Predicting the Onset of Paroxysmal Atrial Fibrillation: The Computers in Cardiology Challenge 2001

GB Moody[1], AL Goldberger[2], S McClennen[2], SP Swiryn[3]

[1]Harvard-M.I.T. Division of Health Sciences and Technology, Cambridge, MA, USA
[2]Harvard Medical School, Beth Israel Deaconess Medical Center, Boston, MA, USA
[3]Northwestern University, Evanston, IL, USA

Abstract

The advent of pacing techniques for preventing the onset of atrial arrhythmias motivates the development of accurate predictors of these arrhythmias, and of paroxysmal atrial fibrillation (PAF) in particular. The goals of the second annual Computers in Cardiology Challenge were to determine if segments of ECG that do not include PAF contain information sufficient (1) to distinguish subjects at risk of PAF from others not at risk, and (2) to predict imminent PAF in at-risk subjects. Via PhysioNet, 18 teams of participants studied training and test databases containing two half-hour ECG recordings from each of 100 subjects (of whom 53 experienced PAF immediately following one of the two recordings). Results indicate that roughly 80% of the subjects can be correctly classified (as at-risk or not), and that imminent PAF can be predicted in roughly 80% of subjects at risk. The most successful approaches were based on analysis of the incidence of premature atrial complexes (PACs) and P-wave variability.

1. Introduction

Following the success of the first Computers in Cardiology Challenge[1], we have introduced a new challenge: to characterize changes in the surface electrocardiogram (ECG) immediately prior to the onset of paroxysmal atrial fibrillation (PAF), in an effort to develop a reliable and fully automated method for predicting the arrhythmia.

No such methods have previously been shown to be reliable. Twelve-lead electrocardiograms[2], signal-averaged P-wave morphology[3, 4], R-R interval dynamics[5, 6], and atrial ectopy[6, 7] have all been studied as possible predictors of the onset of PAF. Sensitive and specific non-invasive markers predicting the onset of PAF have not been determined or independently validated, however. Given recent advances in clinical electrophysiology, a prediction tool that would allow for detection of imminent atrial fibrillation may have future therapeutic ramifications.

Atrial fibrillation (AF) is the most common major cardiac arrhythmia. In the United States alone, it affects an estimated 2.2 million people, with an increasing prevalence in the elderly[8]. As the population ages, the prevalence is expected to rise; currently approximately 6% of the US population over the age of 65 are diagnosed with this arrhythmia. Management consists of heart rate control and/or prevention of recurrent fibrillation, as well as prevention of secondary complications (most often thromboembolism). Paroxysms of atrial fibrillation often precede the onset of more sustained atrial fibrillation. A Japanese study of 234 patients with atrial fibrillation found that 94 (40%) had PAF, and that sustained atrial fibrillation developed within one year in approximately one fourth of patients with PAF[9].

The development of accurate predictors of the acute onset of PAF is clinically important because of the increasing possibility of electrically stabilizing and preventing the onset of atrial arrhythmias with different atrial pacing techniques. Dual chamber atrial pacing may reduce the heterogeneity of atrial refractoriness manifested by P-wave duration changes on the surface electrocardiogram recording. Preliminary studies by Prakash and colleagues[10] have indicated that acute suppression of PAF is possible in selected patients with dual-site right atrial pacing from the coronary sinus ostium and high right atrium. The advances in antitachycardia pacing, drug management, and defibrillation may be applied to prevent the acute onset of PAF prior to the loss of sinus rhythm. The maintenance of sinus rhythm can lead to decreased symptoms, improved hemodynamics, and possibly a decrease in the atrial remodeling that causes increased susceptibility to future episodes of PAF[11]. In addition, there may be a reduction in the risk of thromboembolic events.

2. The PAF Prediction Challenge Database

Participants used a database created for the challenge and made freely available by PhysioNet[12]. The database, which remains freely available at `http://www.physionet.org/physiobank/database/afpdb/`, consists of excerpts of two-channel long-term (Holter) ECG

recordings, and is divided into a learning set and a test set of equal size. The database includes the digitized ECG signals (sampled at 128 Hz per signal, with 12-bit resolution) and a set of unaudited, automatically-generated QRS annotations (containing times of occurrence only; no attempt was made to differentiate ectopic beats from normal beats, or to correct occasional detection errors). Participants were free to use the annotations provided or to create their own based on the digitized ECG signals;

2.1. Learning set

The learning set consists of 50 record sets. Each record set contains two 30-minute records, and two 5-minute "continuation" records. All four records in each record set are excerpts of longer continuous ECG recordings of a single subject; the 50 record sets come from 48 different subjects.

Twenty-five of these record sets (those with names beginning with 'p') come from subjects who have paroxysmal atrial fibrillation (PAF). One record in each pair of consecutively numbered 30-minute records contains the ECG immediately preceding an episode of PAF, which can be verified by examining the associated continuation record. The other 30-minute record in each pair contains the ECG during a period that is distant from any episode of PAF (there is no PAF during the 45-minute periods before the beginning or after the end of the 30-minute record); the associated continuation record can be used to verify that at least the next five minutes do not contain PAF.

The other 25 record sets (those with names beginning with 'n') come from subjects who do not have documented atrial fibrillation, either during the period from which the records were excerpted or at any other time. These subjects include healthy controls, patients referred for long-term ambulatory ECG monitoring, and patients in intensive care units.

2.2. Test set

The test set is similarly constructed of 50 record sets (from 50 different subjects); unlike the learning set, there are no continuation records. The test set records are named t01, t02, ... t100. As in the learning set, pairs of consecutively numbered records come from the same long-term ECG recording of a single subject. Participants in the challenge were told that between 20 and 30 of the record sets in the test set come from subjects with PAF; the actual number (28) was not disclosed until the conclusion of the challenge. The remaining 22 record sets come from subjects without PAF.

3. Organization of the Challenge

The challenge was announced, and the PAF Prediction Challenge Database was posted, on 1 March 2001. Participants used the learning set to develop and optimize fully-automated methods for classifying the record sets and for predicting imminent PAF. Beginning on 21 April, participants were able to submit their algorithms' classifications of the test set to an automated scorer via PhysioNet.

In event 1 (PAF screening), the challenge is to classify the record sets (subjects) in the test set into PAF and non-PAF groups (in other words, can individuals at risk of PAF be identified within a larger population, based on their ECGs?). Results submitted for event 1 were given a numerical score from 0 to 50 based on the number of correct classifications. No distinction was made between false negative and false positive classification errors.

In event 2 (PAF prediction), the goal is to identify which record in each pair immediately precedes PAF, for those subjects in the PAF group (in other words, is the imminent onset of PAF predictable in an individual known to be at risk of PAF?). Since the same records were used for both events, participants were not told which subjects belonged to the PAF group; rather, they were asked to choose one record from each pair as "likely to be followed by PAF". Entries were given numerical scores based on the number of correct identifications; in order to disguise the number of PAF subjects, a point was awarded for each non-PAF subject, so that the scores ranged from 22 to 50 (although the minimum possible score was not known to the participants). The event 2 scores reported in the next section have been corrected by subtracting 22 in each case.

Entrants were permitted to submit multiple sets of results in order to permit experimentation with variations of their algorithms. To discourage attempts to discover the correct classifications by repeated submission of results, participants were allowed to submit up to six entries (total for both events) without restriction, but were required to wait between any additional entries for a period that began at one day and doubled with each subsequent entry.

Participants who wished to present their work in these proceedings needed to obtain initial results no later than 1 May 2001. Entries continued to arrive throughout the summer; the final standings were determined from the entries submitted before noon GMT on 21 September.

4. Results

In each event, the most successful participants achieved scores of roughly 80% of a perfect result. The tables show the best score achieved by each of the top-scoring entrants, together with the date when they first achieved that score and the number of entries needed to do so.

Table 1. Final results for event 1 (PAF screening).

Score	*Entrant*	*Date*	*Entries*
41/50 82%	G Schreier, P Kastner, and W Marko Austrian Research Centers Seibersdorf, Graz, Austria	17 Sep	8
40/50 80%	W Zong and RG Mark Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, USA (unofficial entry)	12 Sep	7
37/50 74%	R Sweeney and colleagues Guidant Corp., St. Paul, MN, USA	8 May	3
36/50 72%	C Maier, M Bauch, and H Dickhaus University of Applied Sciences, Heilbronn, Germany	19 Sep	2
35/50 70%	C Marchesi and M Paoletti Università di Firenze, Firenze, Italy	27 Apr	1
34/50 68%	KS Lynn and HD Chiang Cornell University, Ithaca, NY, USA	28 Apr	6
33/50 66%	CC Yang National Yang-Ming University, Taipei, Taiwan	21 Apr	4
33/50 66%	JA Kors Erasmus University, Rotterdam, The Netherlands	10 Jul	2
32/50 64%	P de Chazal and C Heneghan University of New South Wales, Sydney, Australia	13 Sep	1
32/50 64%	R Loesch	14 Sep	6

Table 2. Final results for event 2 (PAF prediction).

Score	*Entrant*	*Date*	*Entries*
22/28 79%	W Zong and RG Mark Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, USA (unofficial entry)	1 May	1
20/28 71%	G Schreier, P Kastner, and W Marko Austrian Research Centers Seibersdorf, Graz, Austria	19 Aug	2
19/28 68%	P de Chazal and C Heneghan University of New South Wales, Sydney, Australia	28 Apr	1
19/28 68%	C Maier, M Bauch, and H Dickhaus University of Applied Sciences, Heilbronn, Germany	11 Sep	3
18/28 64%	KS Lynn and HD Chiang Cornell University, Ithaca, NY, USA	29 Apr	2
17/28 61%	P Langley, D di Bernardo, J Allen, E Bowers, F Smith, S Vecchietti, and A Murray Freeman Hospital, Newcastle upon Tyne, UK	30 Apr	1
17/28 61%	D Gamberger and T Smuc Rudjer Boskovic Institute, Zagreb, Croatia	23 Aug	2
16/28 57%	CC Yang National Yang-Ming University, Taipei, Taiwan	23 Apr	1
16/28 57%	R Sweeney and colleagues Guidant Corp., St. Paul, MN, USA	8 May	1
15/28 54%	L Almarro UPV, Valencia, Spain	30 Apr	1

5. Conclusions

The questions posed by the challenge are quite difficult. The best-performing algorithms were based on incidence of isolated atrial premature complexes and on P-wave variability. Methods based on time-domain and frequency-domain analysis of heart rate variability were somewhat less successful.

Further study is needed to determine if combinations of these diverse strategies can yield further improvements in PAF screening or prediction accuracy.

Acknowledgements

The Computers in Cardiology Challenges have been conducted using the facilities of PhysioNet, a public service fo the Research Resource for Complex Physiologic Signals, which is supported by a grant from the National Center for Research Resources of the US National Institutes of Health (P41 RR13622). We thank Gerold Porenta and the board of Computers in Cardiology for their continuing and enthusiastic support. GBM thanks his coauthors for their generous contributions of carefully selected data to the PAF Prediction Challenge Database. Thanks also to Isaac Henry, Christoph Maier, Joseph Mietus, and Juan Millet for timely and valuable feedback on the database. Finally, we thank the 16 teams of participants from 10 nations, who accepted this difficult challenge; ten of them report their results elsewhere in this volume.

References

[1] Moody GB, Mark RG, Goldberger AL, Penzel T. Stimulating rapid research advances via focused competition: The Computers in Cardiology Challenge 2000. In Computers in Cardiology 2000. Piscataway, NJ: IEEE Press, 2000; 207–210.

[2] Dilaveris PE, Gialafos EJ, Sideris SK, Theopistou AM, Andrikopoulos GK, Kyriakidis M, Gialafos JE, Toutouzas PK. Simple electrocardiographic markers for the prediction of paroxysmal idiopathic atrial fibrillation. Am Heart J 1998; 135:733–738.

[3] Ishimoto N, Ito M, Kinoshita M. Signal-averaged P-wave abnormalities and atrial size in patients with and without idiopathic paroxysmal atrial fibrillation. Am Heart J 2000; 139:684–689.

[4] Amar D, Roistacher N, Zhang H, Baum MS, Ginsburg I, Steinberg JS. Signal-averaged P-wave duration does not predict atrial fibrillation after thoracic surgery. Anesthesiology 1999;91:16–23.

[5] Vikman S, Makikallio TH, Yli-Mayry S, Pikkujamsa S, Koivisto AM, Reinikainen P, Airaksinen KEJ, Huikuri HV. Altered complexity and correlation properties of R-R interval dynamics before the spontaneous onset of paroxysmal atrial fibrillation. Circulation 1999;100:2079–2084.

[6] Hnatkova K, Waktare JEP, Murgatroyd FD, Guo X, Baiyan X, Camm AJ, Malik M. Analysis of the cardiac rhythm preceding episodes of paroxysmal atrial fibrillation. Am Heart J 1998;135:1010–1019.

[7] Kolb C, Nurnberger S, Ndrepepa G, Schreieck J, Zrenner B, Karch M, Schmitt C. Modes of initiation of paroxysmal atrial fibrillation: an analysis of 157 spontaneously occurring episodes using 12-lead Holter monitoring. PACE 2000; 23(4):607.

[8] Feinberg WM, Blackshear JL, Laupacis A, Kronmal R, Hart RG. Prevalence, age distribution, and gender of patients with atrial fibrillation. Arch Intern Med 1995;155:469–473.

[9] Takahashi N, Seki A, Imataka K, Fujii J. Clinical features of paroxysmal atrial fibrillation: an observation of 94 patients. Jpn Heart J 1981;22:143–149.

[10] Prakash A, Saksena S, Hill M, Krol RB, Munsif AN, Giorgberidze I, Mathew P, Mehra R. Acute effects of dual-site right atrial pacing in patients with spontaneous and inducible atrial flutter and fibrillation. J Am Coll Cardiol 1997; 29:1007–1014.

[11] Prystowsky EN. Management of atrial fibrillation: therapeutic options and clinical decisions. Am J Cardiol 2000; 85:3D–11D.

[12] Moody GB, Mark RG, Goldberger AL. PhysioNet: A research resource for studies of complex physiologic and biomedical signals. In Computers in Cardiology 2000. Piscataway, NJ: IEEE Press, 2000; 179–182.

Address for correspondence:

George B. Moody
MIT Room E25-505A, Cambridge, MA 02139 USA.
george@mit.edu