Personalized and automated remote monitoring of atrial fibrillation
Arnaud Rosier, Philippe Mabo, Lynda Temal, Pascal van Hille, Olivier Dameron, Louise Deléger, Cyril Grouin, Pierre Zweigenbaum, Julie Jacques, Emmanuel Chazard, et al.

To cite this version:
Personalized and Automated
Remote Monitoring of Atrial Fibrillation

Arnaud Rosier 1,2,3,* , Philippe Mabo 1,4, Lynda Temal 5, Pascal Van Hille 5, Olivier Dameron 6, Louise Deléger 7, Cyril Grouin 7, Pierre Zweigenbaum 7, Julie Jacques 8, Emmanuel Chazard 9, Laure Laporte 10, Christine Henry 10 and Anita Burgun 3,11

1 INSERM CIC-IT 804, CHU Pontchaillou, CCP, rue Henri le Guilloux, Rennes, France
2 Service de Rythmologie, Hôpital Privé Jacques Cartier, Groupe GDS, 6 Avenue du Noyer Lambert, Massy 91300, France
3 INSERM UMR_S 1138 Eq 22, Paris Descartes University, Paris, France
4 Cardiology Department, CHU Pontchaillou, CCP, rue Henri le Guilloux, Rennes, France
5 LTSI (Inserm UMR1099), Université de Rennes 1, Rennes, France
6 University of Rennes 1, UMR 6074 IRISA, Campus de Beaulieu, Rennes 35042, France
7 LIMSI-CNRS, Orsay 91405, France
8 ALICANTE SA, France
9 CERIM, Faculté de Médecine, 1 Place de Verdun, Lille, France
10 SORIN CRM SAS, Parc Noveos, 4 Avenue Réaumur, Clamart, France
11 Hôpital Européen Georges Pompidou, Paris, France

1. *Corresponding author. Tel: +33 1 60 13 61 64; fax: +33 1 60 13 66 44. E-mail address: dr.arnaud.rosier@gmail.com

Word Count: ~3300 words (excluding references)
ABSTRACT

(1) Aims
Remote monitoring of cardiac implantable electronic devices is a growing standard; yet, remote follow-up and management of alerts represents a time-consuming task for physicians or trained staff. This study evaluates an automatic mechanism based on artificial intelligence tools to filter atrial fibrillation alerts based on their medical significance.

(2) Methods
We evaluated this method on atrial fibrillation alerts notified in 60 pacemaker recipients. AKENATON prototype workflow includes two steps: natural language processing algorithms abstract the patient health record to a digital version, then a knowledge based algorithm based on an applied formal ontology allows to calculate the CHA2DS2-VASc score and evaluate the anticoagulation status of the patient. Each alert is then automatically classified by importance from low to critical, by mimicking medical reasoning. Final classification was compared to human expert analysis by two physicians.

(3) Results
1783 alerts about AF episode > 5 min in 60 patients were processed. 1749 of 1783 alerts (98%) were adequately classified and there were no underestimation of alert importance in the remaining 34 misclassified alerts.

(4) Conclusion
This work demonstrates the ability of a pilot system to classify alerts and improve personalized remote monitoring of patients. In particular, our method allows integration of patient medical history with device alert notifications, which is useful both from medical and resource-management perspectives. The system was able to automatically classify the importance of 1783 AF alerts in 60 patients, which resulted in an 84% reduction in notification workload, while preserving patient safety.

Keywords: Cardiac implantable electronic devices • Remote monitoring • Atrial fibrillation • Decision support systems • Artificial intelligence

What's new?

- We report for the first time an automatic classification method for remote atrial fibrillation alerts.
- We found our pilot mechanism to be feasible and safe.
- Moreover, it was faster compared with non-assisted physician decisions.
- This technology may help to filter non-clinically relevant alerts and highly reduce the workload for remote monitoring of AF alerts.
- With the increasing use of medical sensors, such techniques will be essential in the future for timely analysis of data sent by connected health devices.
INTRODUCTION

Remote monitoring of patients treated with cardiac implantable electronic devices (CIEDs) has been shown to improve patient follow-up and outcome by enabling early detection of arrhythmias and of technical problems with the devices (1,2). However, although remote follow-up is accepted by international guidelines as a standard alternative to clinical follow-up, adoption in daily practice remains slow (3). Although legal and reimbursement issues may account for part of this inertia, the concern that remote monitoring needs new workflow patterns and leads to an increase in workload is an important and valid issue. Therefore, new organizational models or tools have to be developed. Interesting ways to reduce workload are usually based on alert filtering by trained nurses, as proposed by Ricci & al. in the HomeGuide registry, but such workflows still require human staff to operate, and a transition period to take place (4).

The AKENATON project is a French pilot study designed to test the hypothesis that filtering the importance of alerts could be performed automatically, using algorithms that represent medical knowledge and integrate the patient’s clinical condition with device data. The aim of this study was to design a prototype mechanism for atrial fibrillation (AF) alerts (which represent the most frequent notifications in remote monitoring of CIEDs and are expected to lower patient morbidity), and to evaluate the efficacy and safety of this prototype (5).

METHODS

The design of the AKENATON pilot application was driven by the following scenario: an atrial fibrillation episode detected in a patient with a CIED implant (by means of an automatic mode switch) initiates an alert notification and triggers the automatic classification of this alert. In order to produce a clinically relevant classification of AF alerts, we needed to shift from a strictly device-centered follow-up to a patient-focused perspective, i.e. to take into account the individual clinical condition. Our method of alert classification relies on a knowledge base that includes a formal domain ontology (knowledge base). In order to assist the cardiologist in their decision making, the system takes into account the patient’s condition and extracts relevant data from the patient’s records. The prototype was evaluated on a set of 60 eligible patients sourced from a cardiology department.

Patient selection

The research protocol was approved by the ethics evaluation committee of the INSERM (IORG0003254, FWA00005831) institutional review board (IRB00003888) under the reference 13-103. The study complied with the Declaration of Helsinki, and informed consent was obtained from all participants.

We selected 60 consecutive eligible patients from the retrospective database of the cardiology department of Rennes University Hospital. All included patients were adults (>18 years old). Inclusion criteria were: (1) implantation with a Sorin REPLY dual chamber pacemaker between June 2008 and June 2010, and (2) at least one episode of sustained AF (>30s) present in the device memory during follow-up.

Because remote monitoring in Sorin devices did not allow assessment of an unlimited number of AF episodes for a given patient and in order to lower the duration threshold, the internal memory of each device was cleared, and alerts were generated according to the manufacturer’s specifications with the exception of AF duration: All episodes of at least 30 s were considered, since the minimum duration for significant AF episode remains unclear.
For each patient, the following files were extracted from the hospital information system: (1) medical reports, including clinical reports, radiology reports, and CIED implantation reports; and (2) laboratory test results, in particular international normalized ratio (INR) results.

Medical history analysis and automatic reasoning
Whenever an AF alert is generated, the AKENATON system performs the extraction of the data needed to analyze the patient’s clinical condition (Figure 1), including administrative data (sex, age), narrative patient records for outpatient visit or inpatient stay, laboratory test results, and radiology reports.

![Figure 1 – AKENATON Classification of AF alert based on patient data (details in the text). AF, atrial fibrillation; INR, international normalized ratio.](image)

**Step 1**
Information is extracted both from structured data (e.g., laboratory test results) and from text reports. Personal identifiable information such as patient name is removed from the data, then natural language-processing (NLP) techniques are used to identify key diagnoses and medications from text reports. This step includes the lexical matching of keywords, syntactic and semantic analysis, and detection of negative sentences that are frequent in clinical documents; e.g., the sentence ‘the patient had no sign of heart failure’ would be interpreted correctly as ‘the absence of heart failure’ (6). The final data extracted from the patient medical records included: patient’s age and sex, medical conditions and medications needed for CHA₂DS₂-VASc calculation used to assess the thrombo-embolic risk (7), and INR values and dates.
Along with patient data, the duration and date of AF episodes are imported into the reasoning module. This module consists of a formal web ontology language (OWL 2) as the means to represent the medical knowledge (concepts, relations between concepts) of the domain. Consequently, the system has the following reasoning capabilities.

- Automatic reasoning allows computation of the available data in order to infer implicit information from a given patient record. For instance, even if diabetes is not explicitly mentioned in the patient report, a prescription of metformin is used by the system to infer ‘diabetes’ and the diabetes score is correctly added to the overall CHA\textsubscript{2}DS\textsubscript{2}-VASc score. Automatic reasoning is also used for representing formal definitions that physicians use daily: e.g., a patient taking warfarin with current INR <2 would not be considered an anticoagulated patient.

- An additional benefit is that formal ontology allows representation of data at different levels of granularity, e.g., a patient with ‘lower extremity peripheral arterial disease’ will also match the less specific ‘vascular disease’ criterion, while ‘diabetes type 2’ is recognized as a form of ‘diabetes mellitus.’

As a result, automatic computation of the CHA\textsubscript{2}DS\textsubscript{2}-VASc score and anticoagulation status of the patient is performed, as is the classification of the importance of the AF alert (low, medium, high, or critical).

Classification of the importance of AF alert

In order to determine the importance of AF notification, an ordinal four-point scale was created by evaluating the necessity of a medical intervention and its urgency (Table 1). It is worth noting that our objective here is to score the alert importance rather than the thromboembolic risk itself. A personalized AF alert is ranked accordingly from low to critical, based on the following criteria.

- The patient’s thromboembolic risk is estimated using the CHA\textsubscript{2}DS\textsubscript{2}-VASc score. Three CHA\textsubscript{2}DS\textsubscript{2}-VASc classes were distinguished:
  - Score = 0. A score of 0 is usually assumed to reflect low or no risk, and thus such alerts are of little concern to the physician, especially if AF resumes spontaneously.
  - Score = 1: a patient presenting AF and CHA\textsubscript{2}DS\textsubscript{2}-VASc scores of ≥1 should usually be treated with continuing oral anticoagulation (OAC) therapy. Therefore, alerts from such patients’ devices should be considered of higher importance unless the patient is already being treated (7).
  - Score ≥ 2. The European Society of Cardiology guidelines recommend that these patients should receive OAC therapy, unless their bleeding risk is very high, in order to lower the risk of thrombo-embolic events.

- Duration of AF: the longest duration of AF that can be allowed before a significant increase in the patient’s thromboembolic risk occurs is subject to debate, and may be as low as 5 min (8,9). However, the minimum threshold predetermined at the time this study initiated was 10 min. By contrast, 6 h is the default duration threshold in most remote monitoring systems for reporting AF episodes. In addition, episodes of AF longer than 24 h are considered more important because diagnosed patients may require termination of AF by cardioversion. Therefore, 10 min, 6 h, and 24 h were considered as cutoff values for determining the importance of AF alerts in the current study.

- Current anticoagulation status: we defined current OAC status as current treatment with such drugs (based on medical record information), unless the last INR was strictly >2 for patients currently taking
a vitamin K antagonist INRs >90 days were discarded. Because many hospitalized patients receive non-oral anticoagulation regardless of their usual medication, we also disregarded this.

Table 1 - Classification algorithm used to determine the importance of AF alerts in a given patient (detail in the text).

<table>
<thead>
<tr>
<th>CHA2DS2-VASc (no OAC*)</th>
<th>AF duration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; 10min</td>
</tr>
<tr>
<td>0</td>
<td>low</td>
</tr>
<tr>
<td>1</td>
<td>low</td>
</tr>
<tr>
<td>2+</td>
<td>low</td>
</tr>
</tbody>
</table>

*oral anti-coagulation

Evaluation

For each AF alert, the importance was computed by the system. For each patient, the different results of the automatic computation were compared with those from a manual review of the patient’s medical record and the device’s alert record by two independent domain experts. CHA\textsubscript{2}DS\textsubscript{2}-VASc items, total score, anticoagulation therapy status according to medication and INR, duration of AF, and final alert importance were compared between system computation and manual review, with the latter considered as the reference standard.

Statistical analysis

Cohen’s Kappa and Weighted Kappa were calculated in order to measure agreement between the system and the human reference. Underestimated alert importance cases were considered to bear twice the weight of overestimated alert importance cases.

RESULTS

In the selected population of 60 patients, 1783 AF episodes were detected (mean follow-up 9 months). The number of AF episodes in patients ranged from 1 to 390 (29.7 ± 36.6); 744 episodes (42%) lasted at least 6 h, and 63 episodes (3.5%) lasted at least 24 h.

Classification of alert importance

The importance of alerts computed by the AKENATON pilot was very close to that of manual revision: 1749 of 1783 alerts (98%) were adequately classified by AKENATON (Error! Reference source not found.). At the patient level, 58 of the 60 patients had their alerts correctly classified by the system. Results are summarized in Table 2.

There were 34 misclassified alerts in 2 patients: 33 alerts in a single patient had their importance over-estimated due to an error in the CHA\textsubscript{2}DS\textsubscript{2}-VASc calculation (score of 2 instead of 1), and 1 alert in another patient was overestimated because the OAC therapy this patient was receiving was missed by the language analysis tool.

Cohen’s Kappa value between the system and the human reference was 0.93. Due to the absence of cases with underestimated alert importance, weighted Kappa was the same.
Figure 2 – Distribution of atrial fibrillation (AF) alert importance classified by human (manual) vs. AKENATON (automatic).

Table 2 - Results of AKENATON classification of importance of atrial fibrillation (AF) alerts given by AKENATON vs. the correct human reference classification

<table>
<thead>
<tr>
<th>Human Reference</th>
<th>AKENATON Application</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LOW</td>
</tr>
<tr>
<td>LOW</td>
<td>1502(49)</td>
</tr>
<tr>
<td>MEDIUM</td>
<td>147(22)</td>
</tr>
<tr>
<td>HIGH</td>
<td>88(13)</td>
</tr>
<tr>
<td>CRITICAL</td>
<td></td>
</tr>
</tbody>
</table>

Correct classifications are shown in bold. The number of alerts is given (number of patients within brackets).

CHA\textsubscript{2}DS\textsubscript{2}-VASc and anticoagulation status computation

In general, the CHA\textsubscript{2}DS\textsubscript{2}-VASc score computed by the system was very close to the human reference results. For the score items, 480 values were estimated by the system (8 criteria for 60 patients), of which 466 were correct (Figure 3). CHA\textsubscript{2}DS\textsubscript{2}-VASc score was the same for 46 of 60 patients with both the fully automatic and the human methods, while there was a single error on a CHA\textsubscript{2}DS\textsubscript{2}-VASc criterion for 14 patients (9 false-positive and 5 false-negative), which had little impact on the medical interpretation. CHA\textsubscript{2}DS\textsubscript{2}-VASc score classification (0, 1, or ≥2 as used in clinical practice) was correct for 58 patients (97%).

Anticoagulation was adequately evaluated in 57 patients (95%). The three patients with errors were two patients with interrupted vitamin K antagonist therapy, and one patient whose medical records mentioned OAC without detailing the drug, which was missed by the NLP technique. Of the 17 errors found, 16 were due to errors in this first step, and it is unlikely that these would have occurred if an accurate electronic health record (EMR) had been available.
These few errors had little impact on medical interpretation in terms of alert importance. As a result, 1749 of 1783 alerts (98%) were adequately classified and there were no underestimations.

After anonymization, an average of 27 medical documents per patient were analyzed in 52 seconds, and automatic reasoning took an additional 2 seconds (mean value). These computation times are compatible with a real-time computation triggered by alert notification.
DISCUSSION

We have developed a prototype for remote monitoring and management of AF alerts. In this study, we aimed to assess the safety (specifically, its ability to detect which patients are most at risk of thromboembolic events) and efficacy (reduction in alerts and/or their classification as important or not) of the system. Both criteria are key features for encouraging acceptance of such a system by healthcare providers in the future.

Patient safety
To monitor a patient remotely, the data transmitted by the devices must be regarded as part of the patient’s medical records, and must be combined with the clinical data from EMRs to guide medical decision-making in real time. We found that the automatic extraction of information from EMRs and the classification of alerts were feasible and accurate, with 98% of the 1783 alerts classified correctly. In two cases, the automatic classification of the alerts was associated with an overestimation of the CHA2DS2-VASc score. We did not find any case of underestimation of alert importance. Such results are compatible with optimal levels of patient safety. This study illustrates a novel approach that is likely to become needed much more in the future with the development of remote monitoring and the widespread adoption of EMRs.

Efficacy to reduce the notification workload
Of 1783 AF episodes, 744 lasted at least 6 hours, and would probably have been reported by current remote monitoring systems (providing AF alerts are not switched off). Using the AKENATON system, only 133 of these 744 alerts would have been deemed useful and reported, thus achieving a potential reduction in AF alert quantity of 82%, while preserving 100% of the relevant notifications.

It has been suggested that physicians may need to monitor even short (5 min and more) episodes of AF in patients. Despite the fact that our automatic system assessed all AF episodes without considering their duration, it still produced a reduction in the number of alerts from 1783 possible alerts to 280 relevant alerts (84% reduction); the total number of alert notifications would be a third of what would be reported with current remote monitoring systems’ settings.

Atrial arrhythmias are estimated to represent 60–70% of all pacemaker alerts (10–12). Our findings suggest that the workload related to the management of pacemaker remote monitoring for AF events could only be greatly reduced by such an automatic system, without jeopardizing the clinical benefit of such monitoring. Although others have reported that trained nurses are able to reduce alerts reported to physicians by an average of 85%, we believe our automatic method could be considered as a replacement or added first line management of data (11,13). Furthermore, it could provide key elements of a patient’s medical records to the nurse or physician in a timely manner.

We anticipate that such technology will be applicable to other types of alerts and devices such as ventricular arrhythmias in patients with implantable cardioverter defibrillators (ICDs), and that it would be of benefit to electrophysiologists if it were made widely available for daily practice.

Interoperability between remote monitoring systems and EMRs
Decision support systems such as AKENATON require device and patient data together in order to work. Device data have to be obtained from a sustainable workflow, using a common data representation and convergent pipelines. Initiatives such as IDCO (Implantable Device Cardiac Observation Profile), a standard for representing data from CIEDs, and iCardea, a European project for sharing remote monitoring and patient data together across multiple manufacturers, are milestones in the process of addressing this need (14,15). Moreover, our study is a proof of concept
demonstrating that new knowledge-based methods may assist physicians in interpreting real-time personalized data for the benefit of patients.

**Sensors and other monitoring schemes**
We support the notion that CIEDs represent the first step in using implanted and communicating devices for monitoring patients and therapy. It is expected that the integration of sensors and wireless technologies will transform medical practice in the near future. The huge amount of data generated by such monitoring devices shall not be analyzed by human-based methods. We expect that approaches based on data integration and automated reasoning, such as the method presented in this paper, will be important factors in turning this data into action for the benefit of individual patients, or for mining this data in order to discover new medical knowledge.

**Study limitations**
In our study, AF detection was diagnosed by using a device mode switch. Mode switch criteria are subject to variation in different manufacturers’ devices and models. In addition, a mode switch may be erroneously activated by atrial lead noise (external interference, myopotential for unipolar leads, lead fracture) despite the absence of supraventricular arrhythmia. More importantly, a long AF episode may be detected as multiple shorter AF episodes; this is not really relevant to this study as it is already an issue in current alerts. Nevertheless, despite these well-recognized facts, the vast majority of mode switches still occur appropriately, and transmission of the episodes recorded electrograms and simple algorithms that concatenate subsequent episodes are expected to allow the recognition of these cases.

In addition, we analyzed data from pacemakers, whereas most remote monitoring is currently performed for patients with ICDs. However, we believe that for AF detection, results are generalizable to ICDs, and we plan to extend this proof of concept to all cardiac rhythm-management devices.

Finally, this was a retrospective study, and a multicenter prospective study, including health-cost measurements and clinical safety assessments, would confirm the benefit of such a system.

**CONCLUSION**
This work demonstrates the ability of a pilot system to classify alerts and improve personalized remote monitoring of patients. In particular, our method allows integration of patient medical history with device alert notifications, which is useful both from medical and resource-management perspectives. The system was able to automatically classify the importance of 1783 AF alerts in 60 patients, which resulted in an 84% reduction in notification workload, while preserving patient safety. This method could be extremely helpful for remote monitoring of CIEDs and could be extended to other applications in the field of medical sensors. We expect to achieve further improvement and validation with the increasing use of electronic health records.

**Acknowledgements**
We thank the AKENATON project members (Alicante SA, CERIM, CIC-IT, INSERM U936, LIMSI and Sorin CRM SAS) for their commitment. We also thank N. Roederer (Innovelan Co.) for technical assistance.

**Conflict of interest**
A.R. has received small consultancy fees from Medtronic, St. Jude Medical, and Sorin Group, and research grants from Medtronic, Inc., Boston Scientific Corp. Ph.M. has received speaker honoraria and consultancy fees and research grants from St. Jude Medical, Sorin Group, Medtronic, Inc., Boston

Funding
This work was supported by the Agence Nationale Pour la Recherche – Technologies pour la Santé [ANR-07-TecSan-001].

REFERENCES


Representative Figure

1. Device
   - 60 patients
   - 1783 basic AF alerts
     - AF date & duration

2. Health record
   - Patient Electronic Health Record
     - Age, comorbidities, medication, INR

Step 1: Natural Language Processing

Step 2: CHA₂DS₂-VASc calculation
- Anticoagulation status assessment
- Knowledge Model-based Alert Classification

Personalized AF alert (four point scale: low to critical)