Available online at www.sciencedirect.com

ScienceDirect

Journal of Electrocardiology 50 (2017) 776 – 780

JOURNAL OF Electrocardiology

www.jecgonline.com

PDF–ECG in clinical practice: A model for long–term preservation of digital 12–lead ECG data☆

Roberto Sassi, PhD,[a,*] Raymond R. Bond, PhD,[b] Andrew Cairns,[b] Dewar D. Finlay, PhD,[b] Daniel Guldenring, PhD,[b] Guido Libretti,[c] Lamberto Isola,[c] Martino Vaglio, MS,[c] Roberto Poeta,[d] Marco Campana, MD,[d] Claudio Cuccia, MD,[d] Fabio Badilini, PhD[c]

[a] *Dipartimento di Informatica, Università degli Studi di Milano, Crema, Italy*
[b] *Faculty of Computing and Engineering, Ulster University, Belfast, Northern Ireland, UK*
[c] *AMPS-LLC, New York, NY, USA*
[d] *Fondazione Poliambulanza, Brescia, Italy*

Abstract

Background: In clinical practice, data archiving of resting 12–lead electrocardiograms (ECGs) is mainly achieved by storing a PDF report in the hospital electronic health record (EHR). When available, digital ECG source data (raw samples) are only retained within the ECG management system.
Objective: The widespread availability of the ECG source data would undoubtedly permit successive analysis and facilitate longitudinal studies, with both scientific and diagnostic benefits.
Methods & results: PDF-ECG is a hybrid archival format which allows to store in the same file both the standard graphical report of an ECG together with its source ECG data (waveforms). Using PDF-ECG as a model to address the challenge of ECG data portability, long-term archiving and documentation, a real-world proof-of-concept test was conducted in a northern Italy hospital. A set of volunteers undertook a basic ECG using routine hospital equipment and the source data captured. Using dedicated web services, PDF-ECG documents were then generated and seamlessly uploaded in the hospital EHR, replacing the standard PDF reports automatically generated at the time of acquisition. Finally, the PDF-ECG files could be successfully retrieved and re-analyzed.
Conclusion: Adding PDF-ECG to an existing EHR had a minimal impact on the hospital's workflow, while preserving the ECG digital data.

Keywords: Clinical ECG; ECG digital data; ECG long-time preservation; PDF report

Introduction

As of today, data archiving of clinically recorded 12–lead resting (10s) ECG is mainly achieved by storing a PDF file ("the ECG report") in the hospital electronic health record (EHR) of the patient. In the most frequent scenario, only the medical interpretation is preserved or, at best, a raster image of the ECG print-out is stored. Digital ECG source data is only retained (typically within the cardiology unit) when an adequate ECG management system is in place. Yet, when the hospital is spread geographically across different sites, the archiving of the source data information (*e.g.*, the ECG waveforms) is often dispersed across several systems and, most critically, it is not delivered to the central electronic patient record system.

The widespread availability and preservation of ECG digital source data would undoubtedly bring a series of advantages. For example, it could permit successive secondary data analysis and would facilitate longitudinal studies, with both scientific and diagnostic benefits. Furthermore, the possibility of maintaining the ECG waveforms would benefit the general community, permitting new ways of patient empowerment. Finally, the wealth of ECG digital data would foster big-data mining and epidemiological analysis.

In order to obtain an estimate and somehow a clearer picture of real-life scenario with respect to ECG data storage in health care facilities, we performed an informal survey interviewing 19 IT directors and cardiologists working in major hospitals from three different countries (France, Italy and the United States). The survey consisted of a few simple questions: first, we asked if the cardiology unit was equipped with an ECG management system and whether an EHRs system was in place in the hospital. In the latter case, we further inquired

☆ Preprint submitted to Journal of Electrocardiology August 9, 2017.
* Corresponding author.
E-mail address: roberto.sassi@unimi.it (R. Sassi).

http://dx.doi.org/10.1016/j.jelectrocard.2017.08.001
0022-0736/

about the storage of an ECG record, either in the form of a report printout and/or with the storage of source data. The results of the survey are reported in Table 1. Remarkably, none of the 19 hospitals has an EHR capable to store source ECG data, whereas only a few store the PDF printouts.

While it may not be representative of the entire world, the survey clearly confirms that retrieval of ECG source data, when feasible, can only be achieved from the ECG management system or through re-digitization of paper or PDF report printouts.

To address the challenge of ECG data portability and long-term preservation, a new hybrid ECG archival format (PDF-ECG), was recently proposed [1]. PDF-ECG combines the need (and the benefits) to store a standard ECG report and the digital source data by archiving both the graphical display and the raw waveform information within the same file. With PDF-ECG, digital waveforms can be obtained after proper anonymization directly from the EHR system or from the Contract Research Organization database (in the pharmaceutical context). With PDF-ECG, the patient would be able to preserve his/her own copy of the source data, similarly to what already regularly happens with other clinical exams (*e.g.*, with MRI and CT scans, where together with the paper report, a CD copy of the actual exam is typically provided). PDF-ECG is not another standard or a new digital format (there are arguably already too many, refer to [2]). It is instead a way to obtain long-term preservation of the digital ECG waveforms exploiting the widespread and mature technology of PDF. In technical terms, it is a hybrid archiving format where the printout and the digital data, used to obtain it, coexist in the same object. In a very basic description, the ECG digital data are "attached" to the graphical report in a single PDF/A-3u file (ISO 19005-3:2012), which assures, by definition long-term durability. Today the graphical PDF report and the digital data are two different items, not linked together and often sitting in different servers. With PDF-ECG they are united in a single file. There is no restriction on the actual digital format for the waveforms (at the moment HL7 aECG and DICOM ECG were considered). In its actual implementation, PDF-ECG is slightly more sophisticated: more importantly, the PDF file is built so that it is possible to validate that the digital data and the actual graphical report do correspond. This is an open problem in hybrid archiving and PDF-ECG was designed to cope with it. Finally, PDF-ECG supports digital signatures to guarantee integrity and for liability issues.

Table 1
Informal survey on ECG waveform storage in hospital's EHR.

Country	Question			
	Q_1	Q_2	Q_{2a}	Q_{2b}
France	No	No	–	–
France	No	No	–	–
France	No	No	–	–
France	No	No	–	–
Italy	No	No	–	–
Italy	Yes	No	–	–
Italy	Yes	No	–	–
Italy	Yes	No	–	–
Italy	No	Yes	No	–
France	Yes	Yes	No	–
Italy	Yes	Yes	No	–
Italy	Yes	Yes	No	–
Italy	Yes	Yes	No	–
USA	Yes	Yes	No	–
Italy	Yes	Yes	Yes	No
Italy	Yes	Yes	Yes	No
USA	Yes	Yes	Yes	No
USA	Yes	Yes	Yes	No
USA	Yes	Yes	Yes	No

The four questions were: (Q_1) Do you use an internal ECG management system? (Q_2) In your hospital, is there an EHR system in place? (Q_{2a}) Does it include the ECG? (Q_{2b}) Does the archived ECG include the source data?

PDF-ECG is currently undergoing active development through the activities of a dedicated working group [3] that includes representatives from manufacturers and academia, which is working toward a consensus document. The members of the working group are: Roberto Sassi (University of Milan, Italy), Jean-Philippe Couderc (University of Rochester, USA), Mathias Baumert (University of Adelaide, Australia), Raymond Bond (Ulster University, UK), Fabio Badilini (AMPS-LLC, USA), Brian Young (GE Healthcare, USA), Richard Gregg (Philips Healthcare, USA), Barry Brown (Mortara Instruments, USA), Ramun Schmid (Schiller AG, Switzerland).

While a first technical implementation of PDF-ECG was tested using the waveforms of 93 diagnostic ECG recordings, previously collected with different devices [1], a clinical test was never performed. In this work, we will report the first implementation of PDF-ECG in a clinical setting. The main objectives of the proof-of-concept were:

- to test PDF-ECG in the clinical routine environment and understand potential limitations and/or implementations bottlenecks;
- to store PDF-ECG files in the EHR of the hospital and verify the impact on the usual workflow;
- to verify the feasibility to retrieve and to re-analyze the digital ECG source data from the EHR at a later time.

To achieve the aims of the study, several software tools meant to produce and validate PDF-ECG files were developed. They are now made publicly available and quickly described in the paper.

Material and methods

Clinical proof-of-concept

In order to demonstrate the potential advantages and practical usage of PDF-ECG, a real-world proof-of-concept test was conducted with the collaboration of Fondazione Poliambulanza, a mid-size excellence hospital located in Brescia, in the northern part of Italy. The hospital was accredited by the Joint Commission International (JCI) in 2013, has 600 beds, 1850 employees and 63,000 ECGs are performed every year (2016 data; figures provided by Fondazione Poliambulanza). As with many healthcare facilities (in Italy and in the rest of the world) ECG source data is stored in a common ECG management system (Muse by GE Healthcare) but not in the EHR of the patient, the Galileo system by NoemaLife S.p.A. (recently acquired by Dedalus S.p.A), which only saves a standard PDF report.

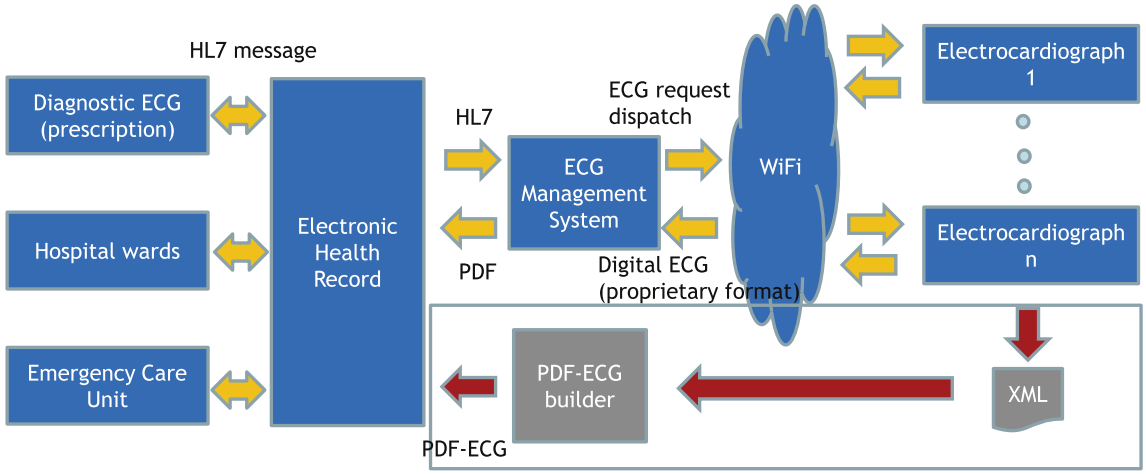

Fig. 1. EHR workflow of the Fondazione Poliambulanza hospital. ECG requests are sent from many hospital units to the EHR, which relays them to the ECG management system. The requests are further dispatched to any electrocardiograph in the hospital. Then, after a recording is actually performed on the patient for whom it was requested, the digital waveforms are sent back, in a proprietary format, to the Muse ECG management system where a cardiologist reads the ECG, a PDF report is generated and sent back to the EHR. With the support of GE Healthcare, digital data were instead collected from the electrocardiograph and, once the PDF-ECG report was produced with a web service, it was sent to the EHR in place of the usual PDF report.

A small set of healthy volunteers (six) was recruited for the test. They were informed about the goals of the experiment and undertook a basic ECG using the hospital routine equipment, *e.g.*, Mac5500 electrocardiographs (GE Healthcare) directly connected with the Muse system. Even though the test ECGs were not part of the routine clinical procedure, the standard workflow used by the hospital was strictly followed, as sketched in Fig. 1. In particular, ECG requests were queued from the EHR to the ECG management system and further dispatched to the electrocardiographs in the hospital. After acquisition on the requested subject, the ECGs digital waveforms were automatically sent back to the Muse ECG management system (using GE Healthcare proprietary format), where they were read (diagnosed) by the hospital cardiologist. A standard (no source data) PDF report was finally sent back to the EHR. As a separate and parallel process, the digital ECG data and the associated diagnoses, were also collected (directly from Mac5500 electrocardiograph with the support of GE Healthcare), stored in a public XML format, and converted to PDF-ECG using specific web services (see Section "Public software for the formalization of PDF-ECG"). Finally, the (standard) PDF previously stored in the EHR system was replaced by the new and waveform-inclusive PDF-ECG.

Public software for the formalization of PDF-ECG

From the very early phase of the Working Group in 2014, it was soon realized and agreed that in order to favor its implementation, PDF-ECG also needed the development of public domain software tools to parallel the advancements of its specifications, a process that has been evolved over the last two years.

Existing code permits today to create and validate PDF-ECGs file starting from an HL7 aECG or a DICOM ECG digital files. The code was originally developed as a stand-alone executable at the University of Milan [4] and subsequently was improved and extended by AMPS-LLC (New York, USA). In order to reach out a larger audience and to allow adequate testing and feedbacks of the new technology, the natural evolution was to design a web application. Recently, AMPS-LLC provided a cloud infrastructure and several web services while the web interface was prepared at Ulster University. The web application (see Fig. 2) is now

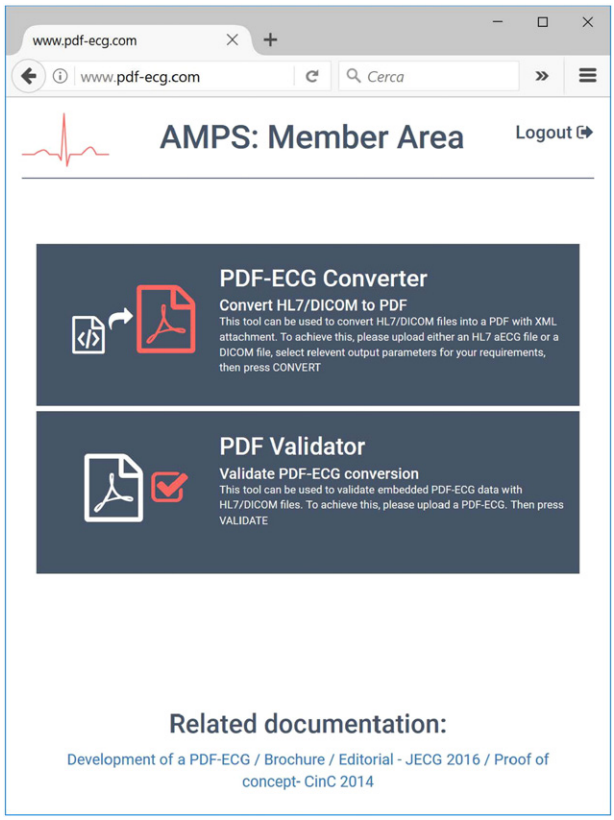

Fig. 2. A screenshot of the web application which connects to the web service to create and validate PDF-ECG files.

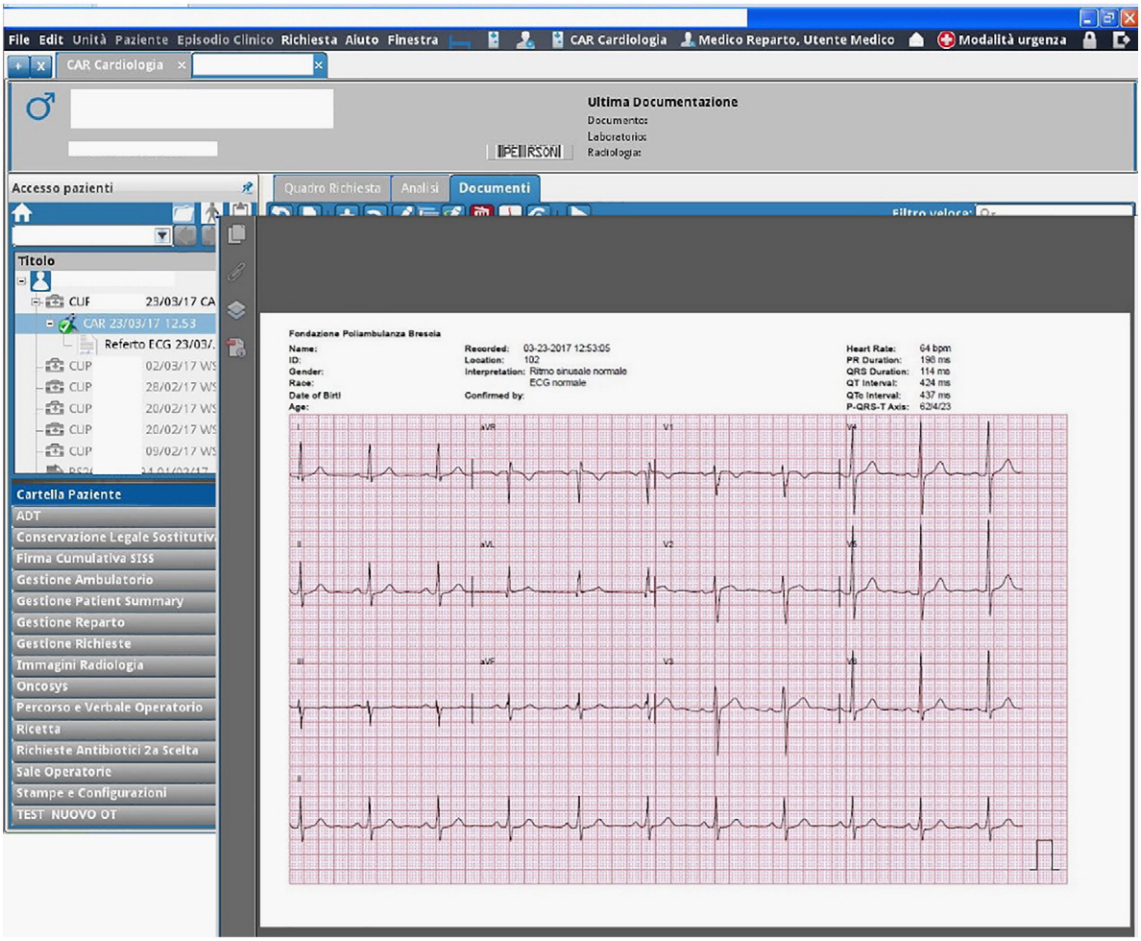

Fig. 3. A snapshot of the NoemaLife S.p.A. EHR system Galileo, with a display of the PDF-ECG file. As expected, it is undistinguishable from a normal PDF report. Only when one looks for the attachments the digital data appear.

available at pdf-ecg.com and will be continuously improved and extended over time. In fact, since the tools are used to implement PDF-ECG, the software development needs to progress over time with further enhancements (*e.g.*, the possibility to apply a digital signature, a probabilistic verification phase to cope with the fact that ECG waveforms might be filtered before printing, *etc.*). At the moment, access to the web application is granted on request to interested researchers.

Results and discussion

Six diagnostic 12-leads ECGs (from 6 different volunteers) were acquired and successfully stored in the EHR of Fondazione Poliambulanza in PDF-ECG format. The overall procedure had minimal impact on the hospital workflow. In fact, to host the PDF-ECG files, no changes were necessary on the EHR software itself. PDF-ECG naturally took the place of the usual standard PDF report (see Fig. 3). Also the ECG requests followed the usual route, as scheduled in the hospital, and the hospital's electrocardiograph devices were employed.

The average size of a PDF-ECG file was 166.7 kB. The addition of the waveforms contributed on average with 35.5 kB, which correspond to a 27.1% file size increase.

Then, later in time, PDF-ECG files were retrieved from the Fondazione Poliambulanza EHR. This step required the direct intervention of the IT division. In fact, for privacy related reason, the Galileo EHR does not permit to locally download any file (they can only be displayed on screen). The restriction could be easily overcome in the future by adding an anonymization layer which strips out the demographic information from PDF-ECG file (this information will be in a specific PDF layer that is, easily removed). The digital data were then successfully extracted (HL7 aECG format) from the PDF-ECG files and re-analyzed, several days from the first analysis from the cardiologist. In particular, we easily computed the J-Tpeak interval (using CalECG, AMPS-LLC, NY, USA [5]) and the V-index ([6], an index of spatial heterogeneity of repolarization).

Limitations of the study

The study was performed in an hospital with a recently deployed IT infrastructure and high-end electrocardiograph devices (of a single vendor). It is not clear if the impact on the hospital workflow would have been larger in a different context (and with machine without direct export of the raw ECG waveforms). Further tests should be implemented in different hospitals to verify the issue.

Conclusions

The execution of the proof-of-concept test demonstrated that adding PDF-ECG to an existing EHR has minimal impact on the hospital workflow. Indeed, no modification was required on the other side, given that, instead of a standard PDF file, a PDF-ECG file was stored for archival. The extra storage space required was minimal (of the order of a few kilobytes for each ECG). However, as the technology for general usage will be developed, storage requirements should be carefully assessed based on the volume of ECG tests performed.

Finally, a web service application permitting the creation and the validation of PDF-ECG files starting from HL7 aECG or DICOM ECG waveforms, was made available. The software tools will progressively implement the specifications defined by the Working Group.

Acknowledgments

Authors are thankful to GE Healthcare for the technical support with the conduction of the hospital test.

RS acknowledges interesting conversations with Pierluigi Ligabò, MD, on the possible advantages (for family doctors and general practitioners) of having digital ECG waveforms in PDF-ECG within patients' national EHRs.

References

[1] Sassi R, Sparagino L, Stockbridge NL, Guadiana JM, Badilini F. Proof of concept for an international long-time preservation ECG format. Computing in cardiology 2014, vol. 41; 2014. p. 461–4.

[2] Bond RR, Finlay DD, Nugent CD, Moore G. A review of ECG storage formats. Med Inform 2011;80(10):681–97.

[3] Badilini F, Sassi R. Development of PDF-ECG: further steps towards the long-term preservation of clinical electrocardiograms. J Electrocardiol 2016;49(5):753–4.

[4] Sparagino L. Proof of concept for an international long-time preservation ECG format. (Master thesis)Università degli Studi di Milano; 2014.

[5] Badilini F, Sarapa N. Implications of methodological differences in digital electrocardiogram interval measurement. J Electrocardiol 2006;39(4 Suppl):S152–6.

[6] Sassi R, Mainardi L. An estimate of the dispersion of repolarization times based on a biophysical model of the ECG. IEEE Trans Biomed Eng 2011;58(12):3396–405.