|  | Std: 3.47 | Std: 6.03 | Std: 2.49 |
|  | Mid: 27.53 | Mid: 9.71 | Mid: **5.5** |
| DCM | Mean: 27.62 | Mean: 13.03[*1] | Mean: 6.80 |
|  | Std: 3.47 | Std: 5.67 | Std: 2.47 |
|  | Mid: 27.53 | Mid: 12.21 | Mid: 6 |
| DDM | Mean: 27.62 | Mean: **10.72[*1]** | Mean: 6.37[*3] |
|  | Std: 3.47 | Std: 6.22 | Std: 2.22 |
|  | Mid: 27.53 | Mid: **9.58** | Mid: 6 |

*1, *2, and *3 represent statistically significant differences between groups.

**TABLE 2 |** Summary comparison of four models.

| Model | Principles | Properties | Performance |
|---|---|---|---|
| QIM | Using QRS-Ints as predictors | Less computation, using information of 3-lead ECG and containing direction information | Second least hits and second least reduced target range |
| DEM | Based on the relationship between distance and morphology difference | More computation, using information of 12-lead ECG and lack of direction information | Second most hits and second most reduced target range |
| DCM |  |  | Least hits and least reduced target range |
| DDM | Based on the mapping between spatial domain and morphological domain | Moderate computation, using information of 12-lead ECG and containing direction information | Most hits and most reduced target range |

lack of direction information is gradually compensated by more and more complete distance network between points, and DEM achieves a similar result to DDM.

Finally, in terms of computational complexity, QIM, and DDM are simpler since the fitted models can be used for prediction directly, while DEM and DCM are more complex due to the additional search for optimal solution that can minimize the cost function $J$. **Table 2** lists a summary comparison of four models.

However, there still exist some limitations in this study. First, the way of picking out modeling sites is relatively random. Theoretically, the next modeling site should be the predicted result if it is reliable; otherwise, the next modeling site is selected by the doctor. Limited by the actual distribution of pacing sites, we take the second place and replace the predicted site with the nearest one among the unused potential modeling sites, which may cause the randomness in modeling site selection due to the different prediction results of four models, for example, when we design a fixed modeling site selection scheme, in which we choose a furthest site from the remaining potential sites in each round. Consequently, QIM, DEM, DCM, and DDM achieve 32, 35, 28, and 33 hits, respectively, indicating that the way of modeling site selection has a certain impact on the research results and the proposed DDM is more suitable for selecting modeling sites by referring to the predicted positions. Because collecting multiple-pace data in clinical practice will increase the risk of patients during operation, this kind of data is difficult to obtain. Therefore, we mainly used the data in the open database collected from three volunteers provided by Charles University in accordance with the strict experimental process. From the perspective of the number of individual patients, our sample is still relatively small, but the total number of test sites used in this paper is relatively large. In the future research, we can also consider the application of four methods to animal experimental data or retrospective clinical data analysis.

## CONCLUSION

To conclude, it is a desirable goal to develop an automated algorithm for the localization of PVC origins. This work provided a novel solution based on the mapping between spatial domain and morphological domain. It performs better with fewer modeling points and is expected to be used to predict the origin of ventricular activation in real-time and guide clinicians to focus on ablation targets.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of Charles University Hospital, Prague, Czechia. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

KH and JS: conceptualization, formal analysis, and methodology. CY: resources, supervision, and project administration. KH, YW, and GZ: software and visualization. KH: writing—original draft preparation. CY and GZ: writing—review and editing. KH and YW: revising and correcting. JS, KH, and CY: clinical interpretation and discussion of findings and their relevance. All authors contributed to the article and approved the submitted version.

## FUNDING

## REFERENCES

Adams, J. C., Srivathsan, K., and Shen, W. K. (2012). Advances in management of premature ventricular contractions. *J. Interv. Card. Electrophysiol.* 35, 137–149. doi: 10.1007/s10840-012-9698-x

Ahn, M. S. (2013). Current concepts of premature ventricular contractions. *J. Lifestyle Med.* 3, 26–33.

Alawad, M., and Wang, L. (2019). Learning domain shift in simulated and clinical data: localizing the origin of ventricular activation from 12-lead electrocardiograms. *IEEE Trans. Med. Imaging* 38, 1172–1184. doi: 10.1109/tmi.2018.2880092

Al-Khatib, S. M., Stevenson, W. G., Ackerman, M. J., Bryant, W. J., Callans, D. J., Curtis, A. B., et al. (2018). 2017 AHA/ACC/HRS guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: a report of the American College of Cardiology/American Heart Association Task Force on clinical practice guidelines and the heart rhythm society. *Circulation* 138, e272–e391. doi: 10.1161/cir.0000000000000549

Aras, K., Good, W., Tate, J., Burton, B., Brooks, D., Coll-Font, J., et al. (2015). Experimental data and geometric analysis repository-EDGAR. *J. Electrocardiol.* 48, 975–981. doi: 10.1016/j.jelectrocard.2015.08.008

Dharmaprani, D., Lahiri, A., Ganesan, A. N., Kyriacou, N., and McGavigan, A. D. (2020). Comparative spatial resolution of 12-lead electrocardiography and an automated algorithm. *Heart Rhythm* 17, 324–331. doi: 10.1016/j.hrthm.2019.08.029

Erem, B., Coll-Font, J., Orellana, R. M., Stovicek, P., and Brooks, D. (2014). Using transmural regularization and dynamic modeling for noninvasive cardiac potential imaging of endocardial pacing with imprecise thoracic geometry. *IEEE Trans. Med. Imaging* 33, 726–738. doi: 10.1109/TMI.2013.2295220

Hsia, H. H., and Xiong, N. (2019). Mapping and ablation of ventricular arrhythmias in cardiomyopathies. *Card. Electrophysiol. Clin.* 11, 635–655. doi: 10.1016/j.ccep.2019.08.005

Josephson, M. E., Waxman, H. L., Cain, M. E., Gardner, M. J., and Buxton, A. E. (1982). Ventricular activation during ventricular endocardial pacing. II. Role of

pace-mapping to localize origin of ventricular tachycardia. *Am. J. Cardiol.* 50, 11–22. doi: 10.1016/0002-9149(82)90003-0

Kobayashi, Y. (2018). Idiopathic ventricular premature contraction and ventricular tachycardia: distribution of the origin. diagnostic algorithm, and catheter ablation. *J. Nippon Med. Sch.* 85, 87–94. doi: 10.1272/jnms.2018_85-14

Kostis, J. B., McCrone, K., Moreyra, A. E., Gotzoyannis, S., Aglitz, N. M., Natarajan, N., et al. (1981). Premature ventricular complexes in the absence of identifiable heart disease. *Circulation* 63, 1351–1356. doi: 10.1161/01.cir.63.6.1351

Li, A., Davis, J. S., Grimster, A., Wierwille, J., Herold, K., Morgan, D., et al. (2018). Proof of concept study of a novel pacemapping algorithm as a basis to guide ablation of ventricular arrhythmias. *Europace* 20, 1647–1656. doi: 10.1093/europace/euy024

Li, A., Davis, J. S., Wierwille, J., Herold, K., Morgan, D., Behr, E., et al. (2017). Relationship between distance and change in surface ECG morphology during pacemapping as a guide to ablation of ventricular arrhythmias: implications for the spatial resolution of pacemapping. *Circ. Arrhythm. Electrophysiol.* 10:e004447. doi: 10.1161/circep.116.004447

Odille, F., Battaglia, A., Hoyland, P., Sellal, J. M., Voilliot, D., de Chillou, C., et al. (2019). Catheter treatment of ventricular tachycardia: a reference-less pace-mapping method to identify ablation targets. *IEEE Trans. Biomed. Eng.* 66, 3278–3287. doi: 10.1109/tbme.2019.2903631

Potse, M., Linnenbank, A. C., Peeters, H. A., SippensGroenewegen, A., and Grimbergen, C. A. (2000). Continuous localization of cardiac activation sites using a database of multichannel ECG recordings. *IEEE Trans. Biomed. Eng.* 47, 682–689. doi: 10.1109/10.841340

Sapp, J. L., Bar-Tal, M., Howes, A. J., Toma, J. E., El-Damaty, A., Warren, J. W., et al. (2017). Real-time localization of ventricular tachycardia origin from the 12-lead electrocardiogram. *JACC Clin. Electrophysiol.* 3, 687–699. doi: 10.1016/j.jacep.2017.02.024

Sapp, J. L., Wells, G. A., Parkash, R., Stevenson, W. G., Blier, L., Sarrazin, J. F., et al. (2016). Ventricular tachycardia ablation versus escalation of antiarrhythmic drugs. *N. Engl. J. Med.* 375, 111–121. doi: 10.1056/NEJMoa1513614

Schulze, W. H., Potyagaylo, D., Schimpf, R., Papavassiliu, T., Tülümen, E., Rudic, B., et al. (2015). "A simulation dataset for ECG imaging of paced beats with models for transmural, endo-and epicardial and pericardial source imaging," in *Proceedings of the 1st First Meeting of the Consortium for EGI Imaging*, Bad Herrenalp.

Yang, T., Yu, L., Jin, Q., Wu, L., and He, B. (2018). Localization of origins of premature ventricular contraction by means of convolutional neural network from 12-lead ECG. *IEEE Trans. Biomed. Eng.* 65, 1662–1671. doi: 10.1109/tbme.2017.2756869

Zhou, S., AbdelWahab, A., Sapp, J. L., Warren, J. W., and Horáček, B. M. (2019). Localization of ventricular activation origin from the 12-lead ECG: a comparison of linear regression with non-linear methods of machine learning. *Ann. Biomed. Eng.* 47, 403–412. doi: 10.1007/s10439-018-02168-y

Zhou, S., Sapp, J. L., AbdelWahab, A., Šťovíček, P., and Horáček, B. M. (2018). Localization of ventricular activation origin using patient-specific geometry: preliminary results. *J. Cardiovasc. Electrophysiol.* 29, 979–986. doi: 10.1111/jce.13622

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.