International Journal of E-Health and Medical Communications
Volume 10 • Issue 1 • January-March 2019

DICOM Metadata Analysis for Population Studies

Milton Rodrigues dos Santos, Health Sciences School, IEETA, University of Aveiro, Aveiro, Portugal

Luis Bastião Silva, BMD Software, Lda, Aveiro, Portugal

Augusto Silva, Department of Electronics, Telecommunications and Informatics, IEETA, University of Aveiro, Aveiro, Portugal

Nelson Pacheco Rocha, Department of Medical Sciences, IEETA, University of Aveiro, Aveiro, Portugal

ABSTRACT

This article reports an experimental study to determine how to use the stored Digital Imaging and Communication in Medicine (DICOM) metadata to perform population studies. As a case study, it was considered three types of medical imaging studies (i.e. routine head computed tomography, thorax computed radiography and thorax digital radiography) stored in the picture archiving and communication systems (PACS) of three healthcare institutions. The final sample consisted of DICOM metadata belonging to 1370360 images, corresponding to 109160 medical imaging studies performed on 72716 patients. The study followed a methodological approach that allows the identification of the number of patients with performed studies by age group and gender, as well as the average number of studies by patient, age group and gender in each one of the three healthcare institutions. The results show the relevance of the aggregation and analyses of DICOM metadata stored in heterogeneous PACS facilities.

KEYWORDS

Big Data, Data aggregation, DICOM Metadata, Medical Imaging, PACS, Population Studies

INTRODUCTION

The enormous amount of information that is acquired, stored and managed within heterogeneous healthcare environments, supported by complex healthcare information networks, is increasingly used to continuous improvement of healthcare provision (Asche, Seal, Kahler, Oehrlein, & Baumgartner, 2017).

The constant technological evolution has given rise to automatic tools to access, process and analyse clinical information. In particular, concerning medical imaging, these tools allow remote and safe analyses of clinical information related to millions of people, comparisons of clinical effectiveness, efficiency, and efficacy, dissemination of good practices, and, therefore, the improvement of the quality of healthcare provision (DeVoe et al., 2011; Klaiman, Pracilio, Kimberly, Cecil, & Legnini, 2013). In this context, advanced processing methods (e.g. artificial neural networks) might optimise the analysis of large amount of healthcare data (Moreira, Rodrigues, Kumar, Al-Muhtadi, & Korotaev, 2017), which might be useful to promote the quality of healthcare provision (Gholipour, Rahim, Fakhree, & Ziapour, 2015).

The population aging and, consequently, the increasing prevalence of chronic conditions, among other factors (Van Oostrom et al., 2016), including technological developments, defensive medicine,

DOI: 10.4018/IJEHMC.2019010101

and the need to perform a large number of medical imaging studies and to achieve results in short periods (Amis Jr. et al., 2007), requires the improvement of the medical imaging procedures, which is reinforced by the fact that many of these procedures involve the use of ionizing radiation (Thrall, 2009; Sachs & Long, 2016).

In addition to the optimization of exposure factors to increase the performance of the medical imaging practice, another aspect that should be considered is the patients' exposure to X-radiation, namely the number of medical imaging studies conducted during their lifetime (Fabritius et al., 2016). Therefore, initiatives should be taken to identify the causes of duplicated and repeated radiation exposures within medical imaging studies (Nol, Isouard, & Mirecki, 2006; Prieto et al., 2009; Hendee et al., 2010; Yuan et al., 2013), namely taking advantage of the information that is part of the Digital Imaging and Communication in Medicine (DICOM) metadata (Tsalafoutas & Metallidis, 2011; Brady & Kaufman, 2015).

The DICOM metadata data produced by different medical imaging modalities and distinct equipment manufacturers can be used in multiple scenarios, namely in efficiency metrics for imaging device productivity (Hu et al., 2011), or to provide data for different initiatives such as, for example, development of medical imaging performance indicators (Prieto et al., 2009); Jahnen, Kohler, Hermen, Tack, & Back, 2011; Dave & Gingold, 2013) or biomedical research (Freymann, Kirby, Perry, Clunie, & Jaffe, 2012).

Important parameters such as image processing and image acquisition parameters, patient dose and geometric information are stored in the Picture Archiving and Communication Systems (PACS). Despite these data being stored on the healthcare repositories, the traditional information systems are not able to access them. These difficulties can be overcome using the Dicoogle (Valente, Silva, Godinho, & Costa, 2016). Dicoogle is an open source solution that allows extracting textual information from the PACS and performing flexible queries over DICOM metadata. The flexibility to index different Information Object Definition (IOD) (i.e. DICOM objects types with different data elements) is due to the replacement the traditional database by an indexation engine. Following this approach, it is possible to index all existent DICOM data elements (text-based) without need to create new fields, new tables, and new relations that would be necessary for a non-proprietary relational database approach (Santos et al., 2015).

Within this context, the study reported in this article was informed by the following research question: is it possible to perform population studies by accessing, processing and analysing the DICOM metadata stored in the PACS of different healthcare institutions to track individual radiation exposure history and to objectively support healthcare management indicators?

Therefore, the study report in this paper aimed to propose and evaluate a method based on a public domain software, the Dicoogle (Valente et al., 2016), able to manage DICOM metadata stored in different PACS of various healthcare institutions, namely to support continuous improvement of clinical procedures.

Concerning the experimental work, the authors access, process and analyse DICOM metadata stored in the PACS of three healthcare institutions related to three imaging modalities (i.e. Computed Tomography - CT, Computed Radiography - CR, and Digital Radiography - DX) and, more specifically, related to three types of medical imaging studies: routine head CT, thorax CR and thorax DX.

This article is organized as follows: The following section, Background, addresses relevant factors that influence the availability of required information to characterize medical imaging provision, especially when various healthcare institutions are involved. The section Methods and Materials presents the various phases of the experimental work. The section Results and the section Discussion are devoted, respectively, to results presentation and discussion. Finally, the section Conclusion and Future Work gathers the conclusion and some prospects for future work.

Related Content

Computerization of Primary Care in the United States
James G. Andersonand E. Andrew Balas (2006). *International Journal of Healthcare Information Systems and Informatics (pp. 1-23).*
www.irma-international.org/article/computerization-primary-care-united-states/2185