

# Exchanging Nursing Oncology Care Data With use of a Clinical Data Ware House

William Goossen, Frank Boterenbrood, Irene Krediet

department ICT innovations in health care

Christelijke Hogeschool Windesheim

Zwolle, the Netherlands

e-mail: wtf.goossen@windesheim.nl

**Abstract**— Isala Clinics Hospital and Icare Homecare under leadership of Windesheim use a multi-method approach to standardize nursing data for oncology care based on Detailed Clinical Models (DCM), and use these DCM as requirements for the Electronic Health Record (EHR) system. Further, the DCM are used for extracting the data from EHRs into a Clinical Data Ware House (CDWH) that will be deployed for different data uses. Data uses include exchanging clinical data from hospital to homecare using Health Level 7 v3 Care Record message, and reuse data for quality reporting. This paper presents the research approach including the design for the infrastructure, and the results from the first phase. The paper illustrates that based on existing materials, widely available for nursing, it is feasible to determine the data needs for exchange and for various data use purposes. In particular the results of standardization and messaging are presented.

**Keywords:** oncology nursing, detail clinical models DCM, electronic health record EHR, nursing informatics, clinical data ware house, HL7 v3 Care Record.

## I. INTRODUCTION

Many development projects towards the electronic health record (EHR) state that one reason for their existence is the abundance of data for secondary uses [1, 2, 3, 4, 5]. It is assumed that the exchange of clinical data for continuity of care and the extraction of clinical data from the EHR can be taken for granted. However, that is not the case. There are several reasons why the electronic exchange and aggregation from the EHR for continuity of care or for secondary use cannot be taken for granted, but needs substantive work, including tackling privacy issues, EHR performance, data use, EHR format, and data aggregation [1, 2, 3, 4, 5, 6, 7, 8].

These distinct tasks are combined into a project for nursing care for oncology patients, currently treated in the Isala Clinics and supported at home by Icare Homecare. The research questions for this project are the following:

1. How can a CDWH for Isala Clinics be developed for basic patient data, generic nursing data and data about oncology nursing, allowing management of EHR data?
2. Can the nursing data be standardized and mapped to specifications for different data uses as electronic exchange, quality indicators and clinical studies?
3. Can we use existing standards (DCM, HL7) to facilitate electronic exchange of nursing oncology care data?

This paper further describes the background, such as earlier research in medical databases, a short introduction of the participants, and the various concepts involved in this project. The Methods describes the selection of useful DCM

for oncology patients. The Results section explains which nursing data and DCM are selected, the requirements for the CDWH, and the legal aspects. A blueprint of the CDWH is presented, and the paper ends with the discussion and conclusions for the project.

## II. BACKGROUND

### A. Clinical Data Ware House

Huff et al are among the firsts to define a clinical data warehouse as a separate system linked to an EHR [5]. Sahama & Croll define a Data Warehouse as a data structure optimized for distribution, mass storage and complex queries [6]. An ISO Technical specification guides a standard approach to CDWH development [7]. These and other inputs lead to the main tasks for this project:

- Privacy regulations prevent different uses of an individual's clinical data. In order to allow electronic exchange and secondary use of clinical data, regulations, consent, and data protection measures have to be followed [1].
- Performance of the EHR might be jeopardized when large queries are carried out on the records directly [3]. Similarly all clinical data required for continuity of care and specified purposes might not be available in one system and need to be collected and combined from multiple data sources [7].
- Clinical data are useful for continuity of care. Electronic exchange of data requires a high level of standardization of nursing data in order to allow their use by receiving EHR [2, 7, 9].
- The EHR format is not necessarily fit for secondary data use. This format can be both its structure e.g. type and number of record tables, nature of the data (free text, numbers, coded), and use versus non-use of standard or local terminology and coding. To use EHR data for exchange and secondary uses it is required to export data from the EHR, standardize and collect data in the format for secondary data use, and import the data into a clinical data ware house (CDWH) [3, 4, 5].
- Data have to be aggregated for secondary data use [3]. Aggregation involves a process of selection of subsets of data relevant for the purpose, and to collect and store them separately [8]. Next, the intended manipulations to the data have to be defined, such as construction of numerators and denominators, the querying of the data set, and their calculations or statistical analysis [8].

Hence, for our project, a new type of CDWH is defined: use of a polyvalent database with structured data, including

the data context, linkage to medical knowledge and meta-information, using unique code from standardized terminologies that facilitate data exchange for continuity of care, and queries for different purposes.

#### B. Participants in the project

The project is carried out by the Windesheim department of ICT innovations in health care. This lector led research group focuses on the quality of information management in health care, meaningful application of health information technologies, and their evaluation. Researchers and students from Windesheim School of Information Sciences (SIS) and School of Health Care (SHC) participate in this project.

Isala Clinics Hospital is the largest Top-clinical hospital in the Netherlands, with 5300 employees and 1000 beds. Per annum Isala Clinics treats more than 502.000 patients in out-patient visits, 47.000 patients via admissions and 48.000 in day care. Besides all basic care and treatment, top clinical functions offered include heart surgery, neonatal intensive care, trauma center, neurosurgery and bone and stem cell transplants. Innovation, education and research is carried out by the Isala Academy, with 180 to 200 annual peer reviewed publications in health research.

Icare Homecare is one of the brands of the Espria concern for home care, child care, maternity care, mental health care, among others. Espria has 16.000 employees at full concern level. Espria offers living, welfare, care and social services. Icare offers the home care, mainly via professional nurses, one specialty concerns the oncology visits in collaboration with Isala Clinics.

#### C. Detailed Clinical Models (DCM) and Health Level 7 (HL7) v3 Care Record Messages

Detailed clinical Models (DCM) link clinical concepts (data) via linkages to standard terminologies and codes to define the semantics, and identify the clinical context and knowledge. DCM handle the conceptual level and add a logical model in Unified Modeling Language (UML) to represent it [9]. For implementations, these UML models are transformed into the required technology.

The HL7 working group 'Patient Care' developed HL7 v3 messages allowing clinical content to be exchanged between EHR's, or systems for secondary data use [7]. Each message has specific dynamics for the care process. In particular sender, receiver and subject data are included, as well as clinical data. The Care Record message is useful for sending structured nursing data between health care facilities for continuity of care [7]. DCM examples for clinical purposes include their transformation into HL7 v3 Care Record messages. This is using DCM in a Model Driven Architecture (MDA) approach [10].

#### D. Electronic Health Records (EHR) in Isala and Icare

Both health care institutions work with EHR systems of quite older generations with a major focus on administrative data and elementary clinical data. Both systems do not include all required nursing data for oncology care, and do not include a full nursing care plan. So one of the outcomes of this project is the functional requirements document

describing for one specialty what their future EHRs needs to be capable of. The Isala Academy is carrying out clinical research and needs the CDWH for this. At this moment a research data base is used and functions well for the specific purpose. However, it does require multiple manual data entry activities, and the consistency with clinical data and research data can be difficult to keep at some times. In addition, quality reporting is increasingly important. Hence, the polyvalent database in the form of a CDWH is seen as a strategic asset to standardize data, to obtain requirements for the EHR, and to facilitate clinical reporting and research by the oncology nurses.

### III. METHODS

The project started in September 2011 and follows different steps including identifying data needs of nurses for oncology care. Next is that these data are standardized using the DCM approach [9]. The oncology nurses reviewed and completed a data set and the relevant DCM.

The literature review of CDWH and analysis of current systems and data uses resulted in a complete set of requirements on which the data warehouse can be built. The final step is the actual functional design and the CDWH implementation.

Along the nursing data content and the technical requirements for the CDWH itself, also specifications for the HL7 Care Record message and its DCM based content where established.

In order to meet privacy regulations data with patient names and id for continuity of care, and data for quality studies, indicators, reporting and research will be anonymous. Using legal inquiry, the way how this can be done is described.

This is all illustrated in Fig 1. Current EHR and auxiliary system data will be mapped to the DCM specifications in the next phase, as are the query definitions.

### IV. RESULTS

#### A. Juridical & ethical aspects

In medical information systems, personal data is stored and processed in patient records. These data are personal and processing data about a person is subject to legislation i.e.: the Dutch Personal Data Protection Act [11], which is in alignment with the European Data Directive [12]. In fact, according to article 16 of the Dutch Personal Data Protection Act, data describing one's health is considered to be special data, of which processing is prohibited except for specific institutions and purposes identified by article 21 [13]. Therefore, special attention is required when processing medical information. This paragraph discusses the impact of juridical and ethical aspects on the design of the CDWH.

According to the Dutch Personal Data Protection Act, several dimensions of data processing are of special interest:

- Transparency. Inform the patient about what data is processed and by whom. Request approval.
- Purpose-specification. The purpose of data processing activities must match the purpose for data collection.

- Legitimate basis. Data processing activities are in line with business goals, patient approval and (inter)national legislation as supervised by the Dutch Data Protection Authority.
- Retention. Data will be stored for a limited time.
- Data Quality. The processes will include activities to ensure optimal data quality.
- Data Security. Shielding data from unauthorized use will be in line with available technical solutions.

In this case, medical data of patients are used for two purposes: data exchange with care partners and scientific research. A legitimate basis for these types of processing can be created by acquiring explicit approval from the patient. Currently, patients are informed about the use for scientific research only. In this project, Isala is advised to inform patients about the data transfer to business partners as well.

As part of this research a limited survey amongst patients indicates that patients may not feel comfortable with their data being used for unknown purposes. Acquiring approval should therefore include an explanation of the (implications of the) types of data which will be processed. Medical data is classified as the most sensitive group of data available; therefore, additional to creating a legitimate basis, data used for scientific research should be made anonymous by removing all personal identifications.

This has implications for the design of the CDWH. In the CDWH, special data marts for scientific research will be defined, providing researchers access to anonymous data only. Data determined for information exchange between care partners will be made available by separate data marts, giving authorized care partners access to identifying information. In the areas of retention and data quality, additional information needs to be acquired. In the remainder of the project, additional privacy enhancing technologies may be identified and added.

#### B. Nursing oncology data set and DCM

In order to find out which general patient information is needed in the transfer from the hospital to homecare for oncology patients, 4 lists were created containing general patient information and information on the care organizations and next, were distributed to 13 nurses working in both organizations. These 4 lists came from HL7v3 CMETSs person and patient and from data of the V&VN/NICTIZ project e-Overdracht [14]. The nurses were asked to score whether or not a specific item was important for them to know in the transfer of oncology patients. A total set of 80 data elements/concepts was determined and integrated into one general questionnaire and implemented in the CDWH. Then, the work group in the project started with reviewing existing DCM on their relevance to this questionnaire. This is necessary for the CDWH data and coding requirements. The DCM are partly available from earlier external projects and are drawn up according to the National DCM guideline for DCM projects in the Netherlands [9], and will be submitted to a national DCM repository [15]. A list of DCM was drawn up of which some need smaller alterations in order to be linked to the specific

group of patients. Available DCM that are found to be important are e.g. pain, nutrition, length, weight, sensitiveness etc. (Table 1).

TABLE I. DCM AVAILABLE AND SELECTED

Table 1: DCM available and selected	
Pain	Sensitiveness
Nutrition	Length
Diabetes treatment	Weight
Usage of appliances	Social information
Decubitus wound classification	Person / patient data (HL7 CMET)
Defecation	

Additional DCM were created to augment the existing examples. After some research, it was concluded that patients with esophagus carcinoma or stomach carcinoma need homecare and that for the nurses who provide homecare, it is important to know which tumor typology the patient has. Therefore, a DCM for the TNM classification of esophagus carcinoma and a DCM for TNM of stomach carcinoma was designed. Also, a start was made on the design of the treatment of esophagus cancer and tube feeding (Table 2). In the next phase of the project, more DCM will be selected or created.

TABLE II. ADDITIONAL DCM'S CREATED

Table 2: Additional DCM's created
Tumor typology esophagus carcinoma
Tumor typology stomach carcinoma
Generic data about cancer treatment esophagus cancer
Tube feeding care

#### C. CDWH Requirements

Analysis of the existing situation in the Isala clinics and Icare, and the literature study for the CDWH revealed a set of functional requirements for the CDWH. These are presented here:

- Data are standardized and structured according to DCM format in the Dutch National Guideline [9], and draft ISO technical specification DTS 13972. [16]
- Data are coded using a unique code per data element and a unique code per value where a value set is applicable. Codes are used from Snomed CT, LOINC, ICD-10, ICF, ICNP, among others.
- DCM's selected and created are accepted by the nursing profession with Isala and with Icare as fit for purpose and properly defined. To achieve this, a verification and adoption process was used.
- The CDWH shall be able to export data in HL7 v3 clinical statement format in order to have payload for the Care Record message [7].
- Specifications of the current Isala Research Database are to be met. For this reason, besides the basic DCM materials, the tumor typology from existing research was used for additional DCM.

- The data specifications and import /export for the EHRs and CDWH are based on national or international standards. The same will have to apply to the queries that will be developed in a later phase.
- Standards based automatic import from EHR to CDWH.
- The CDWH shall be able to import raw data from different systems, store them in the agreed DCM based standard format per data element with codes and contextual information.
- The CDWH shall be able to sort data and support different queries, among those that are Structured Query Language (SQL) based.
- The existing data in the hospital will be queried / extracted from EHR and auxiliary systems to be mapped and coded into the required content format for CDWH storage.
- The EHR and auxiliary systems data will be exported in XML format, or transformed to XML for import in the CDWH.
- The CDWH can export the standards based data to XML format.
- Storage of the standardized data in the CDWH shall not be independent, meaning not dependent on specific applications or application formats.
- The CDWH shall allow sorting, combining, grouping, calculating, and aggregating data using different techniques.
- The CDWH should allow defining different queries, reports, and other functions in order to generate data extractions and /or reports, quality indicators, and research data.

The addition, the requirements where used to draw up an initial blueprint for the CDWH. Figure 2 illustrates the main components of the CDWH approach. In particular the distinction that a separate data type is created per use is seen as beneficial for both privacy reasons and for ease of development and deployment.

#### D. Instructions

Ongoing part of the project is the preparation of teaching materials. These include instructions for nurses with respect to structured and coded data entry in EHR, and for use of results from the CDWH, for instance using quality indicators in practice to improve patient care directly. Further teaching materials will be the CDWH functional design aiming at students and technicians in health care to create such CDWH. A business case package aims at health care managers and decision makers to justify investments.

### V. CONCLUSION

There is a challenge for the nursing profession to reuse nursing data. Like with any goldmine: the mining needs a careful construction and security measures to be safe. We have discussed legal implications, consent for data exchange and consent for data use and data protection measures. In a first pilot, some patients were interviewed and think they need to know exactly what goes over the wire about them.

Nurses determined a baseline set of about 80 data elements, grouped in several DCM. Standardizing patient data via DCM methodology proofs a constructive way for professional content and technological specifications. DCM are available in a repository, allowing a jump start in the project because they adhere all to the same technical format. [9,16]

Requirements for the CDWH could be determined via the literature, however, review of current uses of different data bases in the Isala Clinics, did reveal additional requirements. In particular separating the EHR/IT use of data and the queries against the data in the CDWH is seen in practice as an important requirement to keep up system performance. The solution to create a separate data mart for each data use is not uncommon. However, what we have added here is the storage of identified patient data in the CDWH as part of the daily operations in the hospital, as such not requiring additional consent, and allowing measures for data corrections to be dealt with at that level.

Next, each purpose can identify the data required, ensure specific consent issues handled, and specify the query for the data mart. That data mart will store the data, the consent and only identified data will go in the data mart for the continuity of care exchange, and only anonymous data will go into the data marts for quality reporting and research purposes. Hence, for the patient it is always clear where his / her data are used for and consents can be given separated.

On the technical level, the import from data from several source systems into the CDWH is identified, the storage in CDWH, the allowed use in the data mart, and finally the exports to the relevant subsystems. This architecture is solid and allows a step by step completion. CDWH in itself are not new, however for polyvalent data use, and in particular DCM based data storage brings new dimensions.

For Windesheim, the set of methods and examples serves as a large library of instruments useful for ICT in health care projects.

#### ACKNOWLEDGMENT

This project is funded by SIA RAAK Publiek: project id: 2011-13-41 P. <http://www.innovatie-alliantie.nl/english.html>

Irene Krediet contributed to the nursing DCM work section. Frank Boterenbrood contributed to the CDWH methodology section, the criteria and the legal parts. William Goossen is responsible for the whole of the paper. We wish to thank our 11 students that worked on this project during the September 2011 – January 2012 semester.

#### REFERENCES

- [1] Dick RS, Steen EB (1991). The Computer-based Patient record. An essential technology for health care. Washington DC, National Academy Press.
- [2] Johnson SB. Generic data modeling for clinical repositories. J Am Med Inform Assoc. 1996;3(5):328-39.
- [3] Bellika JG, Sue H, Bird L, Goodchild A, Hasvold T, Hartvigsen G. Properties of a federated epidemiology query system. Int J Med Inform. 2007;76 (9):664-76.
- [4] Sahama TR., Croll PR, (2007). A Data Warehouse Architecture for Clinical Data Warehousing. In: John F.

- Roddick and James R. Warren, Eds.. First Australasian Workshop on Health Knowledge Management and Discovery (HKMD 2007), Ballarat, Australia. Conferences in Research and Practice in Information Technology, Vol. 68.
- [5] Huff SM, Haug PJ, Stevens LE, Dupont RC, Pryor TA. HELP the next generation: a new client-server architecture. Proc Annu Symp Comput Appl Med Care. 1994: 271–275.
- [6] ISO/TS 29585 Health informatics — Deployment of a clinical data warehouse. Geneva, ISO.
- [7] Goossen, W. (2009). Sending electronic nursing discharge messages using the HL7 v3 Care Provision standard. Stud Health Technol Inform, 146, 269-275.
- [8] Goossen WT. (2002). Statistical analysis of the nursing minimum data set for the Netherlands. Int J Med Inf 2002;68(1-3): 205-18.
- [9] Goossen Baremans ATM (Ed). Guideline DCM [Richtlijn DCM] 2010. Amsterdam, Amersfoort, Parelsnoer & Results4Care. Available at [http://results4care.wikispaces.com/DCM\\_UML\\_StyleGuide](http://results4care.wikispaces.com/DCM_UML_StyleGuide). visited February 3, 2013.
- [10] Michael van der Zel & William Goossen (2010). Bridging the gap between software developers and healthcare professionals. Model Driven Application Development. Hospital Information Technology Europe, vol 3, no 2, 20-22.
- [11] Wet Bescherming Persoonsgegevens WBP, [http://wetten.overheid.nl/BWBR0011468/geldigheidsdatum\\_04-10-2012](http://wetten.overheid.nl/BWBR0011468/geldigheidsdatum_04-10-2012), visited February 3, 2013. [Data Protection Law]
- [12] Dutch Data Protection Authority, [http://www.dutchdpa.nl/Pages/en\\_ind\\_wetten\\_eu.aspx](http://www.dutchdpa.nl/Pages/en_ind_wetten_eu.aspx), visited February 3, 2013.
- [13] Wet Bescherming Persoonsgegevens WBP, [http://wetten.overheid.nl/BWBR0011468/geldigheidsdatum\\_04-10-2012](http://wetten.overheid.nl/BWBR0011468/geldigheidsdatum_04-10-2012), visited February 3, 2013. [Data Protection Law]
- [14] Goossen, W. (2011a). Internal publication workpackage 2: Exchanging nursing data for oncology care with a Clinical Data ware house. Zwolle: Christelijke Hogeschool Windesheim.
- [15] Nictiz. Dutch repository of Detailed Clinical Models. Den Haag, Nictiz, <http://www.nictiz.nl/page/Standaarden/Eenheid-van-Taal/Detailed-Clinical-Models>, visited February 3, 2013.
- [16] ISO DTS 13972. Health Informatics: Detailed Clinical Models. Geneva, ISO.

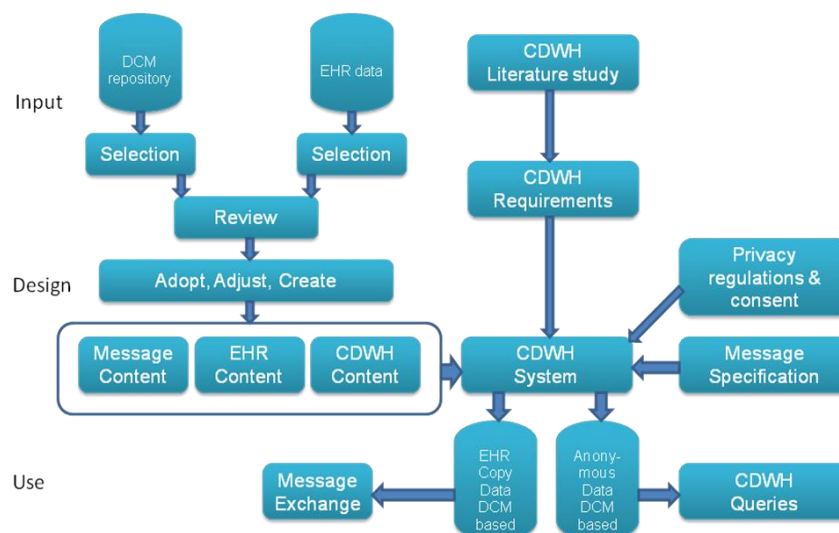


Figure 1. DCM review, adjust, create, CDWH & message specification and uses, and privacy measures for the developmental process.

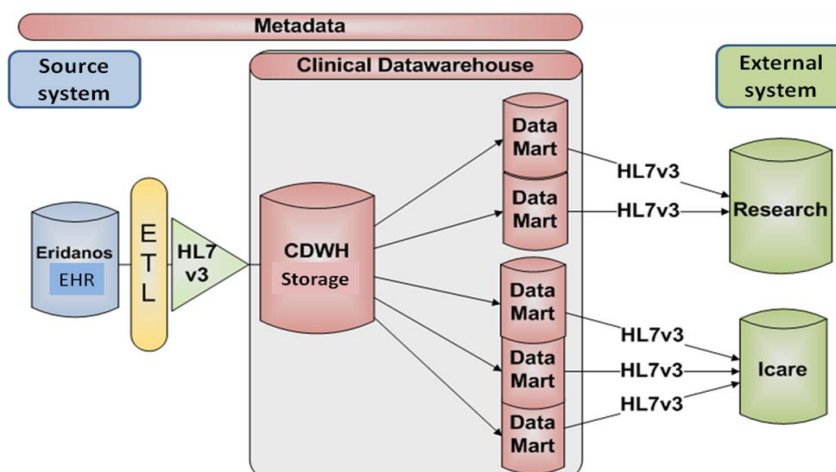


Figure 2. Architecture and main components for the CDWH for oncology nursing.