Modelling Knowledge Represented in Clinical Action Using a Boundary Objects Approach

Grace I. Paterson

Medical Informatics, Faculty of Medicine, Dalhousie University, 5849 University Avenue, Halifax, NS, Canada B3H 4H7, grace.paterson@dal.ca

Clinical documents contain statements and facts about patients under treatment. They help to coordinate care and transmit knowledge among the clinicians contributing to patient care. However, studies show that documentation of important clinical information is generally poor and miscommunication is a leading cause of adverse events. The discharge summary is a formal message between the hospital and the family physician. It has an informational role and serves as a boundary object for coordinating care. The research goal is to create an electronic discharge summary that improves the quality of the data and its interpretation. The method uses the same semantic markup standard, Clinical Document Architecture (CDA) from Health Level 7 (HL7), for structuring clinical and reference documents. It supports coding concurrent with composition of a clinical document. The coding systems and reference documents are building blocks. They function as boundary objects for representing, learning about and transmitting knowledge via the discharge summary. The hypothesis is that an infrastructure based on boundary objects will improve the quality and interpretation of data used in patient care.

Keywords
boundary objects, clinical document architecture, discharge summary, semantic interoperability, infrastructure

1. Introduction
The health record is a central source of data in the patient’s care process and influences the judgment made about the right course of action. It serves as a common point of reference in the information, communication and knowledge spaces inhabited by health professionals. The term, boundary object, was coined to describe objects, such as the health record, that inhabit several intersecting social worlds and satisfy the informational requirements of each [1].

The patient care process is supported by a team of health professionals—nurses, physicians, pharmacists, dieticians and others. Each newcomer to a health profession learns the language of care as part of the process of membership. Wenger defines the term, community of practice, as ‘groups of people who share a concern or a passion for something they do and who interact regularly to learn how to do it better’ [2]. It is the combination of three elements—domain, community and practice—that constitute a community of practice. The experiential knowledge of its members is expressed by unique vocabulary, artifacts and patterns of practice [2].

The health record is a particularly interesting boundary object to study. The hospitals have an infrastructure associated with record keeping. The tools provided by the medical records infrastructure include Patient Care Record committees, classification systems, forms, information systems and policies. The medical records personnel charged with manually coding the information in the record never see the patient. The coded abstracts are used for financial, epidemiological, disease management and health care evaluation purposes.
Reminders about deficient records from health records departments relate to the form of the record and not the content. The administrative sector uses documentation provided by clinicians for purposes of utilization review and quality control [3].

There are two things that need to occur before patient outcomes can be measured to improve medical practice. One is the creation of an electronic patient record so that clinical data can be pooled and analyzed. The second is the development of methods to extract meaning from an often complex medical record [4].

The information needed by nurses for their own purposes seems to be organised in idiosyncratic ways. The conclusion from one systematic review is that nursing work may not easily relate to the design of electronic patient record systems [5]. If a system is not easy and intuitive to use then there may be little point in trying to assess higher levels of evaluation such as the changes to patient outcomes.

The discharge summary is considered the most comprehensive document in the health record [6]. It is the source of information for individual patient care, health care evaluation, populations at risk, research and education [7]. It is a clinician-to-clinician communication. It is written by hospital-based specialists and nurse practitioners for community-based family physicians. A clinician leaves an intellectual footprint in his or her records. The major components are why the patient came, what the clinician found out and what the clinician did about the problem. Boundary objects, such as templates, classification systems and digital libraries, play a role in creating and interpreting the communication.

In our previous research, we found documentation of important clinical information was lacking in the health record. This quality of care issue may reflect a problem of record keeping, a performance gap, a knowledge gap or some combination of the three [8].

For knowledge to jump to a different community of practice it needs shared vocabulary. The discipline of writing clinical protocols and guidelines may constrain the terminology problem to a well-defined and relevant set of terms. The use of a shared and standard terminology suggests good common ground for shared conventional meanings within a community of practice [4]. Boundary objects have been conceptualised as a basic condition for the emergence of a mutual semantic foundation for cooperative work [9].

The motivation for this work is to improve the summarized information to benefit individual patient care and to ground the clinical action knowledge associated with care in explicit references. One component of the infostructure is a discharge summary template that prompts for relevant data. Concurrent coding of the information at the time of capture is a by-product of the process. The coded information enables the development of methods to extract meaning that can be used in knowledge transfer.

The Gricean maxims for the Cooperative Principle—quantity, quality, relation, manner—provide a guiding framework for evaluation of the system design [10]. The principle which clinicians are expected to observe is to make their communication 'such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged' [10]. The system needs to support shared care of a patient based on written dialogue between clinicians.

A boundary infostructure supports team work and coordination across different communities of practice [11]. This paper describes a boundary infostructure built to test the hypothesis that we can improve the quality of discharge summaries for chronic kidney disease patients using a template to prompt medical residents to enter relevant data. It considers how the gaps that may exist in record keeping, performance and knowledge can be filled.

2. Background

The application is discharge summaries for chronic kidney disease patients. This interdisciplinary project is a collaboration between a health informatician and a nephrologist. The collaboration is expressed through research protocols, grant writing and health informatics mentoring for a disease management project [12]. The focus of this paper is the
author’s research into boundary infostructures as a solution to the semantic interoperability problem.

The semantic interoperability problem arises from vocabulary differences in computer-supported cooperative work. Semantic heterogeneity is imposed on nephrologists working in a Canadian hospital setting. Multiple coding systems are needed for diagnostic coding within the community of practice and for shared care with the family medicine community of practice. A discharge summary on a nephrology patient is manually encoded into at least four diagnostic coding schemes to support health care evaluation, financial administration, disease registry and shared care functions. These include the International Statistical Classification for Diseases, 10th Revision, Canada (ICD-10-CA), ICD-9, the International Classification for Primary Care (ICPC) and the Canadian Organ Replacement Registry (CORR). There are controlled vocabularies for laboratory results and medications. A nomenclature, SNOMED, and a thesaurus, UMLS, are included to improve terminological coverage.

Semantic interoperability is a grand challenge research problem in the digital library community [13]. It is also a challenge in the health informatics community. This paper lays out a hybrid solution. It is influenced by the boundary objects work from the digital library community of practice, and the vocabulary and semantic markup standards from the Health Level 7 (HL7) community of practice.

3. Purpose

The research aims at creating a boundary infostructure that constrains the terminology problem sufficiently so that semantic interoperability among electronic health resources is achieved. It poses the question: Is an infostructure created from a semantic markup standard for documents and a network of boundary objects drawn from terminology and reference resources sufficient to improve the quality and interpretation of data used in patient care?

There are four objectives. The first is to identify the terms in the common information space used by the Nephrology community of practice. The second is to model the association between observations and actions and cite the boundary object that explains the association. The third is to improve the quality of the data in the discharge summary. The fourth is to support shared care of a patient based on written dialogue between clinicians.

4. Data Sources

A text corpus is created from all clinical and reference documents that are appropriate for the scope of this project. This is the data source for identifying the set of terms for the common information space.

The principle source for patient information is the paper chart of a patient with chronic kidney disease secondary to hypertension and diabetes mellitus. The paper record documents care in an acute care hospital, a community hospital, a chronic kidney insufficiency clinic, emergency care and dialysis care. He has a chronic condition with complications. His paper chart contains over 300 pages and spans over 20 years.

The boundary objects for information codification include HL7 Vocabulary, HL7 External Vocabularies and controlled vocabularies used for laboratory and medications. A semantic markup standard, HL7’s Clinical Document Architecture (CDA) Release 2.0 is used to structure clinical and reference documents. CDA is a refinement of the HL7 Version 3 classes from the Reference Information Model (RIM).

A clinical activity is something that is consciously done in order to achieve particular results. They can be classified into three groups [14]. Observational activities generate information on the patient’s comfort, function and likelihood of dying. Interventional activities attempt to
change or maintain the health status of the patient. Logistical activities support the process of providing care, such as admission to hospital.

The boundary objects used to produce reference documents are drawn from case writeups in medical education and diabetes, kidney disease and therapeutic guidelines. Almost all reference documents were authored by colleagues in the Dalhousie University community. They include a diabetes reference resource [15], an acute renal disease guide [16] and online therapeutic guidelines [17].

5. Design

A three-pillar model proposed by CEN/TC251 Working Group II (Healthcare Terminologies, Semantics and Knowledge Bases) is adapted for the semantic interoperability solution [18]. It includes a metadata registry for the data model, clinical templates for the grouping of complementary items and electronic terminologies.

The eXtended Markup Language (XML) is a common language to represent Web content. The boundary infostructure is created in XML. This enables linkages between different sections of the infostructure as well as electronic tagging for easier information retrieval. The CDA data model is expressed in an XML Schema. The clinical and reference documents can be created and viewed using the Internet Explorer web browser. Figure 1 shows the boundary object components available for the infostructure solution.

The clinicians are comfortable with a document-centric approach to automation. It is the one that most closely mimics the paper records with which they are familiar. An HL7 template for a CDA document was chosen as the tool to support data entry. An HL7 template supports the educational process by ensuring that residents are prompted to enter relevant data. The design addresses issues relating to gaps in record keeping, performance and knowledge through help files and references to explicit knowledge sources.

The clinical activities are documented in the discharge summary. The data elements are coded using one or more of the terminology resources. The reference documents are stored in a repository that can be linked to the discharge summary via a CDA reference type. The types of references supported by HL7 are: episode link, refers to, has support, has subject and is excerpt of. The CDA semantic markup has an attribute, BibliographicDesignationText, which enables the system to keep track of citations to the same boundary object by different reference documents.

![Figure 1 Boundary Infostructure Components](image-url)
6. METHODS

6.1 Identify the Terms in the Common Information Space

A text corpus is created from items in the patient chart, electronic guidelines and teaching cases. It is manually and automatically processed to get the set of concepts and the terms used to express them. The CLUE Browser Version 5 from Clinical Information Consultancy is used to obtain SNOMED CT and SNOMED Version 3.5 codes. It is a useful tool for understanding the relationship between SNOMED CT concepts. It makes use of description logics to express the meaning of a concept. The National Library of Medicine’s Unified Medical Language System (UMLS) is useful for mapping biomedical text from one coding system to another. The UMLS Knowledge Sources are a useful resource for concept definitions, semantic class for a concept, semantic relations between concepts and concept representation in multiple source vocabularies. The National Cancer Institute for Bioinformatics produced a Metathesaurus Browser that accepts a numeric code from any of the major coding systems and returns the concept in the neighbourhood of another coding system [19].

The text corpus is automatically processed by the UMLS MetaMap Transfer (MMTx) program, which discovers UMLS concepts in text [20]. The same text corpus is also processed by SnoCode® from Medsight Informatique Inc. for encoding text to SNOMED Version 3.5. There is continuing discussion as to what vocabulary standard is sufficient to be used to achieve interoperability. Health informatics research provides evidence that SNOMED is the preferred coding system for therapeutic information [21].

6.2 Model the Association Between Observations and Actions

The objective is to transform information accumulated for clinical purposes into knowledge. The information in an electronic discharge summary has a context and each data element is associated with semantic markup. The CDA constrains the Act classes to Observation, Procedure, SubstanceAdministration and Encounter. The ways for organizing knowledge represented in clinical action are adapted from Abidi’s methods [22, 23].

The structural mapping between clinical and reference documents is based on the CDA semantic markup for document sections. The reference document model uses a common concept template construct composed of observations, actions, decision and policies. As an example, the clinicians create the discharge summary. It is composed of observations (e.g., laboratory results), the actions taken (e.g., procedures), the decisions made (e.g., treatment plan) and the policies associated with creating the information (e.g., therapeutic guidelines). Figure 2 is the model for the clinical action references standardized to CDA and SNOMED.

![Figure 2 Clinical Action Reference Model](image-url)
The HL7 Reference Information Model (RIM) defines two classes specifically for CDA structured documents. The Class:Document is a specialization of the Act class and is used for attributes needed for document management. It has an attribute, Document:bibliographicDesignationText, defined as the ‘citation for a cataloged document that permits its identification, location and/or retrieval from common collections’ [24]. This attribute is useful for retrieval of a document from a specific electronic resource in the infostructure.

The references are categorized using HL7 classes: Observation, Procedure, SubstanceAdministration and Encounter. The HL7 coding systems registry assigns unique identifiers to code sets. These are expressed as Object Identifiers (OID). Our institution has an HL7 member OID root, so code sets in the local realm can be handled. The attributes for a coded concept are the OID for the coding system, the code system name, the code for the concept and the concept name.

To ensure matching accuracy, we define both numerical and vocabulary domains for each attribute value, and standardize the discharge summary values for structured data content.

MMTx software is used to parse the narratives in the text corpus (diabetes, kidney disease and therapeutic guidelines) and to map clinical concepts to unique UMLS concept identifiers. MMTx program takes an input sentence, separates it into phrases, identifies the medical concepts and assigns proper semantic categories to them according to the knowledge embedded in UMLS [20]. The output is manually filtered to select the best match for a clinical concept. The text corpus is also processed by SnoCode® from MedSight Informatique Inc. This natural language processing software encodes the clinical concepts in SNOMED Version 3.5. The UMLS and SNOMED ontologies serve as semiotic vehicles for sharing medical knowledge and matching two concepts. SNOMED CT is a vocabulary in the 2005 version of UMLS and is chosen as the code system for the reference documents.

6.3 Improve the Quality of the Data in the Discharge Summary

The intervention for improving data quality is the HL7 Template for Chronic Kidney Disease Discharge Summary. The HL7 Template is designed for data entry. It has some features that were requested by residents during the analysis and design. It prompts for data required to calculate the Glomerular Filtration Rate (GFR). The result is then used to populate the stage of chronic kidney disease. It provides lookup tables for diagnoses, laboratory results, medications and the reason for the medication. The system is implemented on an Internet server. Figure 3 is a screenshot of the GFR calculator.

![Figure 3 Screenshot of GFR Calculator](image)

Trainees are asked to complete the discharge summary from the same simulated patient chart as used for the baseline measure gathered from the current dictation and transcription method. They insert items into the chart in three ways: by typing into a text box, by choosing...
from a menu option or by selecting entries from a lookup file. They are provided with three forms of instruction: a quick reference guide, a printed manual and an online help file. The user saves the CDA document in XML. The system generates a discharge summary using a style sheet expressed in XSL. The student can view the discharge summary he created at baseline using the dictation system, the one he created using the HL7 Template and the optimal one created by the investigators using the HL7 Template. Unstructured data is entered as text and is not coded at time of data entry. Structured data is coded concurrent with data entry using the following coding systems:

- SNOMED-CT (License Agreement with College of American Pathologists)
- UMLS (License Agreement with National Library of Medicine)
- ICD-10-CA (License Agreement with Canadian Institute for Health Information)
- ICD9
- ICPC (International Classification for Primary Care)
- Primary Renal Diagnosis (includes codes from CORR)
- Cerner Mnemonics for Laboratory Information System
- Drug Identification Number for Medications (Health Canada)

### 6.4 Support Shared Care of a Patient Based on Dialogue Between Clinicians

The CDA document is manually scored using an instrument developed by colleagues and previously tested in a randomized controlled trial. The score sheet grades the discharge summary by section and overall. It includes quality measures of conciseness, correctness, completeness, conformity and clarity. The Gricean maxims for the Cooperative Principle—quantity, quality, relation, manner—provide a guiding framework for the reference standard for the scoring system. Each measure is grounded in explicit references. The conciseness metric is associated with quantity, whereby one should not make the communication more informative than necessary. The correctness and completeness metrics are associated with quality, whereby one should say that for which there is evidence. The conformity metric is associated with relation, whereby one should say things related to the current topic using the relevant terms. The clarity metric is associated with manner, whereby one should avoid ambiguity and be orderly [10].

### 7. Results

The task of identifying the set of terms in the common information space requires improved filtering techniques. Search query categorizations and filters need to be further investigated [23] before implementing the clinical action reference model.

A set of coded concepts output from MMTx and SnoCode® were compared to see how well they matched the original text. The results were context dependent. The outputs of both systems were comparable. SnoCode® was faster to use and exhibited useful features. The output from the coder is given in French with English synonyms, but the input can be in English or French. The system includes the ICD9 and/or ICD10 codes where available and appropriate. Table 1 shows an example of encoded text.

<table>
<thead>
<tr>
<th>Actual Text</th>
<th>SnoCode® in French</th>
<th>SnoCode® in English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short duration of neutropenia</td>
<td>&quot;G-A219&quot; &quot;court&quot; *** ***</td>
<td>Short</td>
</tr>
<tr>
<td></td>
<td>&quot;G-7290&quot; &quot;Duration&quot; *** ***</td>
<td>Duration</td>
</tr>
<tr>
<td></td>
<td>&quot;DC-41210&quot; &quot;neutropenie&quot; ***</td>
<td>Neutropenia</td>
</tr>
<tr>
<td></td>
<td>&quot;288.0D70&quot; ***</td>
<td></td>
</tr>
</tbody>
</table>

"Improving the Quality of Health Information: an international perspective”
The boundary infostructure is currently a network of boundary objects that are linked together via the HL7 template for the Chronic Kidney Disease Discharge Summary. The coded concepts are bound to terminology systems. These include diagnostic coding systems (ICD-10-CA, ICD9, CORR and ICPC), mnemonics for lab tests, Drug Identification Numbers and SNOMED CT.

The HL7 Template for Chronic Kidney Disease Discharge Summary is undergoing beta-testing by the clinical educators, nurse practitioners and family physicians. Early results indicate the system supported the production of a discharge summary that obeys Grice’s Cooperative Principle.

8. Discussion

The boundary infostructure supports electronic data management of patients through data entry into an electronic discharge summary. The terminology supports the process of concurrent coding and serves as a common ground for understanding clinical activity associated with care. The clinical action reference model promises to be valuable for sharing knowledge across the clinical communities of practice. It provides a mechanism for data interpretation. It represents an epistemic categorization of the relationships between observations, actions, decisions and policies grounded in the writings of the clinical community of practice. It is a useful tool that sets the stage so that all stakeholders in the disease management community can learn from one another as well as teach one another in a care setting.

Acknowledgements

The research was supported by a training grant from Canadian Institutes for Health Research (CIHR) Strategic Training Program in Health Informatics.

References


18 Rossi-Mori A, Rector A. Short strategic study on international cooperation on issues of terminology. CEN/TC251N00-33. 2000.


