Unsupervised k-mean classification of atrial electrograms from human persistent atrial fibrillation

Tiago P. Almeida[1], Diogo C. Soriano[2], Xin Li[3], Gavin S. Chu[3,4], João L. Salinet[2], Fernando S. Schlindwein[5,6], Peter J. Stafford[4,5], G. André Ng[3,4,5], Takashi Yoneyama[1]

[1]Aeronautics Institute of Technology, Brazil
[2]Biomedical Engineering, Centre for Engineering, Modelling and Applied Social Sciences, Federal ABC University, Brazil
[3]Department of Cardiovascular Sciences, University of Leicester, Leicester, UK
[4]University Hospitals of Leicester NHS Trust, UK
[5]National Institute for Health Research Leicester Cardiovascular Biomedical Research Centre, UK
[6]Department of Engineering, University of Leicester, Leicester, UK

Abstract

The dichotomous criterion for atrial electrogram (AEG) classification as proposed by commercial systems (normal/fractionated) to guide ablation has been shown insufficient for persistent atrial fibrillation (persAF) therapy. In this study, we used unsupervised classification to investigate possible sub-groups of persAF AEGs. 3745 bipolar AEGs were collected from 14 persAF patients after pulmonary vein isolation. Automated AEG classification (normal/fractionated) was performed using the CARTO criterion (Biosense Webster). The CARTO attributes (ICL, ACI and SCI) were used to create a 3D space distribution. K-mean with five groups was implemented. Group 1 (43%) represents normal AEGs with low ICL, high ACI and SCI. Groups 2 (9%) and 3 (9%) have shown similar low ICL, but Group 3 has shown AEGs with short activation intervals, as opposed to Group 2. Group 4 (23%) suggests moderated fractionation, with high ACI but low SCI. Group 5 (15%) has shown highly fractionated AEGs with high ICL, low ACI and SCI. The three attributes were significantly different among the five groups (P<0.0001), except ICL between Groups 3 and 4 (P>0.999) and SCI between Groups 3 and 5 (P>0.999). The five sub-groups of AEGs found by the k-mean have shown distinct characteristics, which could provide a more detailed characterization of the atrial substrate during ablation.

1. Introduction

Atrial fibrillation (AF) is the most common sustained arrhythmia found in the clinical practice, and it is a leading cause of stroke [1]. Although pulmonary vein isolation (PVI) has been proved effective in treating paroxysmal AF, the identification of critical areas for successful ablation in patients with persistent AF (persAF) remains a challenge due to an incomplete understanding of the underlying pathophysiology of the arrhythmia [1].

Different methods have been introduced to identify atrial regions responsible for AF perpetuation to guide ablation – such as functional re-entries and fractionated atrial electrograms (AEGs) [2, 3]. The latter is of particular interest during persAF ablation: fractionated activity has been linked to i) random activations from meandering wavelets that propagate through the atria; ii) underlying anisotropic conduction in the atrial remodelled tissue and; iii) the occurrence of wave breaks or wave collisions in the atrial tissue [4].

Commercial electroanatomic mapping systems introduced dichotomous criterion for automated AEG classification based on the absence or presence of fractionation. This strategy, however, has been shown insufficient for persAF ablative therapy, resulting inconsistent outcomes possibly due to methodological heterogeneities [4], and poor understanding of the underlying AF dynamics [5-7]. Consequently, the simplistic approach for AEG classification proposed by commercial systems might be incomplete to properly characterise the underlying atrial substrate during persAF.

1.1. Unsupervised classification

Unsupervised classification is often used for unlabelled data – *i.e.*, data without defined categories or groups. It can be used to cluster groups of similar archetype within the data based on predefined characteristics of the data distribution [8].

The k-mean algorithm is one of the many available

methods for unsupervised classification. K-mean divides N observations with P dimensions (variables) into k clusters – defined by the user – so that the within-cluster sum of squares is minimized. The algorithm separates the data into spherical clusters by finding a set of cluster centres, assigning each observation to a cluster based on the squared Euclidean distance as the measure of dissimilarity between a data point and the cluster centres, determining new cluster centres, and repeating this process [9].

In the present work, we sought to investigate unsupervised k-mean classification in a three-dimensional space defined by the attributes calculated by a currently available commercial system. Possible sub-groups of persAF AEGs were investigated, expanding the traditional dichotomous AEG classification proposed by the commercial system.

2. Methods

2.1. Electrophysiological study

The population consisted of 14 patients referred to Glenfield Hospital, UK, for catheter ablation of persAF. All patients were in AF at the start of the procedure. All procedures were performed with full informed consent.

3D LA geometry was created within Ensite NavX^{TM} (St. Jude Medical, St. Paul, Minnesota) using a deflectable, variable loop circular PV mapping catheter (Inquiry Optima, St. Jude Medical). Sequential point-by-point bipolar AEGs were collected from different atrial regions. In all cases, PVs were silent and all patients were in AF during signal collection. Sinus rhythm following AEG-guided ablation was achieved in all cases.

2.2. Signal processing

A total of 3745 AEGs were collected (3413 from the left and 332 from the right atrium), with a sampling frequency of 1200 Hz, and embedded band-pass filtered within 30 – 300 Hz. Each AEG was exported from NavX with 8 s. A stationary wavelet transform filter was implemented based on a previously described method to further reduce both baseline oscillations and high frequency noise [10]. For baseline oscillations, the AEGs were decomposed with Daubechies D11 wavelet into details 8, corresponding to the frequency band between 0 – 2.34 Hz, which was set to zero. For the high frequency noise, the AEGs were decomposed with Haar wavelet into details 7. Level 1 corresponds to frequency band between 300 – 600 Hz, which has no electrophysiologic relevance. Hence, it was assumed that the presence of a white noise – that affects the frequency spectrum homogeneously – would be more evident in this frequency band with variance σ_w^2. An adaptive threshold was calculated for each AEG, accordingly:

$$T_r = \sigma_w^2 \sqrt{2 \cdot \ln N} \qquad (1)$$

where N is the length of the AEG. The threshold T_r represents the amplitude level of the assumed white noise distributed in the AEG. This threshold was then applied in all the levels of the filter bank. At each level, amplitudes higher than the threshold were conserved, while amplitudes below the threshold were suppressed. The resulting filtered AEGs were computed with the levels after thresholding with the inverse wavelet transform.

2.3. Data analysis

The CARTO criterion (Biosense Webster, Diamond Bar, California) for AEG classification has been explained previously [4]. Briefly, the algorithm identifies fractionated intervals based on peaks and troughs on the AEG that occur within a certain amplitude and duration. The algorithm then calculates the number of marked intervals, the average of their duration and the shortest interval (respectively, the Interval Confidence Level – ICL; Average Complex Interval – ACI; and Shortest Complex Interval – SCI). These attributes were calculated on the filtered AEGs, followed by the automated classification considering the CARTO criterion: ICL≥12 (normalised for 8 s), ACI≤82 ms and SCI≤58 ms [4]. The CARTO attributes were used to create a three-dimensional space distribution in which k-mean with Euclidean distance criterion was implemented to identify sub-groups of AEGs. The groups found by the k-mean were then compared with each other.

2.4. Statistics

All values are expressed as median ± interquartile range. Non-parametric unpaired sets of data were analysed using the Mann–Whitney test, while non-parametric unpaired multiple data were analysed using the Kruskal-Wallis test with Dunn's multiple comparisons. Similarities between two probability distributions was estimated with the Kullback-Leibler (KL) divergence. P-values of less than 0.05 were considered statistically significant.

3. Results

K-mean with k=5 was implemented on the distribution formed by the CARTO attributes, as illustrated in Figure 1. 630 AEGs (17% of the total) were classified as fractionated according to the CARTO criterion, while 3172 (83%) were classified as normal.

The AEGs classified as normal by CARTO were further divided in four groups (groups 1 to 4) by the k-mean, each group with very specific characteristics. Group 1 (43%) represents normal AEGs with low ICL, high ACI and SCI. Groups 2 (9%) and 3 (9%) have shown similar low ICL,
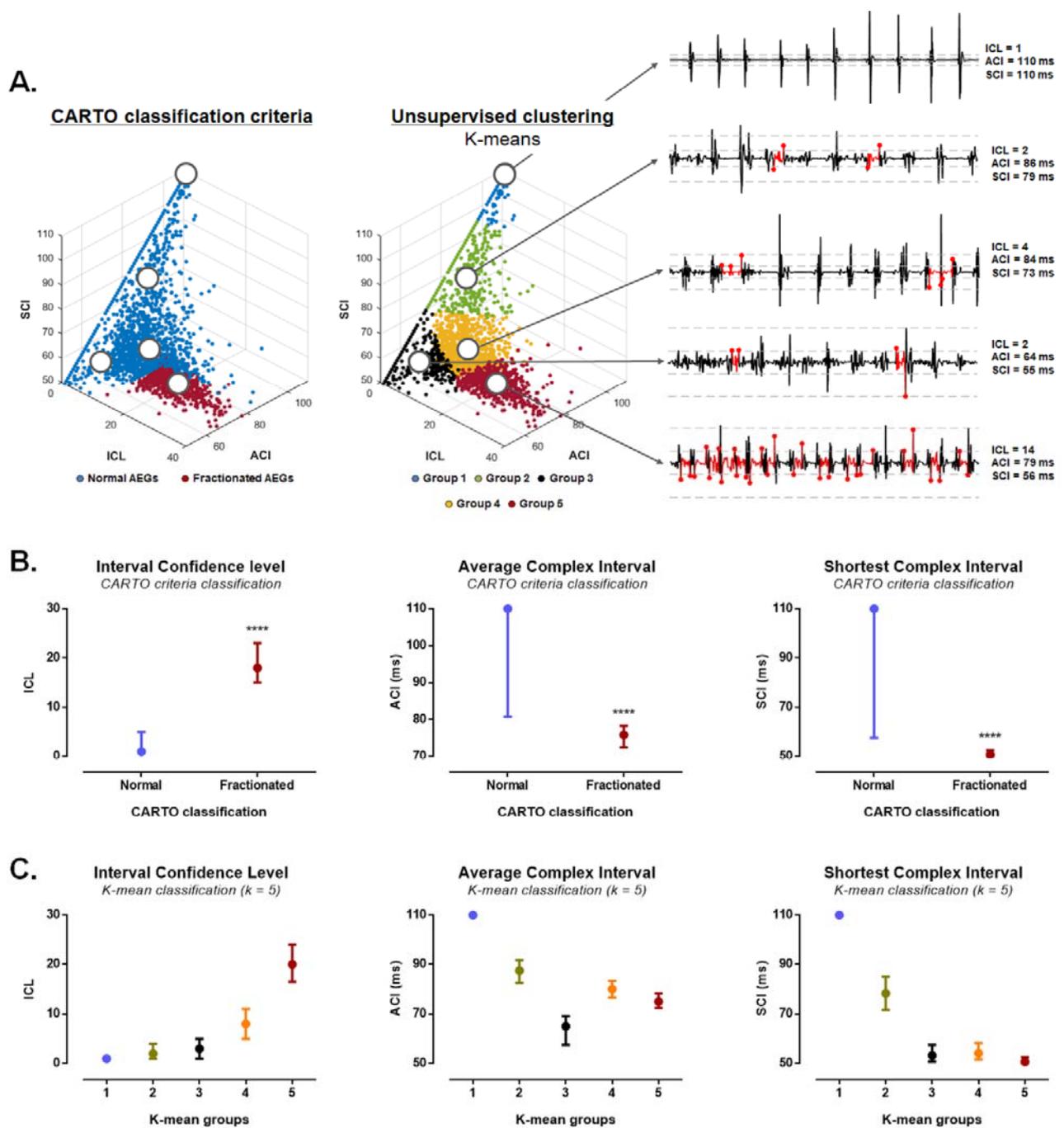
Figure 1. A. Left-hand side: the three-dimensional space distribution of the attributes calculated by CARTO (ICL, ACI and SCI) color-coded with the dichotomous AEG classification from the CARTO criterion (normal vs fractionated AEGs). Centre: the same three-dimensional space distribution color-coded with the clusters found by the k-mean (*k*=5). Right-hand side: illustration of typical AEGs found in each group with the annotations for the AEG classification from the CARTO criterion. B. The attributes calculated by CARTO (ICL, ACI and SCI) for normal and fractionated AEGs. C. The attributes calculated by CARTO (ICL, ACI and SCI) for each cluster found by the k-mean (*k*=5). The three attributes were significantly different among the five groups (P<0.0001), except ICL between Groups 3 and 4 (P>0.999) and SCI between Groups 3 and 5 (P>0.999). **** P<0.0001.

but Group 3 has shown AEGs with shorter activation intervals when compared to Group 2. Group 4 (23%) suggests moderated fractionation, with high ACI but low SCI. Group 5 (15%) presented very similar distribution compared to the AEGs classified as fractionated by CARTO (KL=0.11). 83% of these AEGs coincided with the CARTO classification for fractionation. AEGs in this group have shown very high ICL, low ACI and SCI. The three attributes were significantly different among the five groups (P<0.0001), except ICL between Groups 3 and 4 (P>0.999), and SCI between Groups 3 and 5 (P>0.999). Other values for k have been tested, with the most evident impact on k-mean classification within groups 1 to 4, i.e., the AEGs classified as normal by CARTO would be divided in more or less groups according to k. For instance, for k=4, the AEGs classified as normal by CARTO were further divided in 3 more groups; for k=6, the AEGs classified as normal by CARTO were further divided in 5 groups, and so on. For $k \geq 6$ the groups became less separable. In all cases, group 5 was mostly unaffected.

4. Discussion and conclusion

In the present work, we have implemented a simple unsupervised classification method that revealed possible sub-groups of AEGs further to the traditional the dichotomous AEG classification proposed by commercial systems. AF is regarded as a complex arrhythmia, in which different mechanisms are likely to participate in persAF perpetuation linked to remodeled substrate, such as the rapidly discharging automatic foci [11]; the multiple wavelets hypothesis [12]; the single reentrant circuit with fibrillatory conduction [13]; the conduction dissociation between epicardial and endocardial layers [14]; and functional reentry resulting from rotors [3]. This results in an intricate structure of atrial activations such that the simplistic approach for AEG classification proposed by commercial systems might be insufficient to detect this complexity. Our results support the existence of sub-groups of AEGs with distinct morphological characteristics during persAF, and that the dichotomous AEG classification proposed by commercial systems is insufficient to discriminate them. More specifically, the AEGs classified as normal by CARTO have been divided in four sub-groups of AEGs by the k-mean (for k=5), each of which with particular characteristics that could represent different electrophysiological mechanisms. This could provide a more detailed characterization of the atrial substrate during persAF ablation in future studies.

Acknowledgements

References

[1] H. Oral, B. P. Knight, H. Tada, M. Ozaydin, A. Chugh, S. Hassan, *et al.*, "Pulmonary vein isolation for paroxysmal and persistent atrial fibrillation," *Circulation,* vol. 105, pp. 1077-81, Mar 5 2002.

[2] K. Nademanee, J. McKenzie, E. Kosar, M. Schwab, B. Sunsaneewitayakul, T. Vasavakul, *et al.*, "A new approach for catheter ablation of atrial fibrillation: Mapping of the electrophysiologic substrate," *J Am Coll Cardiol,* vol. 43, pp. 2044-2053, 2004.

[3] S. M. Narayan, D. E. Krummen, K. Shivkumar, P. Clopton, W. J. Rappel, and J. M. Miller, "Treatment of atrial fibrillation by the ablation of localized sources: CONFIRM (Conventional Ablation for Atrial Fibrillation With or Without Focal Impulse and Rotor Modulation) trial," *J Am Coll Cardiol,* vol. 60, pp. 628-36, Aug 14 2012.

[4] T. P. Almeida, G. S. Chu, J. L. Salinet, F. J. Vanheusden, X. Li, J. H. Tuan, *et al.*, "Minimizing discordances in automated classification of fractionated electrograms in human persistent atrial fibrillation," *Med Biol Eng Comput,* vol. 54, pp. 1695-1706, Nov 2016.

[5] A. Verma, C. Y. Jiang, T. R. Betts, J. Chen, I. Deisenhofer, R. Mantovan, *et al.*, "Approaches to catheter ablation for persistent atrial fibrillation," *N Engl J Med,* vol. 372, pp. 1812-22, May 7 2015.

[6] T. P. Almeida, G. S. Chu, M. J. Bell, X. Li, J. L. Salinet, N. Dastagir, *et al.*, "The temporal behavior and consistency of bipolar atrial electrograms in human persistent atrial fibrillation," *Med Biol Eng Comput,* vol. 56, pp. 71-83, Jan 2018.

[7] J. Salinet, F. S. Schlindwein, P. Stafford, T. P. Almeida, X. Li, F. J. Vanheusden, *et al.*, "Propagation of meandering rotors surrounded by areas of high dominant frequency in persistent atrial fibrillation," *Heart Rhythm,* vol. 14, pp. 1269-1278, 2017.

[8] C. M. Bishop, *Pattern recognition and machine learning.* New York: Springer, 2006.

[9] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics,* Berkeley, Calif., 1967, pp. 281-297.

[10] C. Schilling, "Analysis of Atrial Electrograms," PhD, Elektrotechnik und Informationstechnik des Karlsruher Instituts für Technologie, Karlsruhe Institute of Technology, KIT Scientific Publishing, 2012.

[11] M. Haissaguerre, P. Jais, D. C. Shah, A. Takahashi, M. Hocini, G. Quiniou, *et al.*, "Spontaneous initiation of atrial fibrillation by ectopic beats originating in the pulmonary veins," *N Engl J Med,* vol. 339, pp. 659-66, Sep 3 1998.

[12] K. T. Konings, C. J. Kirchhof, J. R. Smeets, H. J. Wellens, O. C. Penn, and M. A. Allessie, "High-density mapping of electrically induced atrial fibrillation in humans," *Circulation,* vol. 89, pp. 1665-80, Apr 1994.

[13] J. Jalife, O. Berenfeld, and M. Mansour, "Mother rotors and fibrillatory conduction: a mechanism of atrial fibrillation," *Cardiovascular Research,* vol. 54, pp. 204-216, May 2002.

[14] N. de Groot, L. van der Does, A. Yaksh, E. Lanters, C. Teuwen, P. Knops, *et al.*, "Direct Proof of Endo-Epicardial Asynchrony of the Atrial Wall During Atrial Fibrillation in Humans," *Circ Arrhythm Electrophysiol,* vol. 9, May 2016.