



## Construction of quality-assured infant feeding process of care data repositories: definition and design (Part 1)



Ricardo García-de-León-Chocano<sup>a</sup>, Carlos Sáez<sup>b</sup>, Verónica Muñoz-Soler<sup>c</sup>,  
Ricardo García-de-León-González<sup>d</sup>, Juan M. García-Gómez<sup>b,e,f,\*</sup>

<sup>a</sup> Department of Information Technology, Hospital Virgen del Castillo, Yecla, Gerencia Área de Salud V-Altaplano, Servicio Murciano de Salud, Spain

<sup>b</sup> IBIME, Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas (ITACA), Universitat Politècnica de València, Spain

<sup>c</sup> Department of Nursing, Universitat d'Alacant, Spain

<sup>d</sup> Department of Paediatrics, Hospital Virgen del Castillo, Yecla, Gerencia Área de Salud V-Altaplano, Servicio Murciano de Salud, Spain

<sup>e</sup> GIBI230 (Grupo de Investigación Biomédica en Imagen), Instituto de Investigación Sanitaria (IIS) hospital la Fe, Spain

<sup>f</sup> Unidad Mixta de Investigación en TICs aplicadas a la Reingeniería de Procesos Sociosanitarios (eRPSS), Instituto de Investigación Sanitaria del Hospital Universitario y Politécnico La Fe, Bulevar Sur S/N, Valencia 46026, Spain

### ARTICLE INFO

#### Article history:

Received 6 May 2015

Accepted 25 September 2015

#### Keywords:

Data quality

Electronic health record

Data analysis

Breastfeeding

Perinatal care

### ABSTRACT

This is the first paper of a series of two regarding the construction of data quality (DQ) assured repositories for the reuse of information on infant feeding from birth until two years old. This first paper justifies the need for such repositories and describes the design of a process to construct them from Electronic Health Records (EHR). As a result, Part 1 proposes a computational process to obtain quality-assured datasets represented by a canonical structure extracted from raw data from multiple EHR. For this, 13 steps were defined to ensure the harmonization, standardization, completion, de-duplication, and consistency of the dataset content. Moreover, the quality of the input and output data for each of these steps is controlled according to eight DQ dimensions: predictive value, correctness, duplication, consistency, completeness, contextualization, temporal-stability and spatial-stability. The second paper of the series will describe the application of this computational process to construct the first quality-assured repository for the reuse of information on infant feeding in the perinatal period aimed at the monitoring of clinical activities and research.

© 2015 Elsevier Ltd. All rights reserved.

### 1. Introduction

The World Health Organization defines breastfeeding as the natural and healthiest way of feeding infants and young children. Exclusive breastfeeding is recommended for the first six months of life, while an adequate complementary feeding is recommended up to two years or older [1]. Currently, breastfeeding is considered one of the main determinants of maternal and child health. Having related to the prevention of risk factors for chronic diseases of childhood and preventable burden of disease [2]. However, European data show that less than 20% of babies are fed following these recommendations [3].

Given the significant health implications of these low rates, the European strategy for the protection, promotion and support of breastfeeding [4] proposes, among other actions, the widespread

deployment of the Baby Friendly Hospital Initiative (BFHI), using global and coordinated actions, along with the development of control mechanisms that may facilitate its implementation. Nevertheless, the absence of these control mechanisms is the main cause of the existing gap between the evidence and the daily clinical practice [5].

In the “Virgen del Castillo” Hospital, the international guidelines described above materialized through an intervention in the improvement of the quality of maternal and child care, which allowed its accreditation as a baby friendly hospital and produced significant improvements in breastfeeding rates [6]. However, the sustainability of such interventions depends on the continuous adaptation to the local context and its systematic inclusion in routine workflows [7]. Thus the improvement group design the process of care of infant feeding (PCIF) (PIEMCA08/13 project) extending the scope to primary care to match the clinical practice from birth until two years of age to the recommendations of the evidence.

For monitoring the clinical process and activating research studies on the PCIF, it was decided to use population data from the

\* Correspondence to: Informática Biomédica (IBIME), Departamento de Física Aplicada, Universidad Politécnica de Valencia, Camino de Vera, s/n, Building 8G, Valencia 46022, Spain. Tel.: +34 963877000x75278.

E-mail address: [juanmig@ibime.upv.es](mailto:juanmig@ibime.upv.es) (J.M. García-Gómez).

Electronic Health Records (EHR) through the construction of a quality-assured data repository. Hence, the aim of the present work is to design the process to construct repositories on birth and infant feeding based on population data from the EHR. The construction process has been implemented by means of 13 quality-assured procedures to assess, recover or filter clinical data to enable their reuse in the monitoring of the PCIF and to study relevant questions on infant feeding.

## 2. Background

The current deployment of health information systems and EHR is a source of valuable data for the systematic monitoring of clinical practice, both for management and research. However, unlike in other health care fields, to our knowledge there is no experience in the use of the data from EHR on infant feeding for the implementation and control of the BFHI, although evidence shows a relationship between the convenience of use population data for monitoring hospital practices and the feasible development and strengthening of BFHI, a type of study only performed by hard and isolated collections of population data [8,9].

Moreover, important forums from the health sciences field support the relevance of basing clinical research and healthcare quality improvement interventions on population studies [10–12]. In this regard, there are experiences that prove the potential of the EHR as a source for population studies when data are subjected to suitable processes to assure the quality of the records [13].

However, the establishment of EHR as a source of knowledge in clinical settings poses important challenges to health organizations. It is of great importance that the data from EHR show an accurate view of the real clinical situation of patients. In this sense, data quality (DQ) is directly related to healthcare quality [14]. Additionally, there are possible errors produced in the registration process by EHR system users [15,16], as well as systematic errors associated with the information system's structure [16].

The lack of systematic control mechanisms of DQ of EHR poses a serious obstacle to ensuring the collection of valid and reliable information for data reuse [16,17]. Multiple factors determine the lack of DQ control. Notable among them is the need for reliable conceptual frameworks that establish the definition and measurement of relevant DQ aspects, usually called DQ dimensions [18]. Moreover, there is a need for computational tools for DQ governance in healthcare and health research [17].

There is a general agreement in defining DQ in terms of fitness for use to fulfil the user's requirements [19,20] or in terms of DQ goals [21]. With such a purpose, many studies defined some DQ dimensions [17–19,22,23] that had to be measured: attributes that represent a single aspect or construct of DQ [18]. Dimensions can conform to data specifications or to user expectations [24].

Recently, Weiskopf et al. [17] performed a systematic review of DQ dimensions and methods in DQ assessment of EHR. From 27 unique terms describing dimensions in the literature, they empirically derived five high-level dimensions: completeness, correctness, concordance, plausibility and currency. In addition to these dimensions, Oliveira et al. [25] formalized a set of DQ problems along different granularity levels of organizational databases, namely multiple sources, multiple relations, single relations and attribute/tuple. Hence, dimensions can be associated with these problems and measured at the different granularity levels.

In this study, we used eight of the dimensions previously defined by the authors in their DQ framework [23] to measure the quality of the input and output data of each step of our process to construct a repository, including: predictive value, correctness, duplication, consistency, completeness, contextualization, temporal-stability, and spatial-stability. The reasons for selecting this

framework were: (1) it helps defining the requirements of the DQ project taking into account the circumstances where data were acquired, (2) it provides a complete set of DQ dimensions, in agreement with those in literature, which serve as the basis for defining what DQ aspects need to be measured and (3) this constitutes an interesting case study in which the applicability of the theoretical framework can be explored.

On the other hand, there are differences on how different EHR systems organize data locally, which can be overcome using standards. Several standards focus on interoperability, both at organizational level (e.g. CONTsys) and at semantic level (e.g. ISO13606 and open EHR). The CONTsys standard (a.k.a. EN ISO 13940 Part1 [26]) provides a system of reference concepts for continuity of care. This standard provides the terminology required to plan, implement and monitor the activities and health status from clinical processes. Besides, the standards for semantic interoperability (such as 13606 and open EHR) have demonstrated their utility in achieving semantic interoperability between systems [27,28]. However, in the mother–child area little development of clinical models following these standards was found.

### 2.1. Methodology

The PCIF improvement group promoted the construction of quality-assured data repositories from the EHR, to enable the monitoring and evaluation of the intervention. Hence, an expert group was set up by clinicians head of PCIF and information technology experts to define the specifications and requirements of the repository for monitoring the PCIF and its information model. Subsequently, this group designed the visual representation and the user interfaces to configure the construction of the repositories. Next, data extraction procedures were defined from the EHR. In our case, the PCIF data is distributed in two EHR systems: perinatal data are housed in the hospital information system (HIS) (Selene Siemens AG) and revisions of the Healthy Child Programme are included in the Primary Care Information System of the health area. A total quality management approach was adopted in the construction of the data repositories in order to avoid posterior suboptimal decisions based on inaccurate data.

In this section, first we describe the definition of the PCIF information model using concepts from the CONTsys norm. Next, we explain the DQ framework adopted to ensure the total quality of the process. Finally, we define the process for the construction of quality-assured repositories.

### 2.2. The PCIF information model

The organizational context of the PCIF information model was defined by CONTsys concepts. The CONTsys standard (CEN STANDARD EN 13940) [26] defines a set of concepts to be used in modeling continuous processes of care. This standard ensures the same meaning for the fundamental concepts governing the organization of continuity of care among heterogeneous environments. Table 1 shows the concepts from CONTsys that we used in the PCIF to define their organizational context.

The clinical concept term is used with the meaning defined by the ISO EN 13606 standard. Both standards CONTsys and ISO EN 13606 complement each other and are connected through the concept EHR extract (part or all of the EHR of a subject of care that is intended to communicate). Fig. 1 shows the PCIF information model using the CONTsys concepts linked to the data sources of our EHR.

### 2.3. Data quality framework

A total DQ management guided the design of the process to construct quality-assured repositories. Given that the process will

Download English Version:

<https://daneshyari.com/en/article/504851>

Download Persian Version:

<https://daneshyari.com/article/504851>

[Daneshyari.com](https://daneshyari.com)