Liet. matem. rink, **47**, *spec. nr.*, 2007, 489–495

Determination of number of clusters of ECG P-wave morphology estimates by testing hypothesis about uniformity of distribution

Renata ŠIMOLIŪNIENĖ, Algimantas KRIŠČIUKAITIS,
Viktoras ŠAFERIS, Violeta ŠIMATONIENĖ (Kaunas University of Medicine)
e-mail: renatasim@gmail.com

Abstract. Cardiac output is controlled by the autonomic nervous system by changing the heart rate and/or the contractions of the heart muscle in response to the hemodynamic needs of the whole body. This control is a result of permanent competition between the sympathetic and the parasympathetic nervous systems. Malfunction of these mechanisms causes the postural orthostatic tachycardia syndrome and/or the chronic fatigue syndrome. Evaluation of functionality and efficiency of the control mechanisms can give valuable diagnostic information in the early stages of dysfunction of the heart control systems and help to monitor the healing process or rehabilitation period after interventions. Quantitative evaluation of ECG P-wave changes evoked by an orthostatic test (which evokes a sudden misbalance in the interplay between the sympathetic and the parasympathetic heart control) by using a newly developed method based on the principal component analysis and clusterization by testing statistical hypothesis of uniformity provides a quantitative estimate for functionality and efficiency of the heart rate control mechanisms.

Keywords: sympathetic and parasympathetic nervous systems, ECG P-wave morphology, principal component analysis, clusterization by testing statistical hypothesis of uniformity.

Introduction

Cardiac output is controlled by the autonomic nervous system by changing the heart rate and/or contractions of the heart muscle in response to the homodynamic needs of the whole body. Sympathetic influences accelerate the heart rate, increase the contractions of the heart muscle whereas parasympathetic influences cause opposite effects. The heart rate is determined by the rate of spontaneous electrical activity of the cells usually located in the sinoatrial node. This rate is the result of permanent interplay between sympathetic and parasympathetic influences [2]. Malfunction of these mechanisms causes the postural orthostatic tachycardia syndrome and/or the chronic fatigue syndrome [14]. Some systematic diseases can impact the autonomic heart control. Impairment of the autonomic heart control is observed in various stages of diabetic neuropathy [16]. Defibrillating shocks delivered to the heart in an attempt to stop severe rhythm disorders inescapably impair the autonomic heart control [3], [11]. The same outcome may be expected after surgical interventions, catheter ablation procedures, etc. Evaluation of the functionality and efficiency of the control mechanisms can provide valuable diagnostic information in the early stages of dysfunction of the heart control systems and help to monitor the healing process or the rehabilitation period after interventions.

Different topology of sympathetic and parasympathetic nerve ganglia in the heart [9] as well as different mechanisms of action of the systems [15] cause different dynamic characteristics of the sympathetic and the parasympathetic heart rhythm control. These differences allow to identify the mechanisms and evaluate them separately by using heart rate variability criteria [15]. More detailed investigations are based on an ECG P-wave morphology analysis [6]. The P-wave shape reflects the spread of the electrical excitation front over the atria, which could be influenced by sophisticated heart rhythm control mechanisms. Electrical excitation of the heart starts in a group of cells with the highest spontaneous electrical activity called the true pacemaker that is normally situated in the sinoatrial node. There exist other groups of cells with spontaneous electrical activity that are situated either in the sinoatrial node or beyond it and are activated by electrical excitation coming from the true pacemaker. They are called latent pacemakers. Such pacemaker cell groups are most densely concentrated in the sinoatrial node [1]. The release of neuromediators from nerve terminals influences spontaneous activity of the closest pacemaker cell groups. As has been reported [9], the topology of neural ganglionated sub plexuses suggests that during a parasympathetic activity the neuromediators are released in the site of the true pacemaker, and the spontaneous activity of the true pacemaker is suppressed. Thus it is possible that a latent pacemaker, whose activity has been less impacted, would take over the role of the true pacemaker. Different sensitivity of the pacemaker cell groups to the parasympathetic activity in the sinoatrial node and other sites of the right atria as reported by [5] supports the above idea of the existence of a mechanism of the true pacemaker shifting within the sinoatrial node or beyond it. This phenomenon has been studied previously [13]. Shifting of the pacemaker site causes specific changes in the heart rate [5]. The electrical excitation front spreads over the atria having originated from the true pacemaker site. If the location of the true pacemaker shifts, this will be reflected as changes in the electrical excitation front spread and ECG P-wave morphology [6]. The more distant are the true pacemaker shifts, the more pronounced are the changes in the P-wave morphology.

In this study we demonstrated how P-wave changes evoked by an orthostatic test (which evokes sudden misbalance in the interplay between the sympathetic and the parasympathetic heart control) could be quantitatively evaluated by using the method based on the principal component analysis. Our preliminary results showed that these quantitative estimates form certain clusters when represented in multidimensional space what could reflect a pacemaker shift between certain anatomical locations. The estimates of centers of these clusters could reflect typical P-wave shapes related to certain pacemaker locations. Such evaluation of ECG P-wave morphology changes during evoked sympathetic/parasympathetic misbalance could reveal valuable information about status of autonomous heart activity control of patient.

Applied methods

Twelve lead ECG signal recordings during orthostatic test were used for searching quantitative estimate of efficiency of the heart rate control mechanisms (balance of the sympathetic and the parasympathetic nervous systems). Control group recordings were made from healthy patients ($n = 20$) in age of 21 ± 2 years. Other recordings were

made from patients ($n = 10$) in age between 45 and 71 years, part of them ($n = 10$) had transient periods of atrial fibrillation.

The vectors of samples representing P-wave in each of the 12 ECG leads consisting of 70 samples were connected one by one into one vector consisting of 840 samples in total. The obtained vector represented P-wave of one ordinary cardiocycle. All 840 sample-long vectors representing P-waves in one ECG recording during an orthostatic test were joined into the following matrix:

$$x = \begin{matrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,1} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{matrix}, \tag{1}$$

where n – number of the P-wave shapes during one cardio cycle, m – number of cardio cycles of one test recording. An example of the standard P-wave two dimensional array during one orthostatic test is given in Fig. 1. The principal component analysis was done for P-wave array of each patient. This method of the signal decomposition represents the signal by finite sum for concrete signal optimal, ortonormal basis functions multiplied by corresponding coefficients (2). It is linear combination of basic functions φ and their coefficients.

$$a_i = \sum_{k=1}^{n} w_{i,k}\varphi_k. \tag{2}$$

Coefficients, which describe P-wave vector of each cardio cycle, are calculated by formula (3):

$$w_i = \Phi^T x_i. \tag{3}$$

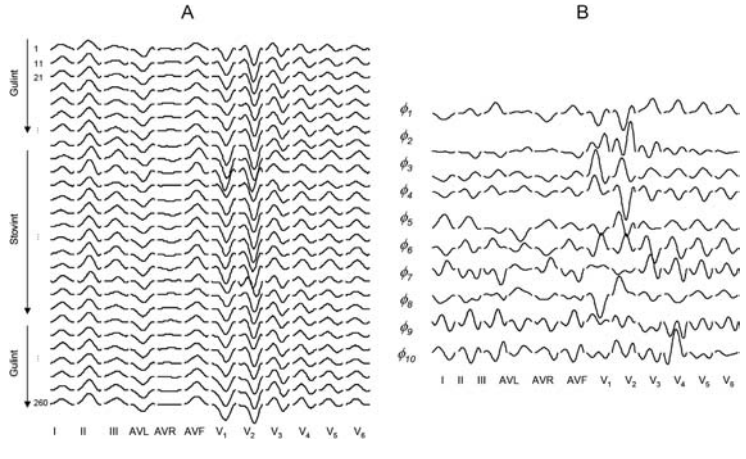

Fig. 1. Variety of P-wave during typical recording (A), the shapes of first 10 principal components (B).

These ortonormal coefficients of the basic functions picture P-wave form during each cardio cycle as a point in the n-dimensional orthogonal space, where n is minimal sufficient number of basic functions. It is defined according to methodology, which is more generally described [7]. It is mentioned in previous researches that several shapes of P-waves during all orthostatic test recordings and images of cardio cycles, which represent these waves group, forming clusters of special form. Those clusters were separated by testing statistical hypothesis of uniformity [12]. For each vector w_i Euclidian distances till all other vectors were calculated:

$$d_{ij} = \sqrt{\sum_{k=1}^{m} (w_{ik} - w_{jk})^2}, \tag{4}$$

k-closest neighbors were selected. Distances till them were denoted $d'_{i1}, \ldots, d'_{ik}$, where $d'_{i1} \leqslant d'_{i2} \leqslant \ldots \leqslant d'_{ik}$. We introduce the null hypothesis, that all points around vector w_i are equally distributed. The lth neighborhood of the vector w_i is called sphere, which radius is equal to d_{il} – the Euclidian distance till lth closest neighbor. Let Q is a part of R^m – the smallest parallelepiped, which contains all data. Each side of this parallelepiped is calculated according to formula (5):

$$a_j = \left(\max_i w_{ij} - \min_i w_{ij} \right), \quad j = 1, \ldots, m. \tag{5}$$

The volume of the parallelepiped is calculated according to formula (6):

$$V_Q = \prod_{j=1}^{m} a_j. \tag{6}$$

Assuming that all points are uniformly distributed in Q, the probability that any point will get in to lth neighborhood of the vector w_i is equal $p = V_{il}/V_Q$, where

$$V_{il} = \begin{cases} \dfrac{\pi^b d_{il}^{2b}}{b!}, & \text{if } d = 2b, \\ \dfrac{2 \cdot (2\pi)^b d_{il}^{2b+1}}{(2b+1)!!}, & \text{if } d = 2b+1, \end{cases} \tag{7}$$

d – dimension of the data, b – whole number of the fraction $d/2$. We define the critical value l', the lowest value satisfying the condition (8):

$$\sum_{i=l'}^{m} C_n^i p^i (1-p)^{m-i} \leqslant \alpha, \tag{8}$$

where α is the significance level, usually $\alpha = 0,05$, $p = V_{il}/V_Q$.

If $l > l'$, then hypothesis H_0 is rejected and lth neighborhood is considered to be an accumulation. The vector w_i is called k-dense if all lth neighbors of vector w_i for $l = 2, \ldots, k$ are accumulations. Here k is an input parameter between 3 and 10. Non

k-dense points are temporarily filtered, that is discarded from the sample. The condition of k-density is tested for remaining points. The filtering is over then all vectors become k-dense. The kth neighborhoods of points remaining after filtration are accumulations. On the basis of accumulations, the nuclei and number of clusters are determined as follows. Each vector and its k closest neighbors construct primary clusters. These clusters are merged together if overlap more than $k/2$ their vectors. The iteration algorithm is over then no one of clusters have more then $k/2$ overlapping vectors. After this iteration m vectors are left. The formation of cluster nuclei is accomplished by excluding cases attributed to the nuclei of several clusters. Let us define the vectors attributed and not attributed to the cluster nuclei, respectively, as $(w_1^{(1)}, w_2^{(1)}, \ldots, w_{m_1}^{(1)})$ and $(w_1^{(0)}, w_2^{(0)}, \ldots, w_{m_2}^{(0)})$.

Each unclassified vector we attribute to the nearest cluster. The distance from an unclassified vector to the nearest cluster is determined from the formula (9):

$$\min_j \| w_i^{(0)} - w_j^{(1)} \|. \tag{9}$$

By repeating this stage for each of the remaining cases we accomplish the classification.

Results

The variety of P-wave shape during typical ECG recording during orthostatic test is presented on Fig. 1A. The shapes of first 10 principal components (basis functions) for optimal representation of these P-wave shapes are presented on Fig. 1B. Numeric estimates of P-wave shapes of all cardio cycles during this recording in three-dimensional space is presented in Fig. 2A. Actually 7 principal components (basis functions) we needed for optimal representation in this case and all further calculations were made using 7 ones, but due to visualization possibilities we presented only first three. The coefficients of principal components were forming from one till three clusters (as in

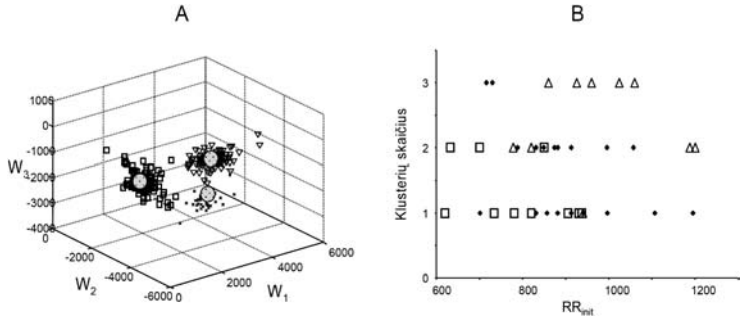

Fig. 2. P-wave shape variety presented as points in 3-dimensional space (A), number of clusters in all recordings against average initial RR interval values during supine position (B). Diamonds – control group, squares – elderly healthy patients, triangles – elderly patients with transient atrial fibrillation periods.

this presented case). We expected slower heart rate during supine position to be related with higher parasympathetic activity and more significant P-wave shape changes and more clusters in multidimensional representation we can expect during orthostatic test. However only weak such correlation one can notice in control group data presented in Fig. 2B. In elderly healthy patients we see 1–2 clusters, while 2–3 clusters we see in elderly patients with transient periods of atrial fibrillation.

Discussion

Cluster analysis of P-wave shape estimates obtained by means of principal component analysis revealed new possibilities for analysis and interpretation of data reflecting electrical excitation spread in the heart. We expect that clusters of P-wave shape estimates in multidimensional space reflect certain anatomical positions of most electrically active cell groups – pacemakers in the right atrium. True pacemaker shift evoked during orthostatic test is reflected as the shift of n-dimensional P-wave shape estimate from one cluster to the another. Cluster dimensions could be predicted by several factors: starting from heart axis movement, noises in the registered signals till sympathetic/parasympathetic impact on electrical excitation spread. However further detail investigations are needed to prove it.

Comparatively big variety in number of clusters in control group probably is a reflection of their physical condition. At the same time only 1 or 2 clusters in elderly patients reflects age related decay in autonomous heart control [8]. Two or three clusters in patients of third group (elderly persons with transient periods of atrial fibrillation) reflect increased parasympathetic activity related with atrial fibrillation [10].

References

1. M.R. Boyett, H. Honjo, I. Kodama, The sinoatrial node, a heterogeneous pacemaker structure, *Cardiovasc. Res.*, **47**(4), 658–687.
2. R. Hainsworth, The control and physiological importance of heart rate, in: M. Malik and A.J. Camm (Eds.), *Heart Rate Variability*, Futura, New York (1995), pp. 3–19.
3. M. Ito, H.P. Pride, D.P. Zipes, Defibrillating shocks delivered to the heart impair efferent sympathetic responsiveness, *Circulation*, **88**(6), 2661–2673 (1993).
4. G. Jacob, F. Costa, I. Biaggioni, Spectrum of autonomic cardiovascular neuropathy in diabetes, *Diabetes Care*, **26**(7), 2174–2180 (2003). Erratum in: *Diabetes Care*, **26**(9), 2708 (2003).
5. I. Kodama, M.R. Boyett, R. Suzuki, H. Honjo, J. Toyama, Regional differences in the response of the isolated sino-atrial node of the rabbit to vagal stimulation, *J. Physiol.*, **495** (Pt3), 785–801 (1996).
6. A. Krisciukaitis, F. Bukauskas, V. Adomonis, K. Lukosevicius, K. Muckus, Changes of orthogonal leads during pacemaker migration, in: *Electrophysiology and Surgery of Cardiac Arhythmias*, Mokslas, Vilnius (1987), pp. 41–46.
7. A. Krisciukaitis, M. Tamosiunas, P. Jakuska, R. Veteikis, R. Lekas, V. Saferis, R. Benetis, Evaluation of ischemic injury of the cardiac tissue by using the principal component analysis of an epicardial electrogram, *Comput. Methods Programs Biomed.*, **82**(2), 121–129 (2006).
8. T.B. Kuo, T. Lin, C.C. Yang, C.L. Li, C.F. Chen, P. Chou, Effect of aging on gender differences in neural control of heart rate, *Am. J. Physiol.*, **277**(6 Pt 2), H2233-9 (1999).
9. D.H. Pauza, V. Skripka, N. Pauziene, R. Stropus, Morphology, distribution, and variability of the epicardiac neural ganglionated subplexuses in the human heart, *Anat. Rec.*, **259**(4), 353–382 (2000).
10. B. Olshansky, Interrelationships between the autonomic nervous system and atrial fibrillation, *Prog. Cardiovasc. Dis.*, **48**(1), 57–78 (2005).

11. L.B. Rigden, R.D. Mitrani, H.N. Wellman, L.S. Klein, W.M. Miles, D.P. Zipes, Defibrillation shocks over epicardial patches produce sympathetic neural dysfunction in man, *J. Cardiovasc. Electrophysiol.*, **7**(5), 398–405 (1996).
12. V. Saferis, L.A. Vilkauskas, Cluster analysis by testing the statistical hypothesis of uniformity, *Statistics in Medicine*, **15**, 817–821 (1996).
13. N. Shibata, S. Inada, K. Mitsui, H. Honjo, M. Yamamoto, R. Niwa, M.R. Boyett, I. Kodama, Pacemaker shift in the rabbit sinoatrial node in response to vagal nerve stimulation, *Exp. Physiol.*, **86**(2), 177–184 (2001).
14. J.M. Stewart, Autonomic nervous system dysfunction in adolescents with postural orthostatic tachycardia syndrome and chronic fatigue syndrome is characterized by attenuated vagal baroreflex and potentiated sympathetic vasomotion, *Pediatr. Res.*, **48**(2), 218–226 (2000).
15. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. *Eur. Heart J.*, **17**, 354–381 (1996).
16. D. Ziegler, D. Dannehl, H. Muhlen, M. Spuler, F.A. Gries, Prevalence of cardiovascular autonomic dysfunction assessed by spectral analysis, vector analysis, and standard tests of heart rate variation and blood pressure responses at various stages of diabetic neuropathy, *Diabet Med.*, **9**, 806–814 (1992).

REZIUMĖ

R. Šimoliūnienė, A. Kriščiukaitis, V. Šaferis, V. Šimatonienė. EKG P-bangos klasių skaičiaus parinkimas, tikrinant statistines hipotezes apie pasiskirstymo tolygumą

Širdies veikla autonomiškai reguliuojama priklausomai nuo organizmo aktyvumo ir su tuo susijusios medžiagų apykaitos (metabolizmo) intensyvumo. Reguliavimą atlieka konkuruodamos simpatinė ir parasimpatinė nervų sistemos. Šių reguliacinių mechanizmų sutrikimai sukelia ortostatinę tachikardiją, nuolatinio nuovargio sindromą. Kiekybinių EKG P-bangos formų vertinimas ortostatinio mėginio metu atliktas naudojant pagrindinių komponenčių analizę ir įverčių klasterizaciją, tikrinant statistines hipotezes apie duomenų pasiskirstymo tolygumą, suteikia diagnostiškai svarbios informacijos ankstyvuose reguliacinių sutrikimų perioduose bei leidžia vertinti reguliacinių sistemų atsitatymą po intervencijų, sutrikdžiusių normalų jų funkcionavimą.