On the P-wave model of a single electrocardiogram lead

Article *in* Seminars in Cardiovascular Medicine · January 2021

DOI: 10.2478/semcard-2021-0001

5 authors, including:

Viktor Skorniakov
Vilnius University
34 PUBLICATIONS **123** CITATIONS
SEE PROFILE

Petras Navickas
Vilnius University
15 PUBLICATIONS **68** CITATIONS
SEE PROFILE

Albinas Stankus
Centre for Innovative Medicine
50 PUBLICATIONS **93** CITATIONS
SEE PROFILE

sciendo
e-ISSN 1822-7767

Original paper

On the P-wave model of a single electrocardiogram lead

Viktor Skorniakov [a,*], Antanas Mainelis [a], Petras Navickas [b,c], Germanas Marinskis [b,c], Albinas Stankus [d]

[a] *Faculty of Mathematics and Informatics, Vilnius University, Vilnius, Lithuania*
[b] *Faculty of Medicine, Vilnius University, Vilnius, Lithuania*
[c] *Vilnius University Hospital Santaros Klinikos, Centre of Cardiology and Angiology, Vilnius, Lithuania*
[d] *State Research Institute Centre for Innovative Medicine, Vilnius, Lithuania Vilnius, Lithuania*

Received 8 June 2020; accepted 9 November 2020

Summary

We describe a parametric model for the P-wave of a single electrocardiogram (ECG) lead trajectory. Though previously met in the bioengineering literature, the model was not treated in a complete parametric fashion. The paper fills the gap by making use of both frequentist and Bayesian approaches. Supporting real data example is provided. Further potential applications are also discussed.

Seminars in Cardiovascular Medicine 2021; 27:1–11

Keywords: electrocardiography, P-wave, statistical modeling

1. Introduction

Electrocardiogram (ECG) is a well-established tool in daily cardiological practice. The list of its applications includes usage for interpretation of the cardiac rhythm, detection of myocardial ischemia and infarction, conduction system abnormalities, preexcitation, long QT syndromes, atrial abnormalities, ventricular hypertrophy, pericarditis as well as other conditions [1]. An important thing to note is that particular applications usually are tied to some specific ECG segment the whole set of which is depicted in Fig. 1 and reflects different phases of the cardiac cycle. For example, diagnosis of myocardial infarction (MI) is based on ST-segment whereas that of atrial fibrillation – on (absence of) a P-wave.

Therefore, it is not surprising that the literature is abundant of items dealing with analysis and modeling of data coming from some distinct ECG segment. In what follows, by analysis and modeling we reference two distinct branches of ongoing research. To modeling we assign all items dealing with an analytic model of the ECG curve corresponding to the single cardiac cycle as well as items dealing with an ECG curve measuring (including whole ECG, distinct segments or certain characteristics) technique, e.g., [2,3]. Analysis branch refers to the set of research items drawing an inference based on some adopted model or, as is usual in medical practice, on visual and/or manual assessment of ECG ([4,5] may serve as exemplary references). There are, of course, research items spanning both branches [6–10]. Here, as a rule, the authors introduce some novel measuring technique or analytic model, and then its capability to describe particular pathology is tested.

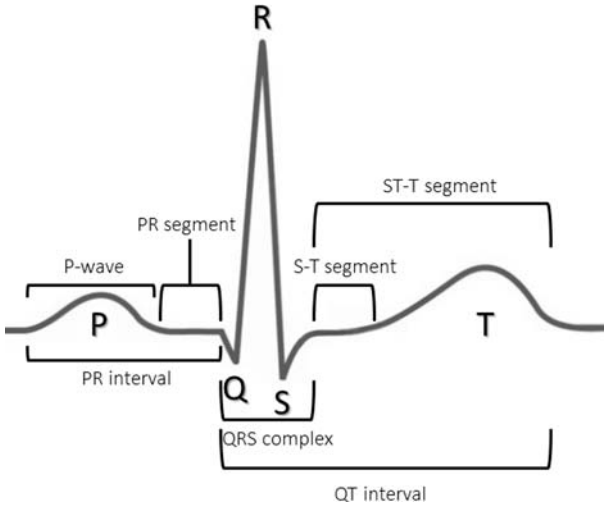

Figure 1. Single ECG wave of a heart in normal sinus rhythm.

* Corresponding address: V. Skorniakov, Faculty of Mathematics and Informatics, Vilnius University, Vilnius, Naugarduko str. 24, LT-03225, Lithuania
Phone: +370 684 88083
E-mail: viktor.skorniakov@mif.vu.lt.

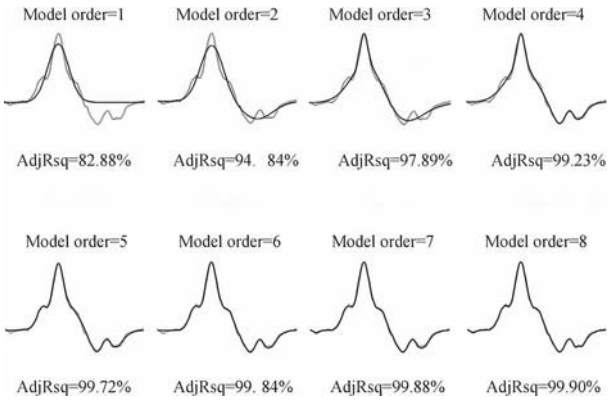

Figure 2. Single P-wave's approximation by Gaussian mixtures having different numbers of kernels [8].

The present paper is devoted to a particular analytic model of the P-wave of the ECG curve. To our best knowledge, the model was considered in [2], and later in [8]. The main idea is to represent the P-wave as a mixture of Gaussian curves perturbed by a random noise stemming from the measuring device. Figure 2, reproduced from Fig. 3 of [8], shows a single P-wave trajectory approximated by mixtures having different numbers of Gaussian curves. Fit seems quite good and thus justifies assumption regarding chosen analytic form.

Turning to biomedical literature cited previously, one finds that, in addition to this model, there are a lot of other analytic and numerical methods to characterize ECG ([11] gives a very exhaustive review of the topic as a whole, and the very recent review [12] focuses on Machine Learning models in particular). Based on these, various accurate computer-aided cardiac diagnosis (CACD) systems are developed. Taking all the said into account, the question regarding our reconsideration occurs. The reasons are as follows. First of all, we consider fully specified parametric version, resulting in a rigorous statistical model, suitable for parametric inference. [2] and [8] discussed semi-parametric versions and focused only on the estimation of model parameters without any discussion regarding further inference. Secondly, we consider a modification of versions given in [2] and [8]. Therefore, at least formally, the model may be termed as a new one. Thirdly, up to date, even most advanced automated systems designed for computerized interpretation of ECG readings still cannot replace human beings [13]. Thus, the new model in the toolbox does not seem redundant. Finally, in contrast to many previously referenced sophisticated models, the model considered exhibits the following list of features:

– simplicity;
– ability to describe dynamics of the P-wave quite accurately by a small set of scalar parameters having a clear physical interpretation;
– ability to conduct simple and rigorously grounded statistical analysis, involving virtually any interesting characteristics computed from the previously mentioned set of parameters.

In our opinion, these features are very valuable when turning to causal inference, and they are usually unavailable in case of application of "black box" methods of Machine Learning. To gain a couple of concrete examples, consider: a) exploratory analysis devoted to looking for relationships between newly emerging markers and the P-wave; b) construction of the confidence interval for a peak time of the P-wave curve in a population of healthy young individuals. In Sect. 3, we outline other possible applications and extensions.

The untouched part of the paper consists of Sect. 2 and the Appendix. In Sect. 2, we provide a specification of the model, two possible ways of parameters' estimation, and comments regarding subsequent statistical inference. Here one also finds exemplary real-data implementation. The Appendix is devoted to computational and some mathematical details.

2. Model

2.1. General setting

We assume that one records single subject's P-wave within a fixed time interval $[0, T]$. In a fully non-parametric approach, an observed trajectory is then described by a continuous time process[1] $\mathbf{X} = (X_t)$, $t \in [0, T]$, having a mean function $\mu_t = EX_t$ and a variance function $\sigma_t^2 = E(X_t - \mu_t)^2$, $t \in [0, T]$. Since we had an ability to take a series of independent measurements of the same subject, the data at hand was given by a discrete set

$$\{X_{k,t_i} : k = 1, \ldots, m; i = 0, \ldots, n\} \tag{1}$$

where:

– $\mathbf{X}_k = (X_{k,t})$ are independent identically distributed (i.i.d.) copies of $\mathbf{X} = (X_t)$;
– X_{k,t_i} denotes an observation made at time point t_i and coming from the trajectory $\mathbf{X}_k$.

Our goal was then to gain a consistent estimator of (μ_t), which could be further used for a

[1] for short, further on we omit time interval $[0, T]$; unless stated otherwise, it is assumed that all processes considered evolve within this particular interval

subsequent statistical analysis, carried out by a physician and aiming at establishing some diagnosis, comparison of controls with those being ill, etc.

Since all variables involved are bounded, by the Law of Large Numbers (LLN),

$$\hat{\mu}_t = \overline{X}_t = \frac{X_{1,t} + \cdots + X_{m,t}}{m} \xrightarrow{P} \mu_t \quad \text{and}$$
$$\hat{\sigma}_t^2 = S_t^2 = \frac{1}{m}\sum_{k=1}^{m}(X_{k,t} - \overline{X}_t)^2 \xrightarrow{P} \sigma_t^2 \tag{2}$$

giving one pair of possible estimators for μ_t, σ_t^2. However, $\hat{\mu}_t$ has a drawback of being completely non-parametric and inconvenient to work with for an ordinary practitioner. It is, therefore, desirable to provide a more handy alternative. The latter is presented in the forthcoming subsection with all notions introduced retained unaltered unless stated otherwise.

2.2. Parametric version

Our single subject's data consisted of $m = 258$ trajectories each of which was measured at equally spaced 251 time points corresponding to $n = 250$ in (1). More details regarding measurement procedure are given in the Appendix. Here, we focus on the model specification. In what follows, m, n always stand for the quantities introduced by the formula (1). Figure 3 depicts all trajectories shifted to the same origin, with a bold line showing $\hat{\mu}_t$, $t = t_0, \ldots, t_n$ given by (2). An observed pattern implies that (μ_t) could be well approximated by some mixture of Gaussian curves $t \mapsto ae^{(\frac{t-\tau}{\sigma})^2}$. By making use of least squares estimation (LSE), this was done in [2] and [8]. However, any further consideration of the distribution of obtained quantities, which is necessary for subsequent statistical inference, was provided. In our refinement, we assumed that

$$X_{k,t_i} = a_0 + \tau_0 t_i + \sum_{j=1}^{p} a_j e^{-(\frac{t_i-\tau_j}{\sigma_j})^2} + \varepsilon_{k,i}. \tag{3}$$

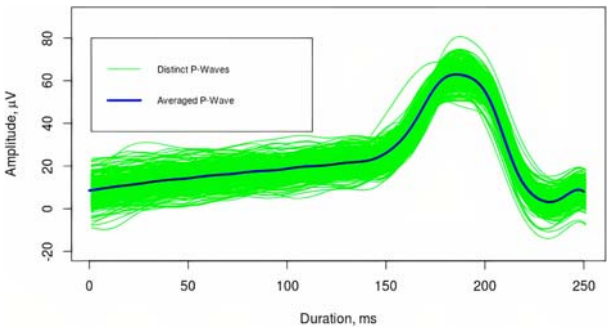

Figure 3. Single subject's readings: averaged P-wave corresponds to point averaged curve.

Where $\varepsilon_{k,i} \sim N(0; \sigma_0^2)$, $k = 1, \ldots, m$; $i = 0, \ldots, n$, are independent random variables (r. vs.) and a_j, τ_j, $\sigma_j \in \mathbb{R}$ are unknown model parameters. Hence, our model differed from counterparts of [2,8] by an addition of a linear term $a_0 + \tau_0 t_i$, which is well seen in Fig. 3, and a fully specified distribution of the remainder $\varepsilon_{k,i}$. Since (3) is a particular version of a non-linear Gaussian regression, our first choice was to estimate unknown parameters by making use of Maximum Likelihood Estimation (MLE) procedure [14], Ch. 5.5. An alternative competing method was empirical Bayes (EB) approach [15]. From Table 1 and Fig. 4 (we do not provide a graphical illustration for the MLE case since the view obtained was essentially the same), it is seen that both methods produced very similar and accurate fit. The details regarding computations are given in the Appendix. Here we restrict ourselves to comments regarding statistical inference.

These are as follows.

1. MLE offers a well-developed asymptotic theory. Since the sample size is large, by taking this approach, one ends with a set of accurate and consistent point estimates (PEs) accompanied by corresponding confidence intervals (CIs). Consequently, point hypothesis testing is readily available. Moreover, Delta method [14], Ch. 3, allows obtaining both PEs and CIs for a very broad class of functions of model parameters, whose values may be of practical interest in applications.
2. EB approach, though being more expensive computationally, in addition to PEs and CIs, offers a flexible framework for interval hypothesis testing. In our opinion, in this particular setting, empirical estimators of prior parameters are very natural. Therefore, the usual criticism of Bayes approach regarding the choice of priors' parameters does not seem to be an obstacle here. However, one can always argue regarding the choice of prior distributions. It is a place to say that we have used normal priors for all parameters. This worked well for us, yet one is free to consider alternatives. We did not settle on that since our goal was to illustrate the work of method and a good fit was the criterion for putting other alternatives aside.
3. Both methods are readily available in every widely used statistical package. Therefore, using the information provided in the Appendix, it is easy to adopt model (3).

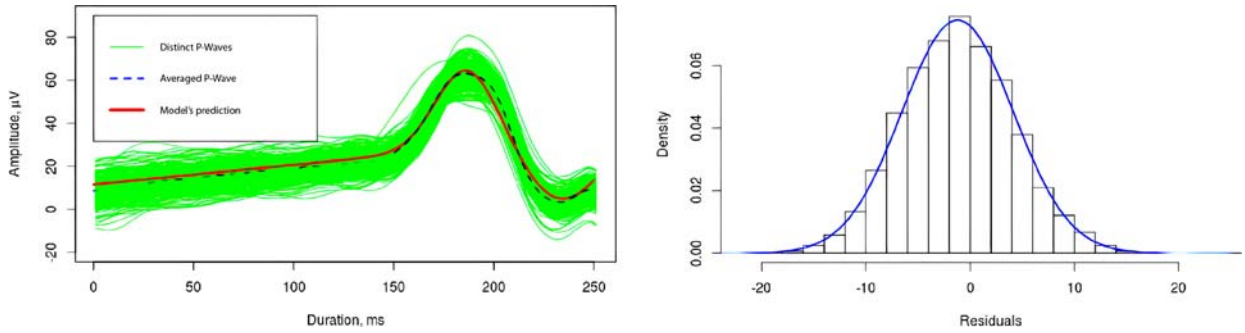

Figure 4. Model prediction and residuals: empirical Bayesian approach.

Table 1.
Summaries of models

Model[a]	Estimated mean equation	$R^2_{adj.}$ (R^2_{raw})	MSE[b]
MLE	$12.692 + 0.084t + 36.086e^{-(\frac{t-185.000}{22.721})^2} - 28.668e^{-(\frac{t-234.000}{26.505})^2}$	0.883 (0.888)	33.363
EB	$12.593 + 0.085t + 35.031e^{-(\frac{t-186.930}{22.514})^2} - 26.4748e^{-(\frac{t-235.282}{28.692})^2}$	0.883 (0.888)	33.353

[a]MLE: Maximum Likelihood Estimation based model; EB: empirical Bayesian approach based model; [b]mean squared error.

3. Possible applications and extensions

We first discuss possible model extensions, which are then also incorporated into subsection devoted to possible applications.

3.1. Extensions

Model (3) describes the pattern of a single subject's P-wave curve. To extend the model to the whole subjects' population, it is reasonable to assume that the model is subject-specific. That is, having a sample of N independent subjects, a subject having number i is described by his own set of parameters $\{(a_{ji}, \tau_{ji}, \sigma_{ji})\}_{j=0}^{p_i}$, $i = 1, \ldots, N$ (note that, for different subjects, we allow different numbers of parameters). By taking such an approach, we identify randomly drawn subject with a random draw of parameters from the corresponding parameters' space. Subject's parameters estimation methods discussed in Sect. 2.2 remain valid since the model given there may be viewed as a conditional one with respect to a fixed realization of a set of parameters. The estimated values $\{(\hat{a}_{ji}, \hat{\tau}_{ji}, \hat{\sigma}_{ji})\}_{j=0}^{p_i}$, $i = 1, \ldots, N$ then may be used for an estimation of the true distribution in the parameters' space (p_i's, though treated as random variables, are observed by a physician after gaining the subject's readings). Being aware of it, one can draw certain clinical insights. Moreover, the introduced model allows to go in an alternative way and apply non-parametric Baeysian modelling, which recently becomes more and more popular in medical applications and appears in very different modelling settings (see, e.g., [16–21]). Putting aside philosophical issues regarding relative merits of frequentist and Bayesian approach, we see that such a wide range of modelling options allows embedding the single subject's model considered by us into very different modelling contexts and obtain practically valuable insights. Several directions regarding such insights are discussed in the forthcoming subsection.

3.2. Possible applications

The importance of P-wave analysis in clinical practise is well established [22–24], and it extends beyond the analysis of cardiac arrhythmias, atrial conduction delays and can be used for prediction of clinical outcome of a wide range of cardiovascular disorders, including ischemic heart disease and congestive heart failure [25,26]. However, the research is going on, and new papers still appear. Therefore, talking about possible applications of our model, we first provide an exemplary list demonstrating different directions of investigations and comment on how our model could be used in these settings. After that, we outline several new applications one could think of after the adoption of the model introduced by us.

3.3. Adoption of the model in the contexts already explored

Table 2 provides a list of publications which focused on the exploration of relationships between P-wave based measures and various clinical conditions. The list below offers our model-based substitutes to all scalar P-wave based measures listed in Table 2.

P-wave Signal Average substitute and assessment of P-wave morphology. Since our model yields the equation of the averaged curve, the substitute here is straightforward. Also, the estimated curve can be used for an investigation of

Table 2.
P-wave related studies

PI	Conditions studied	P-wave based measures used[b]	References
AF	Post-coronary artery bypass graft surgery, Ischemia or angina, Recurrence of AF following cardioversion, After accessory pathway ablation, Following cardiac surgery	P-wave duration, P-wave dispersion, Signal Average	[27–31]
PE	Diagnosis of tuberculous constrictive pericarditis	P-wave terminal force	[32]
CAD	After percutaneous transluminal coronary angioplasty, Acute myocardial infarction, Coronary artery bypass grafting	P-wave duration, P-wave dispersion	[27,33,34]
CHD	Atrial septal defect, Atrial septal aneurysm	P-wave duration, P-wave dispersion, P-wave vector, Signal Average	[35,36]
VHD	Aortic stenosis, Mitral stenosis, Pulmonary stenosis	P-wave duration, P-wave dispersion	[37–39]
HF	Response to cardiac resynchronization therapy, Heart failure in prior myocardial infarction	P-wave morphology, P-wave terminal force	[40,41]
IS	Incidence of ischemic stroke	P-wave morphology, P-wave terminal force, P-wave duration, P-wave maximum area	[25,42]
AH	Pulmonary arterial hypertension	P-wave duration, P-wave dispersion	[43,44]
AD	Rheumatoid arthritis, Systemic lupus erythematosus, Multiple sclerosis, Ankylosing spondylitis	P-wave dispersion, P-wave axis, P-wave duration	[1,45–47]

[a]PI: Problem Investigated; AF: Atrial Fibrillation; PE: Pericarditis; CAD: Coronary Artery Disease; CHD: Congenital Heart Disease; VHD: Valvular Heart Disease; HF: Heart Failure; IS: Ischemic stroke; AH: Arterial Hypertension; AD: Autoimmune Diseases.
[b]different studies have employed at least one of the listed measures ward.

morphology associated with that particular lead for which the model was built.

P-wave duration and P-wave dispersion substitutes. We are inclined to think that linear combinations (or, more generally, functions) of $\sigma_j, j = 1, \ldots, p$, may provide information similar to that yielded by P-wave duration and P-wave dispersion.

Terminal force substitutes. Since P-wave terminal force is defined as the product of the duration and amplitude of the terminal phase of the P-wave in lead V1, one can consider products $a_j \times \sigma_j$ as possible substitutes.

P-wave maximum area. The analytic form of the P-wave equation allows applying numeric integration to compute the area under the curve associated with the particular lead for which the model was built.

3.4. Possible novel usages

Novel usages stem from the fact that one has a clearly defined *statistical* model with explicit distributional assumptions. Below, we list only a few examples of this kind. In our opinion, the provided amount is sufficient for: a) justification of the model's utility; b) generation of other possible applications of similar kind.

Building CIs and testing hypothesis for a particular subject. Both MLE and EB models allow building various CIs intervals for functions $g(\theta)$, where θ denotes the vector of model parameters. Consider, for example, construction of the confidence interval for a) the first peak time of the P-wave curve of the subject investigated, or b) the area under the P-wave. Having such intervals, one can easily test hypotheses whether these parameters fall into the acceptable region and then react accordingly.

Deriving population specific norms for preventive tasks. Recall the discussion given in Sect. 3.1. It was noted there that, applying the model to the sample of subjects drawn from some specific population, one can assess the distribution of parameters in that population and then utilize this information. Consider, for example, prevention. Say, we have constructed the confidence interval for the parameter of interest (e.g., a peak time of the P-wave curve) in a population of healthy young individuals. Having particular individual from the target population, one can then estimate his/her parameter and check whether it is in the acceptable range as suggested in the previous item "Building CIs and testing hypothesis for a particular subject".

Looking for relationships with other clinical markers. Exploration constitutes a very huge part of clinical research. Almost all studies listed in Table 2 are devoted to the tasks of this kind. After applying the model to the sample of N subjects and obtaining their estimates of parameters $\theta_1, \ldots, \theta_N$, one can utilize these values for further statistical analysis. Table 2 shows that there are a lot of directions to take. Moreover, one can apply various statistical models. To gain a few concrete examples, consider:

a) separating those with a present pathology from those with an absent pathology by making use of logistic regression, discriminant analysis, tree-based model or some other

model with $g(\theta_1), \ldots, g(\theta_N)$ acting as independents;

b) exploring relationships by means of regression model in which $g(\theta_1), \ldots, g(\theta_N)$ act as dependents and some other markers (probably appended by the set of various characteristics devoted to accounting for confounding) as independents;

c) conducting a clinical trial in which $g(\theta_1), \ldots, g(\theta_N)$ act as easy to obtain surrogate markers before the final endpoints become available.

Appendix

A.1. Data

The investigation conforms with the principles outlined in the Declaration of Helsinki and approved by the local ethics committee. To obtain the data, single subject's ECG readings having the total duration equal to 5 min and registration rate equal to 1 ms were recorded. The filtered data were used. To extract P-waves, the following algorithm was applied (see Fig. 5).

Algorithm for extraction of distinct P-waves. For each QRS complex:

- find the moment of R peak τ_R
- roll back 300 ms to moment $t_0 = \tau_R - 300$
- take readings $\{X_{t_0}, X_{t_1}, \ldots, X_{t_{250}}\}$, $t_i = t_0 + i$, $i = 1, \ldots, 250$ and treat them as a P-wave of the corresponding cardiac cycle.

A.2. Asymptotic properties of MLEs

Below, we consider asymptotic properties of MLEs of the model described in Sect. 2.2. To be more precise, the following issues are covered: a) asymptotic consistency; b) asymptotic normality; c) asymptotic inference based on a) and b). To make expressions more compact, we introduce some notions.

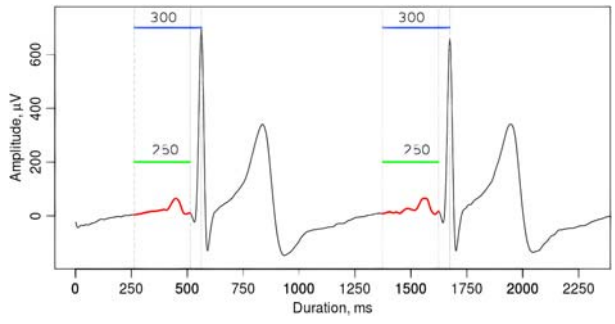

Figure 5. Data extraction visualized: the P-wave consists of the first 251 points extracted from 300 points segment measured backwards from the R peak.

- For $k = 1, \ldots, m$, $X_k = (X_{k,t_0}, \ldots, X_{k,t_n})^T$ denotes an observed vector corresponding to the k-th trajectory $\mathbf{X}_k$.
- For $i = 0, \ldots, n$, $\overline{X}_{\cdot,i} = \frac{1}{m}\sum_{k=1}^{m} X_{k,t_i}$ denotes an averaged trajectory at t_i whereas $\overline{X}_{\cdot} = (\overline{X}_{\cdot,1}, \ldots, \overline{X}_{\cdot,n})^T$.
- For $i = 0, \ldots, 1$ and $j = 1, \ldots, p$, $\kappa_{i,j} = e^{-(\frac{t_i - \tau_j}{\sigma_j})^2}$ and $\kappa_{\cdot,j} = (\kappa_{0,j}, \ldots, \kappa_{n,j})^T$.
- $\theta = (\sigma_0, \sigma_1, \ldots, \sigma_p, a_0, \ldots, a_p, \tau_0, \ldots, \tau_p)^T$ denotes the vector of the model parameters.
- For a random vector (r.v.) $Y \in \mathbb{R}^q$, $f_Y(y)$ stands for its density at $y \in \mathbb{R}^q$.
- For $x = (x_1, \ldots, x_q)^{\mathrm{T}}$, $y = (y_1, \ldots, y_q)^{\mathrm{T}} \in \mathbb{R}^{\mathrm{q}}$, $\|y\| = \sqrt{y_1^2 + \cdots + y_q^2}$, $x \circ y = (x_1 y_1, \ldots, x_q y_q)^T$ and $x \cdot y = \sum_{j=1}^{q} x_1 y_1 + \cdots + x_q y_q$ denote the standard Euclidian norm, Hadamard product and inner product respectively; $x^{\circ d} = x \circ x \circ \cdots \circ x$; finally, $\mathbf{1}_q = (1, 1, \ldots, 1)^T \in \mathbb{R}^q$ (when the dimension is clear, subscript is omitted and $\mathbf{1}$ is written instead of $\mathbf{1}_q$).

From (3) (and Section A.1), it follows that $X_k \sim N_{n+1}(\mu(\theta, t); \sigma_0^2 I_{n+1})$, where I_{n+1} is an identity matrix of order $n+1$, $t = (t_0, \ldots, t_n)^T$ is the time vector, and $\mu(\theta, t) = \mu = (\mu_0, \ldots, \mu_n)^T$ with $\mu_i = a_0 + \tau_0 t_i + \sum_{j=1}^{p} a_j e^{-(\frac{t_i - \tau_j}{\sigma_j})^2}$, $i = 0, \ldots, n$. Taking into account independence of X_k, $k = 1, \ldots, m$, log-likelihood is then equal to

$$\begin{aligned} \ell_m(\theta) &= \frac{1}{m}\sum_{k=1}^{m} \ln f_{X_k}(X_k) \\ &= -(n+1)\ln(\sigma_0\sqrt{2\pi}) \\ &\quad - \frac{1}{2m\sigma_0^2}\sum_{k=1}^{m}\sum_{i=0}^{n}(X_{k,t_i} - \mu_i)^2 \\ &= -(n+1)\ln(\sigma_0\sqrt{2\pi}) - \frac{1}{2m\sigma_0^2}\sum_{k=1}^{m}\|X_k - \mu\|. \end{aligned} \tag{4}$$

Because of the nature of the problem considered, it is reasonable to assume that the parameters' space Θ is a compact interval $[\tilde{\sigma}_L, \tilde{\sigma}_U]^{(p+1)} \times [\tilde{a}_L, \tilde{a}_U]^{(p+1)} \times [\tilde{\tau}_L, \tilde{\tau}_U]^{(p+1)}$ of $(0, \infty)^{(p+1)} \times \mathbb{R}^{(p+1)} \times (0, \infty)^{(p+1)}$ and that the true value θ_{TRUE} lies in the interior of this compact interval. Even without detailed computations, it is then quite evident that, assuming such Θ, ℓ_m is three times continuously differentiable on Θ_{int} with partial derivatives dominated by some measurable integrable function, and one can exchange the order of differentiation and integration provided there is a need. Moreover, computations below show that the Fisher information matrix $\mathcal{J}(\theta_{\mathrm{TRUE}}) = E_{\theta_{\mathrm{TRUE}}}(\frac{\partial}{\partial\theta}\ln f_{X_1}(X_1))(\frac{\partial}{\partial\theta}\ln f_{X_1}(X_1))^T$ is non-singular provided we drop artificial and practically irrelevant cases (see explanation below). That is, all standard assumptions ensuring existence, consis-

tency and asymptotic normality of MLE estimator (see, e.g. [14], Theorems 5.41–5.42) $\hat{\theta}_m$ hold and yield relationships $\hat{\theta}_m \xrightarrow{P} \theta_{\text{TRUE}}$, $\sqrt{m\mathcal{J}(\theta_{\text{TRUE}})}(\hat{\theta}_m - \theta_{\text{TRUE}}) \xrightarrow{d} N(0; I_{n+1})$. In order to employ these formulas practically, one has to obtain an expression for $\mathcal{J}(\theta)$ and replace $\mathcal{J}(\theta_{\text{TRUE}})$ with an estimator $\mathcal{J}(\hat{\theta}_{\text{TRUE}})$. For the sake of the readers' convenience, we provide computation of $\mathcal{J}(\theta)$ and estimating equations for solving $\hat{\theta}_m$. The latter are given only for completeness and may be omitted since, in our opinion, it is more convenient to make use of the computational scheme provided in Section A.3. First, note that $\ln f_{X_1}(X_1) = -(n+1)\ln(\sigma_0\sqrt{2\pi}) - \frac{1}{2\sigma_0^2}\|X_1-\mu\|^2$. Hence,

$$\begin{aligned}\frac{\partial}{\partial\theta_0}\ln f_{X_1}(X_1) &= \frac{\partial}{\partial\sigma_0}\ln f_{X_1}(X_1)\\ &= -\frac{n+1}{\sigma_0} + \frac{1}{\sigma_0^3}\|X_1-\mu\|^2\\ &= \frac{1}{\sigma_0}\big(\xi_{n+1} - (n+1)\big)\\ &= \frac{\xi_{n+1} - E\xi_{n+1}}{\sigma_0}\end{aligned} \tag{5}$$

$$\begin{aligned}&\frac{\partial}{\partial\theta_s}\ln f_{X_1}(X_1)\frac{1}{\sigma_0^2}\sum_{i=0}^{n}(X_{1,i}-\mu_i)\frac{\partial}{\partial\theta_s}\mu_i\\ &= \frac{(X_1-\mu)\cdot\mu'_{\theta_s}}{\sigma_0^2} \quad \text{for } s \geq 1.\end{aligned} \tag{6}$$

Where $\xi_{n+1} = \frac{\|X_1-\mu\|^2}{\sigma_0^2} \sim \chi^2_{n+1}$ and $\mu'_{\theta_s} = (\frac{\partial}{\partial\theta_s}\mu_0, \ldots, \frac{\partial}{\partial\theta_s}\mu_n)^T$. Therefore, $\mathcal{J}(\theta)$ equals to

$$\begin{aligned}&\frac{1}{\sigma_0^4}\begin{pmatrix}\sigma_0^2\operatorname{Var}(\xi_{n+1}) & 0 & \cdots & 0\\ 0 & (\mu'_{\theta_1})^T\operatorname{cov}(X_1)\mu'_{\theta_1} & \cdots & (\mu'_{\theta_1})^T\operatorname{cov}(X_1)\mu'_{\theta_{3p+2}}\\ & \vdots & \ddots & \vdots\\ 0 & (\mu'_{\theta_{3p+2}})^T\operatorname{cov}(X_1)\mu'_{\theta_1} & \cdots & (\mu'_{\theta_{3p+2}})^T\operatorname{cov}(X_1)\mu'_{\theta_{3p+2}}\end{pmatrix}\\ &= \frac{1}{\sigma_0^2}\begin{pmatrix}2(n+1) & 0 & \cdots & 0\\ 0 & \mu'_{\theta_1}\mu'_{\theta_1} & \cdots & \mu'_{\theta_1}\mu'_{\theta_{3p+2}}\\ & \vdots & \ddots & \vdots\\ 0 & \mu'_{\theta_{3p+2}}\mu'_{\theta_1} & \cdots & \mu'_{\theta_{3p+2}}\mu'_{\theta_{3p+2}}\end{pmatrix}\end{aligned} \tag{7}$$

And $\det\mathcal{J}(\theta) = 2(n+1)\sigma_0^{-6(p+1)}\det(M)$, where $M = (\mu'_{\theta_r}\cdot\mu'_{\theta_s})$, $r,s = 1,\ldots,3p+2$, denotes the second block diagonal matrix in (7). Note that $\det(M)$ is the Gram determinant corresponding to the system of vectors $\{\mu'_{\theta_s}\}$, $s = 1,\ldots,3p+2$. Thus, it is nonzero if and only if that system is composed of independent vectors. The dependence is the previously announced artificial condition (consider, for example, the case of $(a_j,\tau_j,\sigma_j) = (a_i,\tau_i,\sigma_i)$ for some $i \neq j$, which should very rarely, if ever, occur in practise). Having estimate $\hat{\theta}_m$, one can always check the condition empirically by computing $\det M(\hat{\theta}_m)$. If it happens so that $\det M(\hat{\theta}_m) \approx 0$, it is advisory to reconsider the model and try to minimize the number of parameters. Otherwise, one can apply the asymptotic theory given above. For this, it remains to note that, for any $j = 1,\ldots,p$,

$$\mu'_{\theta_s} = \begin{cases}\mathbf{1}_{n+1} = \mathbf{1} \quad \text{for } \theta_s = a_0,\\ (t_0,\ldots,t_n)^T = t \quad \text{for } \theta_s = \tau_0,\\ (\kappa_{0,j},\ldots,\kappa_{n,j})^T = \kappa_{\cdot,j} \quad \text{for } \theta_s = a_j,\\ \frac{2a_j}{\sigma_j^2}((t_0-\tau_j)\kappa_{0,j},\ldots,(t_n-\tau_j)\kappa_{n,j})^T\\ \quad = \frac{2a_j}{\sigma_j^2}(t-\tau_j\mathbf{1})\circ\kappa_{\cdot,j} \quad \text{for } \theta_s = \tau_j,\\ \frac{2a_j}{\sigma_j^3}\big((t_0-\tau_j)^2\kappa_{0,j},\ldots,(t_n-\tau_j)^2\kappa_{n,j}\big)^T\\ \quad = \frac{2a_j}{\sigma_j^3}(t-\tau_j\mathbf{1})^{\circ 2}\circ\kappa_{\cdot,j} \quad \text{for } \theta_s = \sigma_j.\end{cases} \tag{8}$$

Combining these expressions with (4)–(6), we see that estimating equations $\frac{\partial}{\partial\theta}\ell(\theta) = 0$ for solving $\hat{\theta}_m$ can be written as follows:

$$\sigma_0 = \sqrt{\frac{1}{(n+1)m}\sum_{k=1}^{m}\|X_k-\mu\|^2} \tag{9}$$

$$\begin{aligned}&(\overline{X_\cdot}-\mu)\cdot\mathbf{1} = 0,\\ &(\overline{X_\cdot}-\mu)\cdot t = 0,\\ &(\overline{X_\cdot}-\mu)\cdot\kappa_{\cdot,j} = 0,\end{aligned} \tag{10}$$

$$\begin{aligned}&(\overline{X_\cdot}-\mu)\cdot\big((t-\tau_j\mathbf{1})\circ\kappa_{\cdot,j}\big) = 0,\\ &(\overline{X_\cdot}-\mu)\cdot\big((t-\tau_j\mathbf{1})^{\circ 2}\circ\kappa_{\cdot,j}\big) = 0\end{aligned} \tag{11}$$

for $j = 1,\ldots,p$. Since μ does not depend on σ_0, from (9), it follows that an estimator of σ_0 is completely determined by values of the rest estimators. Next, for any $A \subset \{1,2,\ldots,3p+2\}$, let θ_A denote a sub vector of θ which is composed from coordinates θ_s, $s \in A$, and let $A_1, A_2, A_3 \subset \{1,2,\ldots,3p+2\}$ be the subsets for which it holds $\theta_{A_1} = (\tau_0, a_0, a_1, \ldots, a_p)^T$, $\theta_{A_2} = (\tau_1,\ldots,\tau_p)^T$, $\theta_{A_3} = (\sigma_1, \ldots, \sigma_p)^T$. Noting that $\mu = K\theta_{A_1}$ with $K = (t\mathbf{1}_{n+1}\kappa_{\cdot,1}, \ldots, \kappa_{\cdot,p})$ and rewriting equations (10) in an equivalent form

$$\overline{X_\cdot}\cdot\mathbf{1} = \mathbf{1}^TK\theta_{A_1}, \qquad \overline{X_\cdot}\cdot t = t^TK\theta_{A_1},$$

$$\overline{X_\cdot}\cdot\kappa_{\cdot,j} \overset{j\geq 1}{=} \kappa_{\cdot,j}^TK\theta_{A_1},$$

one arrives to conclusion that

$$\theta_{A_1} = \big(K^TK\big)^{-1}K^T\overline{X_\cdot} \tag{12}$$

Since K depends only on θ_{A_2}, θ_{A_3}, one needs to solve these parameters in order to obtain solutions for σ_0 and θ_{A_1}. One can rearrange the first set of the equations (11) into the fixed point equation for θ_{A_2}, then plug into this equation the value of θ_{A_1} given by (12), and solve θ_{A_2} iteratively for any fixed θ_{A_3}. Equations for θ_{A_3}, however, remain quite complicated and do not seem to admit substantial simplification. Nonetheless, one can make use of the derivations given above to find initial solutions for σ_0, θ_{A_1} and θ_{A_2} provided the initial solution for θ_{A_3} is already at hand. The latter solution then may be supplied to any standard software package offering MLE procedure.

However, it is more natural to take into account the nature of the problem considered and to obtain the initial solution as explained in the forthcoming Section A.3. Therefore, we do not settle down on further development of the above route.

There remains to say a word regarding statistical inference. The latter is quite straightforward: having any well-behaved function $g : \Theta \to \mathbb{R}^q$, one should employ the Delta method (see [14], Ch. 3) which allows both asymptotic testing and building of CIs.

Finishing this subsection, we remind that, in our particular case, $p = 2$, $m = 258$, and $n = 250$. However, to keep a more general framework, we will remain with unplugged values without emphasizing this in the sequel.

A.3. Computations

A.3.1. Preliminaries

From the results of Section A.2, it follows that analytic solution of parameters was unavailable in case of MLE framework; the same applied to EB framework as well. Therefore, parameters were estimated numerically. Computations were performed by making use of R statistical computing environment [48]. In addition to base R, the following packages were employed: pracma [49] for finding peaks of ECG curves; mcmc [50] for Bayesian analysis. MLE estimation was based on quasi-Newton method. For Bayesian analysis, Metropolis–Hastings algorithm was applied. For each model, initial solutions had to be supplied. The latter were obtained by making use of the following algorithm.

Algorithm for generation of initial values.

1. Fix the number of Gaussian curves p and compute averaged trajectory $\{\hat{\mu}(t_i)|i = 0, \dots, n\}$, where $\hat{\mu}(t_i)$ is given by (2).
2. Set $t_{\text{reg}} = 140$, $\sigma_0 = \sqrt{\frac{1}{n}\sum_{i=0}^{n} s_i^2}$, $s_i^2 = \frac{1}{m-1} \times \sum_{k=1}^{m}(X_{k,t_i} - \bar{X}_{\cdot,i})^2$, and obtain initial estimates $\hat{a}_0$, $\hat{\tau}_0$ by fitting simple linear regression model on the subset $\{\hat{\mu}(t_i)|i = 0, \dots, t_{\text{reg}}\}$
3. Compute $res_0 = \{\hat{\mu}(t_i) - \hat{a}_0 - \hat{\tau}_0 t_i = res_0(t_i)|i = 0, \dots, n\}$
4. For $j = 1, \dots, p$:
   (a) $\hat{a}_j$ = amplitude of the peak of the first wave of res_{j-1};
   (b) $\hat{\tau}_j$ = time corresponding to the peak of the first wave of res_{j-1};
   (c) $\hat{\sigma}_j = \sqrt{\frac{-(t^{(0)}-\hat{\tau}_1)^2}{\log(\frac{res_{j-1}(t^{(0)})}{\hat{a}_j})}}$, where
   $$t^{(0)} = \begin{cases} \max\{t \in \{0, \dots, \hat{\tau}_j - 1\} \big| |res_{j-1}(t)| \\ \quad \geq 0.8|\hat{a}_j|\}, \quad \text{for } \hat{\tau}_j \geq \frac{n}{2}; \\ \min\{t \in \{\hat{\tau}_j + 1, \dots, n\} \big| |res_{j-1}(t)| \\ \quad \geq 0.8|\hat{a}_j|\}, \quad \text{for } \hat{\tau}_j < \frac{n}{2}; \end{cases}$$
   (d) $res_j = \{res_{j-1}(t_i) - \hat{a}_j e^{(\frac{t_i - \hat{\tau}_j}{\hat{\sigma}_j})^2} | i - 0, \dots, n\}$.
5. Take $\{\hat{\sigma}_j, \hat{a}_j, \hat{\tau}_j\}_{j=0}^{p}$ as an initial solution.

Remark 1. To justify computation of $\hat{\sigma}_j$ in 4.c, note that $t \mapsto ae^{(\frac{t-\tau}{\sigma})^2}$ has two inverse branches on $(0, a) : t_{\pm}(x) = \tau \pm \sigma\sqrt{-\log(\frac{x}{a})}$. The choice of threshold value 0.8 is arbitrary. However, in [2] it is noted that varying threshold values do not affect results significantly and the model is robust with respect to this quantity.

In Sect. 2.2, we have already mentioned that all priors for parameters were normal. The hyper-parameters were obtained as follows. First, for each distinct trajectory, we have obtained estimates $\{(\hat{a}_{kj}, \hat{\tau}_{kj}, \hat{\sigma}_{kj})\}_{j=0}^{p}$, $k = 1, \dots, m$, in the same way as explained above in the "Algorithm for generation of initial values". That is, the only difference was that we have used distinct trajectories X_k, $k = 1, \dots, m$ instead of an averaged one. After that, we have assumed that $\sigma_j \sim N(\overline{\sigma}_j, s^2_{\overline{\sigma}_j})$, $a_j \sim N(\overline{a}_j, s^2_{\overline{a}_j})$, $\tau_j \sim N(\overline{\tau}_j, s^2_{\overline{\tau}_j})$ were independent with

$$\overline{\sigma}_j = \frac{1}{m}\sum_{k=1}^{m}\hat{\sigma}_{kj}, \qquad s^2_{\overline{\sigma}_j} = \frac{1}{m-1}\sum_{k=1}^{m}(\hat{\sigma}_{kj} - \overline{\sigma}_j)^2,$$

$$\overline{a}_j = \frac{1}{m}\sum_{k=1}^{m}\hat{a}_{kj}, \qquad s^2_{\overline{a}_j} = \frac{1}{m-1}\sum_{k=1}^{m}(\hat{a}_{kj} - \overline{a}_j)^2,$$

$$\overline{\tau}_j = \frac{1}{m}\sum_{k=1}^{m}\hat{\tau}_{kj}, \qquad s^2_{\overline{\tau}_j} = \frac{1}{m-1}\sum_{k=1}^{m}(\hat{\tau}_{kj} - \overline{\tau}_j)^2.$$

Consequently, the posterior used to perform computations was proportional to

$$e^{\ell(\theta)} = \prod_{j=0}^{p}\varphi(\sigma_j|\overline{\sigma}_j, s_{\overline{\sigma}_j})\varphi(a_j|\overline{a}_j, s_{\overline{a}_j})\varphi(\tau_j|\overline{\tau}_j, s_{\overline{\tau}_j}),$$

where $\ell(\theta)$ is given by (4) and $\varphi(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \times e^{-(\frac{x-\mu}{\sigma})^2}$ denotes the density of the r. v. having normal distribution $N(\mu; \sigma^2)$. The values of the hyper-parameters are given in Table 3.

Note that, for $j \geq 1$, σ_j entered the posterior squared meaning that there was no positivity constraint on these parameters, and only σ_0 had to be positive. However, from Table 3 it is seen that the coefficient of variation $C_{\sigma_0} = \frac{s_{\overline{\sigma}_0}}{\overline{\sigma}_0} \approx 0.099$ was small. Moreover, the obtained initial value

Table 3.
Hyper-parameters of the priors[a]

	σ_0	σ_1	σ_2	a_0	a_1	a_2	τ_0	τ_1	τ_2
Mean	16.811	24.602	25.246	9.355	18.561	–11.979	0.095	199.124	219.733
St. dev.	1.658	5.639	6.875	5.979	34.288	26.360	0.042	21.972	20.778

[a] each parameter was assumed to follow normal distribution with the mean and standard deviation given in the corresponding column.

and MLE estimate were equal to 16.811 and 16.172 correspondingly. Hence justification of the choice of the normal prior used by us.

A.3.2. Factual implementation with R

```
# packages
library(pracma);library(mcmc)
# data
path <- "/media/visk/Local Disk/Science/P wave ECG/data.csv";
df<- read.csv(file = path)

# I) initial assignments
m <- 258 # number of trajectories
n <- 250 # number of points in a single trajectory
p <- 2 # number of Gaussian curves
t <- seq(from=0, to=n, by=1) # time vector
r <- 140 # number of points for estimation of a0, tau0
meanX <- rowMeans(df, na.rm = FALSE, dims = 1) # averaged X

# F) auxiliary functions
# F.1) function for computation of estimates from a given curve
getEstimatesFromCurve <- function(curve){
    # vectors for parameters storage
    a <- matrix(data = 0, nrow=p, ncol=1);tau <- matrix(data = 0,
    nrow=p, ncol=1);sigma <- matrix(data = 0, nrow=p, ncol=1);
    # computation of a0, tau0
    regLine <- lm(curve[1:r]~t[1:r]);a0 <- as.numeric(regLine$
    coefficients[1]);tau0 <- as.numeric(regLine$coefficients[2]);
    res <- as.vector(x = curve - a0-tau0*t) # subtract linear part
    #computation of the rest parameters
    for (j in (1:p)){
      peakPos <- findpeaks(res, minpeakdistance=1, npeaks=1,
      sortstr=TRUE)
      peakNeg <- findpeaks(-res, minpeakdistance=1, npeaks=1,
      sortstr=TRUE)
      a[j] <- (peakPos[1]>peakNeg[1])*peakPos[1] - (peakPos[1]<=
      peakNeg[1])*peakNeg[1]
      tau[j] <- (peakPos[1]>peakNeg[1])*peakPos[2] + (peakPos[1]<=
      peakNeg[1])*peakNeg[2]
      # time, step and threshold for computation of sigma[j]
      tsigma = tau[j]; s = ifelse(tsigma>n/2,1,-1);threshold =
      abs(0.8*a[j])
      while (abs(res[tsigma])>threshold) {tsigma = tsigma-s}
      sigma[j]=sqrt(-(tsigma-tau[j])^2/log(abs(res[tsigma]/a[j])))
      res <- res - a[j]*exp(-((t-tau[j])/sigma[j])^2)
      }
    return(c(sigma, a0, a, tau0, tau))
}

# F.2) function for computation of log-likelihood
getLL <- function(posterior = FALSE, hyperPar = NULL){
    ll<-function(par){
      # par[1] ~ sigma0, par[2:(p+1)] ~ sigma[1:p]; par[p+2] ~ a0,
      par[(p+3):2*(p+1)] ~ a[1:p]; par[2*p+3] ~ tau0, par[(2*p+4):
      3*(p+1)] ~ tau[1:p];
      f1=0
      f1=sapply(1:p,function(k){f1=f1+(par[p+2+k]*exp(-((t-
      par[2*p+3+k])/par[1+k])^2))}); f1=par[p+2]+t*par[2*p+3]
      f=n*m*log(par[1])+(1/(2*par[1]^2))*sum((df-f1)^2)
      if (posterior){f = -f+sum(sapply(1:c(3*(p+1)),
      FUN = function(i){return(dnorm(par[i],hyperPar[i,1],
      hyperPar[i,2]))}))}
      return(f)
    }
    return(ll)
}

# M) MLE block (run I and F blocks first)# estimation
sigma0 <- mean(apply(df,2,sd))
initialVals <- c(sigma0,getEstimatesFromCurve(curve = meanX))
mleEst <- optim(par=initialVals, fn=getLL(),
    lower=c(rep(1.0e-5,p+1),rep(-Inf, c(2*(p+1)))),
    method="L-BFGS-B", control=c(maxit=100000))
parVec <- mleEst$par # for parameters order, see comments in getLL

# B) Bayesian block (run I and F blocks first)

# computation of hyperparameters:
# parameters' order corresponds to that used previously, i.e.
# sigma0, sigma, a0, a, tau0, tau
hyperPar <- matrix(data = 0,nrow = 3*(p+1), ncol = 2)
hyperPar[1,1] <- mean(apply(df,2,sd)) # sigma0mean
hyperPar[1,2] <- sd(apply(df,2,sd)) # sigma0sd
# all the rest:
hyperM <- matrix(data = 0,nrow = 3*(p+1), ncol = 2)
hyperM <- sapply(X = 1:m, FUN =function(i){getEstimatesFromCurve
(curve = df[,i])})
hyperPar[2:c(3*(p+1)),] <- t(apply(X = hyperM,MARGIN = 1,FUN =
function(prow){return(c(mean(prow),sd(prow)))}))
# estimation
sigma0 <- mean(apply(df,2,sd))
initialVals <- c(sigma0,getEstimatesFromCurve(curve = meanX))
set.seed(2018)
batchSize <- 100000
mc <- metrop(getLL(posterior=TRUE, hyperPar=hyperPar),initial=
initialVals, nbatch = batchSize, scale = 0.5)
parVec <- apply(mc$batch[1:batchSize,],2,mean)

# S) summaries
y_est=parVec[p+2]+t*parVec[2*p+3]
for (j in 1:p){y_est=y_est+(parVec[p+2+j]*exp(-((t-parVec[2*p+3+j])/
parVec[1+j])^2))}; plot(df[,1],type="l",col='green', ylim=c(-20,90))
for (i in 2:n) {lines(df[,i], type="l", col='green')}
lines(x=t,y=y_est, col='red', type='l', lwd=3); lines(x=t,y=meanX,
col='blue', type='l', lwd=2, lty=2)
err = (df-y_est); mse=sum(err^2)/(n*m); sRES=sum((df-y_est)^2); sTOT<-
sum((df-mean(as.matrix(df)))^2)
R2=1-sRES/sTOT; R2_adj=1-(sRES/(n-2-(p+1)*3-1))/(sTOT/(n-1));
hist(as.numeric(matrix(data = unlist(err),ncol = 1)))
```

Declarations of interest

none.

Author contributions

Viktor Skorniakov: Conceptualization, Methodology, Software, Formal analysis, Writing – Original Draft, Visualization, Project administration, **Antanas Mainelis**: Methodology, Software, Formal analysis, Writing – Original Draft, Visualization, **Petras Navickas**: Investigation, Resources, Writing – Original Draft, Visualization, **Germanas Marinskis**: Methodology, Writing – Review & Editing, **Albinas Stankus**: Conceptual-

ization, Methodology, Resources, Writing – Original Draft.

Funding

this research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Acar RD, Bulut M, Acar S, Izci S, Fidan S, Yesin M, et al. Evaluation of the P wave axis in patients with systemic lupus erythematosus. *J Cardiovasc Thorac Res* 2015;7(4):154–7.

[2] Suppappola S, Sun Y, Chiaramida SA. Gaussian pulse decomposition: an intuitive model of electrocardiogram waveforms. *Ann Biomed Eng* 1997;25(2):252–60.

[3] Herreros A, Baeyens E, Peran JR, Johansson R, Carlson J, Olsson B, editors. An algorithm for phase-space detection of the P characteristic points. 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2007.

[4] Herreros A, Baeyens E, Johansson R, Carlson J, Perán JR, Olsson B. Analysis of changes in the beat-to-beat P-wave morphology using clustering techniques. *Biomed Signal Process Control* 2009;4(4):309–16.

[5] Furniss GO, Panagopoulos D, Kanoun S, Davies EJ, Tomlinson DR, Haywood GA. The effect of atrial fibrillation ablation techniques on P wave duration and P wave dispersion. *Heart Lung Circ* 2019;28(3):389–96.

[6] Filos D, Chouvarda I, Tachmatzidis D, Vassilikos V, Maglaveras N. Beat-to-beat P-wave morphology as a predictor of paroxysmal atrial fibrillation. *Comput Methods Programs Biomed* 2017;151:111–21.

[7] Martis RJ, Chakraborty C, Ray AK. A two-stage mechanism for registration and classification of ECG using Gaussian mixture model. *Pattern Recognit* 2009;42(11):2979–88.

[8] Censi F, Calcagnini G, Ricci C, Ricci RP, Santini M, Grammatico A, et al. P-wave morphology assessment by a Gaussian functions-based model in atrial fibrillation patients. *IEEE Trans Biomed Eng* 2007;54(4):663–72.

[9] Srinivasan N, Wong M, Krishnan S, editors. A new phase space analysis algorithm for cardiac arrhythmia detection. Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat No 03CH37439). IEEE; 2003.

[10] Diery A, Rowlands D, Cutmore TR, James D. Automated ECG diagnostic P-wave analysis using wavelets. *Comput Methods Programs Biomed* 2011;101(1):33–43.

[11] Martis RJ, Acharya UR, Adeli H. Current methods in electrocardiogram characterization. *Comput Biol Med* 2014;48:133–49.

[12] Lyon A, Mincholé A, Martínez JP, Laguna P, Rodriguez B. Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances. *J R Soc Interface* 2018;15(138):20170821.

[13] Schläpfer J, Wellens HJ. Computer-interpreted electrocardiograms: benefits and limitations. *J Am Coll Cardiol* 2017;70(9):1183–92.

[14] Van der Vaart AW. Asymptotic statistics. Cambridge University Press; 2000.

[15] Carlin BP, Louis TA. Bayes and empirical Bayes methods for data analysis. Chapman and Hall/CRC; 2010.

[16] Ohlssen DI, Sharples LD, Spiegelhalter DJ. Flexible random-effects models using Bayesian semi-parametric models: applications to institutional comparisons. *Stat Med* 2007;26(9):2088–112.

[17] Dunson DB, Herring AH, Engel SM. Bayesian selection and clustering of polymorphisms in functionally related genes. *J Am Stat Assoc* 2008;103(482):534–46.

[18] Scarpa B, Dunson DB. Bayesian hierarchical functional data analysis via contaminated informative priors. *Biometrics* 2009;65(3):772–80.

[19] Barcella W, Iorio MD, Baio G, Malone-Lee J. Variable selection in covariate dependent random partition models: an application to urinary tract infection. *Stat Med* 2016;35(8):1373–89.

[20] Canale A, Lijoi A, Nipoti B, Prünster I. On the Pitman–Yor process with spike and slab base measure. *Biometrika* 2017;104(3):681–97.

[21] Paulon G, De Iorio M, Guglielmi A, Ieva F. Joint modeling of recurrent events and survival: a Bayesian non-parametric approach. *Biostatistics* (Oxford, England). 2018.

[22] Magnani JW, Williamson MA, Ellinor PT, Monahan KM, Benjamin EJ. P wave indices: current status and future directions in epidemiology, clinical, and research applications. *Circ Arrhythm Electrophysiol* 2009;2(1):72–9.

[23] Laucevicius A. Non-invasive diagnosis of cardiac arrhythmias. In: Lip GY, Godtfredsen J. Cardiac arrhythmias: a clinical approach. Mosby; 2003. p. 189–214.

[24] Laucevičius A, Aidietis A, Černiauskas R, Marinskis G, Aganauskienė J, Jokšas V. The signal averaged P wave electrocardiogram – methodological problems and perspective. *Kardiologijos seminarai* 1998;4(2):114–9.

[25] He J, Tse G, Korantzopoulos P, Letsas KP, Ali-Hasan-Al-Saegh S, Kamel H, et al. P-wave indices and risk of ischemic stroke: a systematic review and meta-analysis. *Stroke* 2017;48(8):2066–72.

[26] Platonov PG. P-wave morphology: underlying mechanisms and clinical implications. *Ann Noninvasive Electrocardiol* 2012;17(3):161–9.

[27] Chandy J, Nakai T, Lee RJ, Bellows WH, Dzankic S, Leung JM. Increases in P-wave dispersion predict postoperative atrial fibrillation after coronary artery bypass graft surgery. *Anesth Analg* 2004;98(2):303–10.

[28] Dilaveris PE, Andrikopoulos GK, Metaxas G, Richter DJ, Avgeropoulou CK, Androulakis AM, et al. Effects of ischemia on P wave dispersion and maximum P wave duration during spontaneous anginal episodes. *Pacing Clin Electrophysiol* 1999;22(11):1640–7.

[29] Budeus M, Hennersdorf M, Perings C, Wieneke H, Erbel R, Sack S. Prediction of the recurrence of atrial fibrillation after successful cardioversion with P wave signal-averaged ECG. *Ann Noninvasive Electrocardiol* 2005;10(4):414–9.

[30] Aytemir K, Amasyali B, Kose S, Kilic A, Abali G, Oto A, et al. Maximum P-wave duration and P-wave dispersion predict recurrence of paroxysmal atrial fibrillation in patients with Wolff–Parkinson–White syndrome after successful radiofrequency catheter ablation. *J Interv Card Electrophysiol* 2004;11(1):21–7.

[31] Lazzeroni D, Parati G, Bini M, Piazza P, Ugolotti PT, Camaiora U, et al. P-wave dispersion predicts atrial fibrillation following cardiac surgery. *Int J Cardiol* 2016;203:131–3.

[32] Ren Y, Qiu J, Li Z, Li C. P-wave terminal force in lead V1 is a predictive indicator for the diagnosis of tuberculous constrictive pericarditis. *Heart Lung* 2019;48(2):155–8.

[33] Akdemir R, Ozhan H, Gunduz H, Tamer A, Yazici M, Erbilen E, et al. Effect of reperfusion on P-wave duration and P-wave dispersion in acute myocardial infarction: primary angioplasty versus thrombolytic therapy. *Ann Noninvasive Electrocardiol* 2005;10(1):35–40.

[34] Baykan M, Çelik Ş, Erdöl C, Durmuş İ, Örem C, Küçükosmanoğlu M, et al. Effects of P-wave dispersion on atrial fibrillation in patients with acute anterior wall myocardial infarction. *Ann Noninvasive Electrocardiol* 2003;8(2):101–6.
[35] Ho TF, Chia E, Yip W, Chan K. Analysis of P wave and P dispersion in children with secundum atrial septal defect. *Ann Noninvasive Electrocardiol* 2001;6(4):305–9.
[36] Deveci OS, Aytemir K, Okutucu S, Tulumen E, Aksoy H, Kaya EB, et al. Evaluation of the relationship between atrial septal aneurysm and cardiac arrhythmias via P-wave dispersion and signal-averaged P-wave duration. *Ann Noninvasive Electrocardiol* 2010;15(2):157–64.
[37] Turhan H, Yetkin E, Atak R, Altinok T, Senen K, Ileri M, et al. Increased P-wave duration and P-wave dispersion in patients with aortic stenosis. *Ann Noninvasive Electrocardiol* 2003;8(1):18–21.
[38] Ozmen N, Cebeci BS, Kardesoglu E, Celik T, Dincturk M, Demiralp E. P wave dispersion is increased in pulmonary stenosis. *Indian Pacing Electrophysiol J* 2006;6(1):25.
[39] Guntekin U, Gunes Y, Tuncer M, Gunes A, Sahin M, Simsek H. Long-term follow-up of P-wave duration and dispersion in patients with mitral stenosis. *Pacing Clin Electrophysiol* 2008;31(12):1620–4.
[40] Holmqvist F, Platonov PG, Solomon SD, Petersson R, McNitt S, Carlson J, et al. P-wave morphology is associated with echocardiographic response to cardiac resynchronization therapy in MADIT-CRT Patients. *Ann Noninvasive Electrocardiol* 2013;18(6):510–8.
[41] Liu G, Tamura A, Torigoe K, Kawano Y, Shinozaki K, Kotoku M, et al. Abnormal P-wave terminal force in lead V 1 is associated with cardiac death or hospitalization for heart failure in prior myocardial infarction. *Heart Vessels* 2013;28(6):690–5.
[42] Kamel H, Soliman EZ, Heckbert SR, Kronmal RA, Longstreth Jr W, Nazarian S, et al. P-wave morphology and the risk of incident ischemic stroke in the Multi-Ethnic Study of Atherosclerosis. *Stroke* 2014;45(9):2786–8.
[43] Şap F, Karataş Z, Altin H, Alp H, Oran B, Baysal T, et al. Dispersion durations of P-wave and QT interval in children with congenital heart disease and pulmonary arterial hypertension. *Pediatr Cardiol* 2013;34(3):591–6.
[44] Bandorski D, Bogossian H, Ecke A, Wiedenroth C, Gruenig E, Benjamin N, et al. Evaluation of the prognostic value of electrocardiography parameters and heart rhythm in patients with pulmonary hypertension. *Cardiol J* 2016;23(4):465–72.
[45] Guler H, Seyfeli E, Sahin G, Duru M, Akgul F, Saglam H, et al. P wave dispersion in patients with rheumatoid arthritis: its relation with clinical and echocardiographic parameters. *Rheumatol Int* 2007;27(9):813–8.
[46] Razazian N, Hedayati N, Moradian N, Bostani A, Afshari D, Asgari N. P wave duration and dispersion and QT interval in multiple sclerosis. *Mult Scler Relat Disord* 2014;3(5):662–5.
[47] Aksoy H, Okutucu S, Sayin B, Ercan E, Kaya E, Ozdemir O, et al. Assessment of cardiac arrhythmias in patients with ankylosing spondylitis by signal-averaged P wave duration and P wave dispersion. *Eur Rev Med Pharmacol Sci* 2016;20(6):1123–9.
[48] R Development Core Team. R: A language and environment for statistical computing. 2013.
[49] Borchers HW. pracma: Practical Numerical Math Functions. 2018. https://cran.r-project.org/web/packages/pracma/index.html.
[50] Charles J. Geyer LTJ. mcmc: Markov Chain Monte Carlo. 2019. https://cran.r-project.org/web/packages/mcmc/index.html.