AMPS-QT

The AMPS Quarterly - Issue n.23 - 3Q2014 Editors: Fabio Badilini PhD, FACC and Martino Vaglio MS

AMPS-QT is a quarterly journal dedicated to all the people and organizations involved in the world of cardiac safety. Published by AMPS LLC, it covers all aspects of methodology and software technology related to clinical trials and Thorough QT studies.

Editorial

Right when we thought we were done, at least for a while, as far as ECG standards were concerned, we received a mail from Juan Guadiana of the Federal Aviation Administration, last January, about the need of the FAA to acquire a tool capable of converting both HL7 xml, and commercial ECG formats into PDF files, something that actually none of our customers ever asked before, except for report-printing purposes. We could immediately see the advantages of this approach: PDF viewers are ubiquitous and free of charge. We also believed we saw an apparently obvious and immediate drawback: the loss of information in the conversion process. After few mail exchanges with Juan and some due diligence on the PDF format capabilities (the new PDF/A-3u standard) Dr. Roberto Sassi, of the University of Milan, and Dr. Fabio Badilini, AMPS Chief Scientist, came quickly to the realization that a great opportunity was at hand to combine fairly elegantly both worlds (digital ECG and PDF format) thus allowing a new method to both display, with good accuracy, and store, without losing any information, the digital ECGs in an already existing, and widely used, standard format. Consulted about the concept the FDA immediately recognized the value and encouraged AMPS to pursue the idea. A proof of concept was needed and the collaboration with Dr Roberto Sassi, the FAA and the FDA culminated in the work that was recently accepted and very well received at the Computing in Cardiology Conference in Boston last September. In this issue of AMPS-QT, our guest author Dr Sassi provides our readers with a summary of the work done to reach this proof of concept. As usual, enjoy!

A Noteworthy Contribution:

Towards an International Long-Time Preservation ECG Format.

Roberto Sassi PhD, University of Milan, Milan, Italy

Digital ECG is recognized and recommended as a suitable way to store and preserve records; however, and despite a number of standards being available, including the annotated ECG standard (aECG) adopted more than 10 years ago by the FDA [1,2], ECG paper reports remain a preferred and frequent choice of clinicians.

ECG systems by most manufacturers provide the possibility of exporting traditional 12-lead printouts in common formats, *e.g.,* ISO 32000-1:2008 Portable Document Format (PDF). These printouts permit some qualitative interpretation; however, most often only a portion of the data is included and, even worse, any analysis from the graphical picture is limited by the printing resolution of the system. On the contrary, raw digital ECG data would facilitate subsequent quantitative assessments, with a clear advantage in diagnosis.

As of today, ECG records are thus either stored as paper or PDF reports, or in one of the digital standards available [1] (being aECG only one of the possible options). Thus they cannot be easily interchanged in the (frequent) case that clinical or research analyses may require it.

In the proof of concept we recently presented at the Computers in Cardiology meeting held in Boston, we proposed a new standard (PDF-ECG) that combines the benefits of a digital ECG and a standard graphical report. This proof of concept is a PDF/A-3u (Unicode) envelope containing the ECG graphic (with no predefined layout), the ECG digital data in aECG format, and a digital signature. A means to ensure content match between the image and the digital data is also provided. PDF-ECG satisfies, within a unique structure, both the needs for a simple graphic report, accessible without the installation of

specific proprietary software, and for delivering all the acquired information for further processing with specialized software.

The HL7 aECG standard

The HL7 aECG standard was created in response to the FDA's digital electrocardiogram initiative introduced in 2001 [2]. The FDA required both ECG waveforms and annotations, generated for regulatory submission, to meet a specific format. At the time, no current standard for ECG waveforms met all the FDA's needs. As a result, the FDA, together with representatives from pharmaceutical industry, central core laboratories, and device manufactures worked together within Health Level Seven (HL7), a standards developing organization, to create a format to meet the FDA requirements. Like other HL7 v3 formats, aECG is XML based, which means that the implementation of an aECG object must follow a specific XML schema, that defines in details the rules for a given file to be valid. These rules include the specific names of the XML mark-ups (tags, or elements), as well as which tags are mandatory and which are optional, according to an official Implementation Guide document [3].

Even if it was initially developed to meet specific needs of US regulatory bodies, and also due to available software [4], aECG has been designed to be a broad multi-purpose ECG standard and results to be an ideal platform for data interchange.

The PDF/A-3u standard

PDF is a digital format, originally developed by Adobe Systems Inc., for representing printed documents. It was derived from a subset of the PostScript language, and while Adobe holds patents to PDF, it subsequently made the specifications publicly available. As of today, PDF files are used to represent a large body of information around the world, and while this information must be preserved for a long time, multiple generations of technology might hinder accessibility. For these reasons, a second ISO standard (19005-1:2005) defines a file format based on PDF, known as PDF/A-1, which is meant for long term preservation of information, independently of the tools employed for creating or viewing the document. In practice, PDF/A-1 restricts the number of PDF components and features which might be used. Moreover, PDF/A-1 provides standard metadata to represent context, history and logical structure of the information within the document.

The latest recent revision (ISO 19005-3:2012), PDF/A-3, permits the embedding of files of any format (including XML, CSV, CAD, etc.).

The PDF-ECG proof-of-concept

PDF-ECG was conceived as an *hybrid archiving format*, where a source document in a less preservation-robust format is embedded within a PDF/A file. More specifically, PDF/A-3u (Level U conformance) was selected to be Unicode compliant.

The main content of a PDF-ECG file is a standard ECG report, such as (but not limited to) a 3×4 or 6×2 leads printout. No constraints are set on the graphic report, (except a few technical implementation details described below) and any manufacturer can shape it to its needs. The file can be opened with any standard PDF viewer, thus hiding the further complexity to users who might not need to access the digital data (*e.g.,* patients).

The aECG file is embedded into the PDF/A file as a compressed object, reducing significantly its footprint. For hybrid archiving, PDF/A-3 mandates that an explicit association must be made between each associated embedded file and the PDF container (or one of its objects) by means of the AFRelationship key. Typically, such association is set to Data (meaning that the data can be employed to rebuild the PDF content) or Alternative (meaning that the two representations are equivalent), as we did for PDF-ECG. While the aECG data can be extracted using standard software by any user, its main purpose is to be employed by automated specialized algorithms or clinical tools used by clinicians to perform further analyses.

Unfortunately, there is no standard way to validate the relationship between the main content and any embedded files in PDF/A-3. Therefore, even if the two representations were declared as Alternative, it would be possible, by mistake or deliberately, to produce a PDF-ECG document where the aECG data did not correspond to what displayed in the PDF file. To avoid such possibility we designed a two-fold verification scheme. First, the institution or entity creating the PDF-ECG file has the option to sign the document digitally, to avoid subsequent modifications. Since the signature verification procedure is embedded in most viewers, its validation does not require additional software. Thus, the digital signature puts the correctness of the stated AFRelationship under the responsibility of the signer. However, there are situations in which this is not possible or inappropriate. Also, a digital signature does not protect from non-deliberated mistakes and technical problems. For this reason, we put a significant effort in constructing PDF-ECG such that it may be validated after its creation by a third party. To obtain this, we set forth the following (still preliminary) specifications for the PDF files:

- Report pages must contain a main layer. Its name must be "LAYOUT_[hs]:[vs]" where hs and vs are the horizontal (in mm/s) and vertical (in mm/mV) scales, respectively, *e.g.,* LAYOUT_25:10 represents the classical 25 mm/s and 10 mm/mV ECG scales.
- For any aECG <sequenceSet>, a new layer is added (sibling of the main layout).
- For any HL7 signal, a further layer is created (child of the sequenceSet layer in which it is contained). The name of the layer describes its content according to the scheme:
  "[CODE](counter*)_[s]:[i]:[e]:[offset]"
  where CODE is the HL7 name of the signal, s and e are the index of the first and last samples printed and i is the increment (in samples) employed. In fact, most often the ECG report only contains a portion of the information in the digital file. Also, offset is the vertical offset on the page, in postscript points. counter* (optional) is employed to enumerate signals printed more than once (e.g., in the 3×4 format, MDC_ECG_LEAD_II(2)_0:1:2600:325)
- Every continuous ECG waveform, plotted on the report, must be contained in a single independent PDF stream contained in its layer.

At validation, after deflating the compressed PDF streams, the position of the points (composing a ECG waveform on the graphic report) are extracted. The layer naming convention facilitates identifying each lead. Then, the vertical offset (offset) is removed and the positions of the points are rescaled to mV by using the scale extracted from the name of the main layer, e.g.

$$\mathbf{mVPoints_i} = \frac{\mathbf{25.4 \times (pdfPoints_i - offset)}}{\mathbf{72 \times 10}}$$

Finally, the value of each $mVPoints_i$ is compared with the corresponding sample in the embedded aECG file. PDF 1.7 uses single precision floats or 32-bit fixed point numbers, so the comparison in a genuine file will report a small difference $\ll 1$ μV, due to rounding errors (unless using the same identical numerical library, which is not practical). However, when the aECG embedded file and the PDF report differ, the difference is significantly larger than a few μV.

No validation is implemented at the moment for non-waveform data (e.g., name of the patients, age, etc.), although it could be easily implemented in the future.

Feasibility of the proof of concept was assessed on 93 digital ECGs randomly selected from a few ECG clinical studies previously submitted to the ECG Warehouse for FDA review. The selection was purposely heterogeneous and included ECG records from different manufacturers with different sampling rates and resolutions. All ECG files were first checked to be valid aECG objects, using a commercial software package [5]. Then, each digital file was rendered in a vector standard report (either 3×4 or 6×2). A PDF/A-3u file encapsulating the digital data was prepared along the lines of what previously described. The conformance to the PDF/A-3u format of the files was checked with the commercial Adobe Acrobat preflight tool.

A dedicated software application was finally developed and employed to validate that the graphic report in the PDF-ECG and the embedded HL7 file corresponded. The corresponding maximum validation discrepancy was always smaller than 0.02 μV (28 times smaller than the finest resolution among the 93 files).

Conclusions

The new PDF-ECG document can be redistributed and opened with most available PDF readers, without specific software. Moreover, analysis software could be easily adapted to accept the file as input for subsequent analyses, which could be indeed overlaid on the printout itself. PDF-ECG long-term preservation characteristics make it suitable for long-term archival in hospital patients managements systems, core laboratory and research center databases. Finally, the PDF-ECG formats permits third-party validation, solving a problem which might prevent the adoption of this form of hybrid archiving.

The PDF-ECG format is still in a very preliminary stage, but we hope that manufacturers and institutional bodies might be interested in setting up a working group for its further formalization

References

[1] Bond RR, Finlay DD, Nugent CD, Moore G. A review of ECG storage formats. Int J Med Inf 2011;80:681–697.

[2] Brown BD. The FDA's Digital ECG Initiative and Its Impact on Clinical Trials. In: Morganroth J, Gussak I, editors. Cardiac Safety of Noncardiac Drugs, Humana Press; 2005, p. 301–327.

[3] Brown BD, Badilini F. HL7 Version 3 Implementation Guide: Regulated Studies – Annotated ECG, Release 1. Health Level Seven International; 2005.

[4] Badilini F, Isola L. Freeware ECG viewer for the XML FDA format. Proceedings of the 2nd OpenECG Workshop, Berlin, Germany: 2004, p. 31–34.

[5] Vaglio M, Isola L, Gates G, Badilini F. Use of ECG quality metrics in clinical trials. Comput Cardiol 2010;37:505–508.

Products News

Latest Releases

In Q3 2014 we have released:

- Version 1.2.0 of AMPS CER-S (Continuous ECG Recordings Suite) including the new platform for Holter format conversion.

Looking forward

In Q4 of 2014 AMPS is planning to release:

- The AMPS beat-to-beat continuous ECG solution

While, in early 2015, we are planning to release:

- A new major version of CER-S including the following platforms:
  - Continuous ECG beat detection and classification
  - ECG beat editor
  - Arrhythmia detection and Arrhythmia editor

AMPS Notebook

Fabio Badilini and Gianfranco Toninelli actively participated at the **Computing in Cardiology Conference** that was held in Cambridge, MA from September 7th to 10th, 2014. Here three abstracts were presented:

- Gianfranco Toninelli, coauthored by AMPS team and the University of Brescia, talked about "ECG Quality Assessment Using Data Mining",
- Roberto Sassi, from the University of Milan, introduced PDF-ECG, fully presented in this issue of AMPS-QT.
- Dave Mortara enlightened a paper, coauthored by Fabio Badilini, entitled "A Quantitative QT Hysteresis Model".

Fabio Badilini will be attending the **American Heart Association**, Scientific Session held from November 15th to 19th in Chicago, Illinois.
He will also be attending **CSRC Annual Meeting**, held in Washington DC on December 11th and 12th.