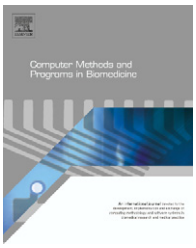
COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 105 (2012) 22–30

journal homepage: www.intl.elsevierhealth.com/journals/cmpb



Development of a clinical data warehouse from an intensive care clinical information system

Marleen de Mul[a,*], *Peter Alons*[b], *Peter van der Velde*[c], *Ilse Konings*[c], *Jan Bakker*[d], *Jan Hazelzet*[e]

[a] *Erasmus University, Institute of Health Policy and Management, PO Box 1738, 3000 DR Rotterdam, The Netherlands*
[b] *Atos Origin Nederland, Enterprise Solutions/BI-CRM, Papendorpseweg 93, 3528 BJ Utrecht, The Netherlands*
[c] *Erasmus University Medical Center, Department of Information Technology, PO Box 2040, 3000 CA Rotterdam, The Netherlands*
[d] *Erasmus University Medical Center, Department of Intensive Care, PO Box 2040, 3000 CA Rotterdam, The Netherlands*
[e] *Erasmus University Medical Center, Department of Pediatric Intensive Care/Department of Information Technology, PO Box 2040, 2000 CA Rotterdam, The Netherlands*

ARTICLE INFO

Article history:
Received 31 July 2009
Received in revised form
18 January 2010
Accepted 14 July 2010

Keywords:
Data warehouse
Intensive care
Medical records systems, computerized
Clinical information system

ABSTRACT

There are relatively few institutions that have developed clinical data warehouses, containing patient data from the point of care. Because of the various care practices, data types and definitions, and the perceived incompleteness of clinical information systems, the development of a clinical data warehouse is a challenge.

In order to deal with managerial and clinical information needs, as well as educational and research aims that are important in the setting of a university hospital, Erasmus Medical Center Rotterdam, The Netherlands, developed a data warehouse incrementally. In this paper we report on the in-house development of an integral part of the data warehouse specifically for the intensive care units (ICU-DWH). It was modeled using Atos Origin Metadata Frame method. The paper describes the methodology, the development process and the content of the ICU-DWH, and discusses the need for (clinical) data warehouses in intensive care.

1. Introduction

Information technology in health care is still a topical subject, and reports like IOM's *Crossing the Quality Chasm* have stimulated developments in physician order entry, decision support systems and shared patient records [1]. Despite all the efforts, many health care organizations still have stand-alone information systems that do not communicate with each other. Therefore it remains an enormous task to use data collected at the point of care or in the supporting administrative processes for other purposes than they were collected for (e.g. management information, quality assessment and research). More importantly, clinical information systems (CIS) such as electronic patient records, are often designed to support hands-on care for individual patients, but are not well suited for analyses on an aggregated level, for example on groups of patients with the same disease.

Data warehousing is one of the techniques that seems promising for healthcare information systems. In its simplest definition, a data warehouse is a copy of transaction data specifically structured and optimized for query and analysis [2]. Use of data warehouses in healthcare is not new. The first articles on data warehousing date from two decades ago. In

* *Corresponding author*. Tel.: +31 10 40 81 86 8; fax: +31 10 40 89 09 2.
E-mail addresses: demul@bmg.eur.nl (M. de Mul), peter.alons@atosorigin.com (P. Alons), J.vandervelde@erasmusmc.nl (P. van der Velde), i.l.m.konings@erasmusmc.nl (I. Konings), Jan.bakker@erasmusmc.nl (J. Bakker), J.a.hazelzet@erasmusmc.nl (J. Hazelzet).
0169-2607/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2010.07.002

the period 1995–2000, it was a topic in healthcare management journals, but many articles provided only viewpoints, and lacked technical and empirical data [e.g. 3,4]. Certainly back then, the focus was on financial data warehouses, and data warehouses are still used for research into in-hospital costs [e.g. 5–7]. Since then, the focus has shifted somewhat to data warehouses for the bioinformatics domain [e.g. 8]. To date, there are but a few published examples of *clinical* data warehouses, using data from CIS, that are implemented and in use [9–13]. Because of the various care practices, data types and definitions as well as perceived incompleteness of clinical information systems, the development of a clinical data warehouse is a challenge [14].

Intensive care units (ICUs) might benefit from clinical data warehouses, as they can be characterized as information-rich environments with a high degree of automation and information technology. The last decades many ICUs especially in the larger university and teaching hospitals have implemented clinical information systems (CIS), sometimes called patient data management systems (PDMS) in critical care. These systems sample and store data from the monitors and other bedside devices, as well as manually entered observations. CIS are used for charting, fluid balance, medication lists and care planning. Some CIS also have physician order entry functionalities, or connections to the hospital information system (e.g. the laboratory system). Since these systems are designed for the point of care, the aggregation of data for management and research is limited. Querying the database of a CIS requires technical skills and knowledge of the database structure, competences which most managers, doctors, and nurses lack. Moreover, querying can be a burden for the operational database because it slows down performance [14].

There is still little evidence in the literature of the presence and use of data warehouses containing ICU data. Brammen et al. [15] provided the only example of a data warehouse that uses ICU data as a source for research on the interface of intensive care and genetics. The paper does not delve into the technical details of the development, but it is clear that this data warehouse has a specific and narrow focus that fits current research interests but is not prepared for future needs. Other studies on data warehousing in intensive care use a combination of a data warehouse (containing only administrative or financial data) and a separate query on the ICU's CIS [e.g. 16–19]. From this we conclude that data warehousing in intensive care is emerging, but not yet documented very well. Moreover, there is little experience with data warehouses containing clinical data about ICU patients, to be used for quality monitoring and research.

In this paper we report on the in-house development of an ICU-part of the data warehouse at Erasmus Medical Center, Rotterdam, The Netherlands (ICU-DWH). This part of the data warehouse collects data from the CIS installed at the ICUs, as well as data from the hospital information system. It was modeled using Atos Origin Metadata Frame method. The paper starts with a short analysis of the characteristics of intensive care, to understand the information needs of ICU staff and management. After that we describe the Metadata Frame method and its application in this project. We also present some examples of the way the ICUs currently use the data warehouse. In Section 5.1 we discuss the advantages of the Metadata Frame method, and present our lessons learned.

2. Background

Every clinical department is, or at least should be, interested in quality information and other issues requiring the analysis of patient data. However, intensive care units have some characteristics that make clinical data warehouses even more interesting and relevant. A first characteristic is the high availability of and dependency on technology and data. The devices on the ICU support vital functions of patients, while producing a lot of information about the patient: heart frequency, blood pressure, lab results, fluid intake, medication, etc. Producing a 'snapshot' of the status of an individual patient requires expert knowledge, but generating valuable information about groups of patients is even more complex. Many data elements have to be combined, which is difficult in a CIS.

Secondly, the cost of and access to intensive care are important (management) issues. The ICU is the most expensive ward in a hospital, ICU beds are scarce, and should be allocated to those patients who benefit from an ICU admission. Both to make a good admission and discharge policy and to monitor its application, data about the patient *population* of the ICU is crucial, yet not available in a CIS.

Related to cost and scarcity, a third characteristic of intensive care is the obligation to be accountable for the quality of care. In many countries, for example, standardized mortality rates (SMR) and other quality measures are benchmarked between ICUs and reported to health inspectorates. Many quality measures and prediction models have been developed for ICUs, and because of the drive to measure and improve quality ICUs are continuously in need of performance data. Quality measures such as "number of ventilator associated pneumonia infections per 1000 ventilator days" can be derived from the data in a CIS, but this requires a complex query. Likewise, severity of illness models (e.g. SOFA[1] [20]) or mortality predication models (e.g. APACHE[2] IV [21]) require complex queries if they are not available in a CIS.

A fourth and final characteristic that is relevant for data warehouse developments is the emergence of intensive care medicine as an independent discipline. Although the first ICUs date from the 1940s, it was not until the 21st century that intensive care medicine was regarded as a medical specialism. Knowledge on the best treatment of critically ill patients is still growing, and CIS can be a powerful data source for scientific research, especially for university hospitals.

Because of all these characteristics of intensive care, the need to analyze ICU data from different perspectives and on different levels (individual patient, patient group, department) is self-evident. Therefore a data warehouse combined with easy-to-use querying and reporting tools would be a big step forward. However, the CIS used in intensive care units usually have very complex data models. There is a lot of technical knowledge needed for making the right queries, while the

[1] Sepsis-related Organ Failure Assessment.
[2] Acute Physiology And Chronic Health Evaluation.