CREATION OF SNOMED CT CODE SET FOR RICK HANSEN SPINAL CORD INJURY REGISTRY (RHSCIR) DATA ENTRY FORMS

Fariba Behzadi
B00520159
far697721@dal.ca

Performed at
Division of Neurosurgery
3rd Floor - Halifax Infirmary
1796 Summer Street
Halifax, Nova Scotia

In partial fulfillment of the requirements of the Master of Health Informatics Program,
Dalhousie University
Internship Report for the period May 25, 2010 – December 10, 2010
Date Submitted: December 19, 2010
# Table of Contents

1. ACKNOWLEDGEMENTS ........................................................................................................... 3  
2. EXECUTIVE SUMMARY ........................................................................................................ 4  
3. Description of the organization ............................................................................................ 5  
   3.1 Rick's Story ....................................................................................................................... 5  
   3.2 Rick Hansen Spinal Cord Injury Registry project ................................................................. 6  
   3.3 Purpose of the RHSCIR Project .......................................................................................... 7  
   3.4 RHSCIR Project Objectives ............................................................................................. 7  
      3.4.1 Subject recruitment .................................................................................................... 9  
4. Author's part of project ......................................................................................................... 10  
5. Future work .......................................................................................................................... 24  
6. The Health Informatics Relation to the Author's Internship .................................................. 25  
7. Recommendations ................................................................................................................. 26  
8. References ............................................................................................................................. 28
1. ACKNOWLEDGEMENTS

I especially thank Dr. Paterson my supervisor for her guidance, encouragement and useful suggestions on the project.

I also thank Ginette Thibault-Halman, research coordinator Division of Neurosurgery for providing all the information.

I thank Deepak Swain who worked with me for a period of time on the project.

I am grateful to The Rick Hansen Institute for giving me the opportunity to work on this wonderful project.

I am thankful to all my instructors in the MHI program.
2. EXECUTIVE SUMMARY

The Rick Hansen Spinal Cord Injury (RHSCI) Institute granted this internship opportunity to the author, which began on May 25th and was extended to 10th December 2010. The author’s part of the project included the creation of a SNOMED CT code set for the Rick Hansen Spinal Cord Injury Registry (RHSCIR) data entry forms. This internship is a part of phase one of RHSCIR project.

There were fifteen data entry forms plus an updated diagnosis form to be coded. The author used SNOMED CT coding software to begin the process.

After the coding was done the author learned more regarding data exchange and the national RHSCIR database, which is the ultimate goal of phase one of the RHSCI Registry project.

Finally, clinical statements were generated from coded value sets. The author formed a clinical statements library by assembling the clinical statements and a Health Level Seven International Clinical Document Architecture (HL7 CDA) was built for data sharing and data exchange.
3. Description of the organization

3.1 Rick’s Story

On June 27, 1973, Rick Hansen sustained a Spinal Cord Injury (SCI) as a result of a car crash at the age of fifteen. He was paralyzed from the waist down, but slowly adapted to his new life and set new goals. His willpower and tenacity helped him achieve many of those goals, including winning nineteen international wheelchair marathons and six medals as a Paralympian. On March 21, 1985, Rick left Vancouver and wheeled around the world on his ‘Man In Motion World Tour’. “He inspired a nation, raised awareness and received over $26 million in donations for the cause. Following the Tour, the Rick Hansen Foundation was born. Among the Foundation’s accomplishments is the creation of the Rick Hansen Institute” (Rich Hansen Foundation, 2010).

“It began as a Foundation project and has grown in both size and success over the years. Today this independent institute is a Canada-wide collaboration dedicated to finding a cure and improving the quality of life for people who live with SCI and related-disabilities. The Foundation continues to provide leadership and assistance to the Institute in order to accomplish financial sustainability and expand its programs internationally” (Rich Hansen Foundation, 2010).
3.2 Rick Hansen Spinal Cord Injury Registry project

“The Rick Hansen Spinal Cord Injury Registry (RHSCIR) Project initiated in June 2003 to develop a national prototype database that would link researchers, clinicians, the spinal cord injury support communities as well as SCI patients across Canada to move translational research to evidence-based practices” (Mcginness & Kingsmill, 2009).

RHSCIR is deployed using a two-phase approach that relies on partnerships with participating RHSCIR sites where patients with SCI are admitted. Currently, there are over 1200 registered participants in the RHSCIR Project. Phase one is currently underway, and requires participating RHSCIR sites to install a local version of the RHSCIR software, which is the Subject to a Data Sharing Agreement that enables SCI data to be collected and stored locally, which is de-identified by local RHSCIR staff and transmitted (on a quarterly basis) to the national RHSCIR database in Vancouver, BC.

RHSCIR is currently limited to traumatic SCI cases as they present themselves at participating RHSCIR sites. Eventually, Spinal Cord Injury Solution Network (SCISN) intends to collect data about historical SCI cases from participating RHSCIR sites to capture both traumatic and non-traumatic SCI individuals across Canada. (Mcginness & Kingsmill, 2009).
3.3 Purpose of the RHSCIR Project

The primary purpose of the RHSCIR project is to provide researchers, clinicians, and health care professionals with a research and reporting tool that aggregates, stores and retrieves comprehensive national spinal cord injury data. This data tracks specific outcome actions for people with traumatic forms of SCI throughout Canada.

The purpose of creating a national RHSCIR database is to generate a clinical and epidemiological based information management system. The national RHSCIR database can promote collaboration and research, accomplish SCI information goals, and ensure data quality, remaining current with changing trends and issues in health care management by building relationships with government and industry sponsored clinical trials. (Mcginness & Kingsmill, 2009).

3.4 RHSCIR Project Objectives

The objectives of the RHSCIR Project are:

1. Clinical support and management is one of the objectives of the project to encourage extend a competent, efficient and effective national Spinal Cord Injury data recovery and management reporting service. Standardization of national SCI database created by management of the registry database and secure data warehouse.
2. The second objective of the project includes: research support to create a clinical and epidemiological based information service in order to endorse collaboration between scientists and clinicians, as well as provincial, national and international data exchange and collaboration.

3. Partnership with participating RHSCIR sites where patients with SCI are admitted and data quality improvement to demonstrate flexibility and compliance in helping partners achieve their SCI information goals.

4. Ensuring the quality of data stored in the national RHSCIR database. To achieve this goal there should be the establishment of standardized definitions, an identification of standard procedures of data retrieval from the sources, the establishment of an internal quality assurance mechanism to monitor data quality, the establishment of quality control monitoring tools for local Registry review for data accuracy, and an evaluation of Registry effectiveness and its impact on outcomes.

5. Business planning and future development.

The timeline of the project starts from the time of injury and is ongoing for the lifetime of the participant. This includes follow-up questionnaires at one, two, five, ten years and every five years thereafter from the date of injury and/or until the participant withdraws from the RHSCIR project.
3.4.1 Subject recruitment

Inclusion Criteria:

All of the following conditions must be present to be eligible to participate in the RHSCIR Project:

1. Patients with a traumatic SCI as defined as an injury resulting from the transfer of energy.

2. Patients admitted to a participating centre;

3. Patients diagnosed with SCI or classified as ASIA (American Spinal Injury Association) A, B, C or D or Cauda Equina (including those individuals who progress to ASIA E by discharge);

4. Patients who speak English or French, or when this is not the case, have a family member, friend, or a medical professional who can translate;

5. Patients of any age; who can provide either informed consent or assent

Exclusion Criteria:

1. Individuals with non-traumatic SCI (i.e. due to infection, cancer, spinal stenosis etc.).
2. Patients who do not speak English and a translator is not available (all attempts will be made to provide a translator) will not be approached for consent to participate in the project. (Mcginness & Kingsmill, 2009).

4. Author’s part of project

The author was employed by Rick Hanson Spinal Cord Injury institute (RHSCI) as a Health Informatics Intern to create a SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms) code set for the Rick Hansen Spinal Cord Injury Registry (RHSCIR) data entry forms. The author’s internship started on 25th May 2010.

She was provided with twenty one files of patients with spinal cord injuries admitted in Halifax Infirmary plus the XML (Extensible Markup Language) files of patients’ information. Horizon Patient Folder access for CDHA (Capital District Health Authority) was provided to the author as well. Each file was scanned and organized into several sections consisting of patient information such as emergency admission, hospitalization information, rehabilitation and discharge summary.

Special attention was given to the discharge summaries, which were usually typed from the surgeon’s or physician’s dictated tapes. The discharge summary is considered to be the most important part of the information extracted from patients chart. The discharge summary outlines the patient’s admission
diagnosis, previous medical history, allergies, social history, physical
examination at the time of admission, treatment course in the hospital,
discharge diagnosis and discharge recommendations.

The author compared the co-morbidity section of each discharge summary to
build a basic idea on the data collection style in patients’ files. Coding was
done by the head office in Vancouver which did not match any standard coding
system. After inquiring, the author was notified of internal coding system used
only by RHSCI.

Protocol of the RHSCI registry was provided to the author consisting of all
information on the project including RHSCI registry data set. The RHSCI
registry data set included all forms needed to be filled based on the patient’s
information. The RHSCIR data set included different forms: Demographics,
Sociodemographics, Medical history (General), Medical history(Injury),
Admission/ Discharge information, Consent status, Neurology, Interventions,
Hospital, Questionnaires, Diagnosis, Procedure, Respiratory, Pain and Trauma.

Every form consists of several sections and every section has several options.
For example the Medical History (General) form consists of a section called
Smoking currently, which has three options: (1) yes (2) No (3) Unknown
(Mcginness & Kingsmill, 2009).
For a period of time, there were two MHI students working on this project. The author was ultimately responsible for all forms, but initially worked on a subset of seven forms for further assessment and SNOMED CT coding.

SNOMED CT is a terminological source that can be implemented in software applications to communicate with clinically applicable information. SNOMED CT enables applications to provide effective support and delivery of high quality healthcare to individual people and populations.

SNOMED CT has an international and multilingual scope, which can be confined to communicate unique concepts and terms of particular organizations or localities. SNOMED CT’s terminology consists of concepts, terms, and interrelationships between them (Spackman, 2008).

SNOMED CT provides a common language that provides a consistent way of collecting, sharing and aggregating the data. Therefore, healthcare knowledge is more functional and accessible worldwide among specialists and sites of care when it is SNOMED CT coded (CûtÈ & Bisson, 2004).

There are a variety of tools for SNOMED CT coding. The author started working with Snocode software but due to some technical problems this was not practical. Snocode opens with Microsoft Word and the procedure for coding becomes complex as the output provided extensive details, which were not related or needed for this project. The Snoflake browser was the next choice; it
is an online SNOMED CT browser. The Snoflake browser uses a variety of techniques to find, arrange and retrieve codes.

One of the features of the Snoflake browser is a multi parent-child relationship viewer, which makes it easier to navigate SNOMED CT relationships between different concepts.

Figure 1; from Snoflake Browser (DATALINE SOFTWARE, 2009)

The Snoflake browser needs login with username and password every time and by default it logs off after a period of time (fifteen minutes without any activity), which was inconvenience for the author. The author was introduced to other SNOMED CT tool by her supervisor. CliniClue Xplore is a desktop browser,
which provides access to SNOMED CT constituent through a range of different views. Each view is displayed in a frame of the main CliniClue window. The frames can be resized by dragging the boundaries between them (Features of CliniClue Xplore, 2010). CliniClue is a Registered Trademark of the Clinical Information Consultancy Ltd. The Clinical Information Consultancy Ltd is a small independent consultancy that specializes in the use, representation and communication of clinical information (The Clinical Information Consultancy, 2007).
Figure -2; from Cliniclue Xplore (Features of CliniClue Xplore, 2010):

One of the advantages of SNOMED CT is that it consists not only of clinical vocabularies, but also non-clinical words. SNOMED CT currently contains over 360,000 medical concepts. Concepts are arranged into hierarchies and may have multiple parents.

Each concept is represented by a unique number and several concepts can be used simultaneously to describe a complex condition. The objective of the work
was to search for RHSCIR terms in the SNOMED CT system and determine how to express an RHSCIR term using SNOMED CT.

Forms included: Demographics, Sociodemographics, Medical History (General), Medical history (Injury), Admission/ Discharge information, Hospital and Diagnosis. The new upgraded diagnoses form, which contained two sections of Trauma and Non Trauma, was added to the project. The author had to determine whether the terms on these forms are SNOMED CT vocabulary or Health Level Seven International (HL7) vocabulary.

“Health Level Seven International (HL7) is the global authority on standards for interoperability of health information technology with members in over 55 countries. HL7’s vision is to create the best and most widely used standards in healthcare” (About HL7, 2007).

Interoperability between two or more systems includes their ability of communicating and exchanging data.

Functional interoperability is exchanging the information and Semantic interoperability is using the information that has been exchanged (Bonney, 2009).

HL7 was founded in 1987. HL7 Version 2 standards aim to support hospital workflows; they define a series of electronic messages to support administrative, logistical, financial, and clinical processes.
Development of HL7 version 3 (current iteration) began in 1995 with the aim to support all healthcare workflows. HL7 version 3 messages are based on an XML encoding syntax (Paterson, Shepherd, Wang, Watters, & Zitner, 2002).

The Reference Information Model (RIM) is the foundation of the HL7 Version 3 development process and is the final source from which the information-related content of all HL7 version 3 protocol requirement standards are retrieved.

RIM expresses the information content of the HL7 Working Group in Unified Modeling Language (UML). UML is a standardized modeling language in the field of software engineering that includes a set of graphic data to create a visual model. RIM is based on six "core" classes:

- **Act**, which represents any action that occurs.
- **Entity** represents any physical thing that takes part in health care.
- **Role** describes the task that entities play or provide as they participate in health care acts.
- **Participation** refers to an association between a Role and an Act.
- **Act_relationship** is the association between a pair of Acts.
- **Role_link** is the connection that exists between two co-dependent roles.

(Stevens, 2003)
HL7 uses two main types of code. The first type covers the specialized codes used for structural attributes and is defined by HL7 itself. The second type covers externally defined terms and codes such as SNOMED CT.

The author worked on Excel worksheet for SNOMED CT coding expressions to generate value set for each concept.
After coding all forms present in RHSCIR protocol, the author started coding the Diagnoses Form. The form contained two sections: Trauma and Non-Trauma. The Trauma section consisted of Cervical, Thoracic, Lumbar, and Other Trauma. The Non-Trauma section consists of Oncology, Myelopathy, Deformity-Major, Deformity-Spondylo, Inflammatory Athropathy, Degenerative, Infections, Co-morbidity-Other associated condition, Complications-Neurological, Complications-Local, Complications-Implant or Graft Related, Miscellaneous.
Cervical part of Trauma section was divided into: C0-C2 Trauma and C3-C7 Trauma. C0-C2 Trauma was sub-divided into more detailed conditions, which were coded in ICD 10 and integral coding of RHSCIR.

Dislocation has been coded as **S131: TC1**. The first part of the code (S131) is supposed to be ICD 10 coding, which the correct code according to Canadian Institute for Health Information is **S13.1** (International Statistical Classification of Diseases and Related Health Problems, 2009). The second part of the code is an internal coding, which means: **Trauma Cervical first row (TC1)**.

The ICD 10 description of S13.1 is ‘Dislocation of cervical vertebra’. This description does not cover the whole meaning of the field. The code should cover different descriptions of this field as: dislocation is in C0-C2, it is due to trauma, it is an open trauma or a closed trauma (not applicable here otherwise the code should determine that too).

The closest SNOMED CT description is: 44264009 traumatic dislocation of joint of cervical vertebra (disorder). This code covers all the information such as dislocation is due to trauma and trauma happens in cervical vertebra but the code does not determine the exact location of the trauma.

To overcome the lack of specificity, the author began to generate SNOMED CT expressions using compositional grammar for compound expressions. (Spackman & Gutai, 2008)
Compositional Grammar links the codes and their concepts with other codes using signs such as +, =, |, (), {}, :, and comma (,). An expression may consist of a single concept, followed by a description associated with that concept.

Therefore ‘Dislocation of Occipital-cervical due to trauma’ could express with three different SNOMED CT descriptions as:

<table>
<thead>
<tr>
<th>SNOMED-CT CODE</th>
<th>SCOMED-CT DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>44264009</td>
<td>Traumatic dislocation of joint of cervical vertebra (disorder)</td>
</tr>
<tr>
<td>363698007</td>
<td>Finding site (attribute)</td>
</tr>
<tr>
<td>410731008</td>
<td>Joint structure of first cervical vertebra (body structure)</td>
</tr>
</tbody>
</table>

Regarding combinatory method for joint terms in SNOMED CT compositional grammar, the clinical statement would be:

44264009|traumatic dislocation of joint of cervical vertebra|:363698007 |finding site|= 410731008|joint structure of first cervical vertebra|

Another example is ‘Three column burst fracture without dislocation’ from Diagnosis form C3-C7 Trauma section, which coded as **S122XO:TC20**.

First part of the code (S122) is ICD10 and it’s described as: Fracture of other specified cervical vertebra. This code does not describe the field. The author made a list of the SNOMED CT codes that each covers some parts of the field.
Pre-coordination means that all required concepts, that we need to be covered are included in the terminology in advance, so that a single code can capture the intended meaning. Example: 207903001 covers closed fracture, cervical vertebra and burst.

Post-coordination means that the desired meaning is represented by assembling one or more codes into an expression.
In the above post-coordinated expression all the needed descriptions are assembled together.

Although the concept has been coded it still cannot be shared with other sites and sources. The author had to learn about Clinical Document Architecture.

Clinical Document Architecture (CDA) is a part of the HL7 version 3 that is XML-based and intended to indicate the encoding, structure and semantics of clinical documents for exchange of information.

Based on XML, HL7, RIM and coded vocabularies, CDA makes documents both machine readable and human readable.

CDA’s priorities to patient care are promoting the longevity of clinical records and enabling policy-makers to control information requirements. CDA’s primary scope is the standardization of clinical documents in exchanging health care information. The CDA specifies the exchange markup of documents but does not specify the creation or management of documents (Health Level Seven, 2000), (Dolin, 2004).
CDAs consist of three main parts: Header, Body and Clinical Statement. “The HL7 Version 3 Clinical Statement is as a common pattern for the development of all types of clinical messages, used for the exchange of information between different computer systems” (Benson, 2009).

The merged clinical statement for a section of value set mentioned in Figure-4 is:

```
<observation classCode="OBS" moodCode="EVN" negationInd="false">
  <codeSystem="2.16.840.1.113883.6.96"
    codeSystemName="SNOMED CT">
    code="102478008" displayName="pre-existing condition">
    value="22298006" displayName="Myocardial infarction">
  </code>
</observation>
```

Each value from the field has a clinical statement. A clinical statement library is formed by assembling the statements together to build a CDA for the purpose of data sharing and data exchange.

### 5. Future work

1. HL7 CDA Implementation Guide to be completed.

The HL7 CDA Implementation Guide specifies a standard for the electronic submission of Spinal Cord Injury Registry (SCIR) entry forms. As a requirement of phase one of the Rick Hansen Spinal Cord Injury Registry (RHSCIR) Project,
all RHSCIR participating sites must install a local version of the RHSCIR software, subject to a data sharing agreement, that enables SCI data to be collected and stored locally and then transmitted (on a quarterly basis) to the national RHSCIR database in Vancouver, BC.

The Implementation Guide gives the opportunity to retrieve accurate and high quality information from all sites using the same guideline and software.

2. The HL7 CDA Implementation Guide needs to be mapped to pre-existing templates, some which are available from “Integrating the Healthcare Enterprise” (IHE). IHE is a scheme initiated by healthcare professionals to improve the way computer systems in healthcare share information. IHE promotes the coordinated use of established standards such as HL7 to address specific clinical need in support of best patient care (About IHE, 2010)

3. Testing the efficiency of the Implementation Guide. The HL7 CDA needs to be tested and determine that if it is implementable.

6. The Health Informatics Relation to the Author’s Internship

The scope of the RHSCI Registry Project deals specifically with management, storage and use of SCI data, both locally and at the national office within a defined privacy and security framework. The overall objectives of creating a national RHSCIR database is to create a clinical and epidemiological based information management system that ensures data quality and completeness
of the information retrieved from participants’ files. The objectives include: research support, partnerships and quality improvement, quality data and information, and business planning and future development. The creation of a SNOMED CT code set for (RHSCIR) data entry forms is included in phase one of the project. SNOMED CT coding improves the accuracy of data collection and accomplishes the goal of collecting and storing a comprehensive, national SCI data.

7. Conclusion and Recommendations

The quality of SNOMED CT coding can be improved by considering these recommendations:

1. Use of standard coding systems instead of internal coding to improve data quality and integrity.

2. Avoid non-clinical specific expressions such as: mild and moderate. For example in the co-morbidity section of medical history form, there is an option for kidney disease, which has been qualified as moderate to severe. Although mild and moderately low kidney functions are commonly used in the general population, from a clinical context the better way of expressing the disease would be refining it in stages or levels.

A SNOMED CT concept, such as "kidney disease", can be further refined with a severity code. The value set for severity is: moderate, severe, mild, mild to
moderate, moderate to severe, fatal, and life threatening severity. If the severity code were used, it would fit with current RHSCIR forms; regardless of whether or not it fits with how clinicians document the condition.

It will also fit with related initiatives. “An international classification for kidney health is being developed to facilitate research, patient care, and policy development. An inclusive classification system that avoids labeling is proposed that could be used in disparate international settings” (Catherine M Clase, 2004).

3. Avoidance of combining the values with ‘and/or’. This way of pointing to the condition can complicate the coding. Conditions can be coded separately or in a combined manner, which covers the specific required definition.
8. References


DATALINE SOFTWARE. (2009). *Snoflake browser, the free online SNOMED database*. Retrieved from http://snomed.dataline.co.uk


