

# Rigor in electronic health record knowledge representation: lessons learned from a SNOMED CT clinical content encoding exercise

Karen A. Monsen,<sup>1</sup> Robert S. Finn,<sup>1</sup> Thea E. Fleming,<sup>2</sup>  
Erin J. Garner,<sup>1</sup> Amy J. LaValla,<sup>1</sup> and Judith G. Riemer<sup>3</sup>

<sup>1</sup>School of Nursing, University of Minnesota, Minneapolis, MN, USA,

<sup>2</sup>Gillette Children's Specialty Healthcare, St. Paul, Minneapolis, MN, USA, and

<sup>3</sup>Consultant, Riverside, CA, USA

Rigor in clinical knowledge representation is necessary foundation for meaningful interoperability, exchange and reuse of electronic health record (EHR) data. It is critical for clinicians to understand principles and implications of using clinical standards for knowledge representation within EHRs.

*Purpose:* To educate clinicians and students about knowledge representation and to evaluate their success of applying the manual lookups method for assigning Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) concept identifiers using formally mapped concepts from the Omaha System interface terminology.

*Methods:* Clinicians who were students in a doctoral nursing program conducted 21 lookups for Omaha System terms in publicly available SNOMED CT browsers. Lookups were deemed successful if results matched exactly with the corresponding code from the January 2013 SNOMED CT-Omaha System terminology cross-map.

*Results:* Of the 21 manual lookups attempted, 12 (57.1%) were successful. Errors were due to semantic gaps, differences in granularity and synonymy or partial term matching.

*Conclusions:* Achieving rigor in clinical knowledge representation across settings, vendors and health systems is a globally recognized challenge. Cross-maps have potential to improve rigor in SNOMED CT encoding of clinical data. Further research is needed to evaluate outcomes of using of terminology cross-maps to encode clinical terms with SNOMED CT concept identifiers based on interface terminologies.

**Keywords** Education, informatics, interoperability, knowledge representation, Omaha System, SNOMED CT, terminologies

## INTRODUCTION

The expected and desired outcome of electronic health record (EHR) implementation is the flow of standardized data across settings, vendors and health systems resulting in the use of data to improve healthcare quality and population health at decreased cost (1–7). Achieving these goals depends on the rigorous use of specialized semantic tools in EHRs to accurately convey the

---

Correspondence: Karen A. Monsen, Associate Professor, School of Nursing, University of Minnesota, 5-140 Weaver-Densford Hall, 308 Harvard Street SE, Minneapolis, MN 55455, USA. Tel: 621-624-0490. E-mail: mons0122@umn.edu

intended clinical meaning through computer-readable codes (1–7). However, rigor in knowledge representation is daunting, given the complexity of human languages and diversity of clinical knowledge domains.

The Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT®) is the most comprehensive, multilingual clinical healthcare terminology in the world (4). It is intended to be used for consistent reporting in EHRs and for making information more accessible for research and decision-making (4–6). It has been chosen as a national clinical standard in many countries (4,6). For example, in the United States, Meaningful Use Stage 2 rule identifies SNOMED CT as a clinical terminology standard that must be used within certified EHRs (3,4). Therefore, it is critical for clinicians to understand the principles and implications of using SNOMED CT and other clinical standards for knowledge representation within EHRs (1). SNOMED CT is owned, maintained and distributed by the International Health Terminology Standards Development Organization (IHTSDO). The approximately 300 000 SNOMED CT concepts are organized by several hierarchies, and within each hierarchy, there are multiple levels of granularity (sub-concepts). Major hierarchies within SNOMED CT include clinical finding/disorder, procedure/intervention, observable entity, body structure, organism, substance, pharmaceutical/biologic product, specimen, special concept, physical object, physical force, event, environment or geographical location, social context and staging and scales (4). Every concept has a unique numeric concept identifier and one or more associated human readable terms called descriptions (4).

There are numerous specialized collections of SNOMED CT concepts that support clinician data entry in the EHR (Table 1) (3,7–15). Diverse clinical standards are incorporated as subsets or mapped within SNOMED CT, such as the International Classification of Diseases (ICD-9 and ICD-10); nursing interface terminologies (Clinical Care Classification, International

**Table 1.** Unified medical language system terminology resources: subsets and specialized maps (7).

---

#### Subsets

- CORE problem list subset of SNOMED CT
- Kaiser convergent medical terminology (CMT) subsets
- Nursing problem list subset of SNOMED CT
- Route of administration subset of SNOMED CT

#### Interface terminology maps

- International classification of nursing practice (ICNP) (9)
- North American nursing diagnosis association (NANDA) (14)
- Nursing intervention classification (NIC) (14)
- Nursing outcome classification (NOC) (14)
- Omaha System (11)

#### International classification of disease maps

- ICD-9-CM diagnostic codes to SNOMED CT map
  - ICD-9-CM procedure codes to SNOMED CT map
  - SNOMED CT to ICD-10-CM rule-based mapping and tool (NLM)
  - SNOMED CT® to ICD-10 rule-based mapping (prepared by IHTSDO)
  - SNOMED CT® to ICD-9-CM mapping (prepared by IHTSDO)
  - SNOMED CT® to ICD-9-CM rule-based mapping to support reimbursement (draft)
-

Classification for Nursing Practice, NANDA, Nursing Intervention Classification and Nursing Outcome Classification); multidisciplinary interface terminologies (Omaha System); and the Nursing Problem List Subset of SNOMED CT. The purpose of incorporating multiple interface terminologies, structured vocabularies and standards within SNOMED CT is to create a common reference terminology for use within EHRs, with the goal of facilitating the transfer of meaningful information (4).

The challenge of using the vast resource of SNOMED CT within EHRs is being addressed using artificial intelligence and human intelligence methods (16). Artificial intelligence methods such as terminology servers and databases are being developed to manage SNOMED CT encoding of local and standard terminologies and to facilitate the integration and sharing of information across systems (4–6,15). Automated methods for standardized code assignment must be supplemented and validated by human intelligence methods, including browser lookups and terminology maps (1,16). Lee et al. described a manual browser lookups method for assigning SNOMED CT concept identifiers to clinical terms that includes searching for terms and associated synonyms and other phrases to find relevant SNOMED CT concepts (16). Finally, terminology cross-maps for interface terminologies that specify correct SNOMED CT concept assignment have been developed by various terminology experts (7–15). Terminology cross-maps provide explicit linkages across terminologies by specifying codes for two or more terminologies related to common concepts. For example, the term “Pain” is a shared concept across several interface terminologies and SNOMED CT (7–15).

Terminology cross-maps for interface terminologies are freely available from the terminology organizations or by request from the US National Institutes of Health National Library of Medicine Unified Medical Language System (7–15). Scholars continue to expand and validate these terminology cross-maps using established protocols (9). To illustrate, a brief excerpt of the problem concepts within the January 2013 SNOMED CT – Omaha System terminology cross-map is provided in Table 2. The entire Omaha System-SNOMED CT terminology cross-map consists of 810 Omaha System concepts mapped to SNOMED CT concept identifiers (7).

Rigor in encoding clinical concepts is necessary to ensure accuracy of clinical knowledge representation and address threats to validity such as semantic gaps and synonymy or partial term matching (16–18). A semantic gap refers to a code, concept or term that does not fully represent the intended meaning, resulting in a difference in the intended meaning vs. the original meaning (17). A difference in granularity (level of detail) of two concepts is a type of semantic gap (17,18). Synonymy or partial term matching refers to incorrect matches to functionally redundant terms or synonyms, when two or more terms or descriptions within a terminology mean the same thing (18).

Evaluation of SNOMED CT concept identifier lookups success using terminology cross-maps provides experiential learning for clinicians and other informatics students about the importance of rigor in code assignments for accuracy in clinical knowledge representation (7–16,19–23). The purpose of this educational exercise was to evaluate the success of the manual lookups method (16) for assigning SNOMED CT concept identifiers to selected terms from a multidisciplinary interface terminology, the Omaha System (6,11,12).

**Table 2.** Excerpt of SNOMED CT – Omaha System terminology cross-map, January 2013 (7).

<b>Omaha System problem number</b>	<b>Omaha System problem term</b>	<b>SNOMED CT concept identifier</b>	<b>SNOMED CT description</b>
1	Income	365 552 003	Finding of income details (finding)
2	Sanitation	129 683 001	Sanitation finding (finding)
3	Residence	365 508 006	Finding of residence and accommodation circumstances (finding)
4	Neighborhood/workplace safety	129 685 008	Neighborhood/workplace safety finding (finding)
6	Communication with community resources	129 686 009	Finding related to communication with community resources (finding)
7	Social contact	129 874 008	Social interaction finding (finding)
8	Role change	129 876 005	Role change (finding)
9	Interpersonal relationship	225 706 007	Interpersonal relationship finding (finding)
10	Spirituality	129 886 006	Spiritual finding (finding)
11	Grief	224 965 009	Grief finding (finding)
12	Mental health	36 456 004	Mental state finding (finding)
13	Sexuality	118 199 002	Finding relating to sexuality and sexual activity (finding)
14	Caretaking/parenting	129 879 003	Parenting finding (finding)
14	Caretaking/parenting	129 880 000	Finding related to dependent adult care (finding)
15	Neglect	95 930 005	Victim of neglect (finding)
16	Abuse	95 921 002	Elderly person maltreatment (event)
16	Abuse	386 702 006	Victim of abuse (finding)
16	Abuse	397 940 009	Victim of child abuse (finding)
17	Growth and development	105 726 004	Age AND/OR growth finding (finding)
19	Hearing	118 230 007	Hearing finding (finding)
20	Vision	301 978 000	Finding of vision of eye (finding)
21	Speech and language	286 832 009	Speech and language finding (finding)
45	Oral health	116 337 000	Oral cavity finding (finding)
23	Cognition	373 930 000	Cognitive function finding (finding)
24	Pain	22 253 000	Pain (finding)
25	Consciousness	106 167 005	Consciousness related finding (finding)
26	Skin	106 076 001	Skin finding (finding)
27	Neuro-musculo-skeletal function	102 957 003	Neurological finding (finding)

(Continued)

**Table 2.** Continued

<b>Omaha System problem number</b>	<b>Omaha System problem term</b>	<b>SNOMED CT concept identifier</b>	<b>SNOMED CT description</b>
27	Neuro-musculo-skeletal function	106 028 002	Musculoskeletal finding (finding)
28	Respiration	106 048 009	Respiratory finding (finding)
29	Circulation	106 063 007	Cardiovascular finding (finding)
30	Digestion-hydration	129 908 006	Digestive-hydration finding (finding)
31	Bowel function	249 562 008	Bowel finding (finding)
46	Urinary function	106 098 005	Urinary system finding (finding)
47	Reproductive function	118 232 004	Reproductive finding (finding)
48	Pregnancy	118 185 001	Finding related to pregnancy (finding)
49	Postpartum	118 213 005	Postpartum finding (finding)
50	Communicable/infectious condition	191 415 002	Communicable disease (navigational concept)
35	Nutrition	106 088 004	Nutritional system finding (finding)
36	Sleep and rest patterns	106 021 008	Finding of sleep rest pattern (finding)
37	Physical activity	129 861 002	Physical activity finding (finding)
38	Personal care	129 817 005	Personal hygiene finding (finding)
39	Substance use	409 069 009	Finding related to substance use (finding)
40	Family planning	13 197 004	Contraception (finding)
41	Health care supervision	129 825 007	Healthcare supervision finding (finding)
42	Medication regimen	422 979 000	Medication regimen behavior finding (finding)

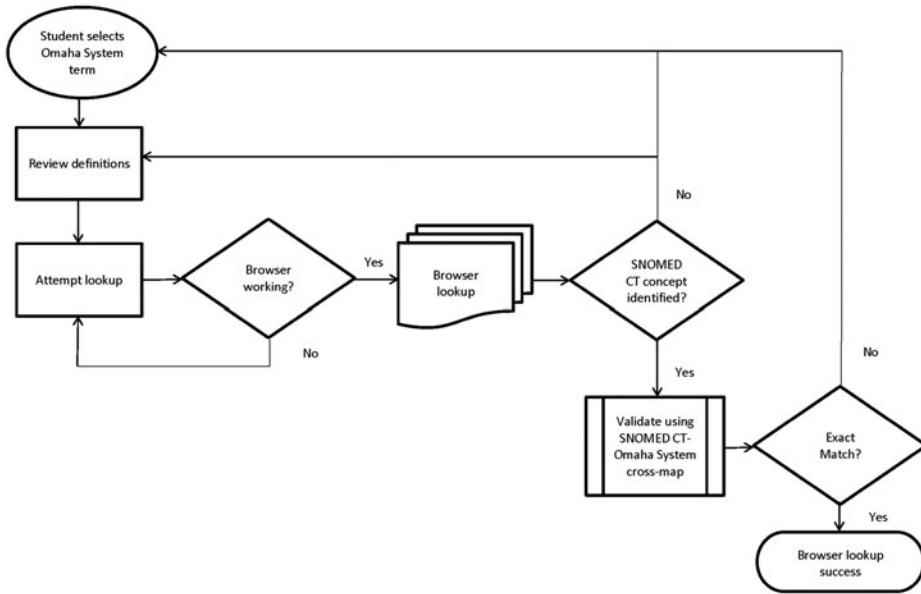
## METHODS

This educational exercise was a formal study conducted at a Midwestern university for six weeks from March to May 2013 and was exempt from review by the University Institutional Review Board. The Omaha System (11,12) was selected as an exemplar of an interface terminology that has been mapped within SNOMED CT and is available in the public domain. Similar to SNOMED CT, multi-disciplinary Omaha System Problem Classification Scheme concepts are organized within hierarchies with multiple levels of granularity including domains, problems and signs/symptoms. Omaha System problem terms are sub-concepts of four domains of health: environmental, psychosocial, physiological and health-related behaviors. Each problem term has a definition and a set of signs/symptoms that are unique to the problem.

Graduate students enrolled in Doctor of Nursing Practice studies at a Midwestern University (38 students working in pairs) searched for selected Omaha System terms in freely available SNOMED CT browsers, employing the manual lookups method for encoding clinical datasets with SNOMED CT (16). These graduate nursing students were experienced clinicians; however, this study was their first formal exposure to clinical knowledge representation and the SNOMED CT lookups method. At the beginning of the study, a terminology expert (the first author) conducted a hands-on, interactive training session with the students during which they completed a preliminary lookup according to the study procedures. First, the student pairs independently selected terms from the Omaha System that were of interest to them. Second, to establish a conceptual foundation for semantic interoperability, students reviewed the definitions of their selected terms in the literature. Third, students used the lookups method described by Lee et al. to assign a SNOMED CT Concept identifier to the selected Omaha System term (16). The SNOMED CT browsers used by students included the United States-based Unified Medical Language System Browser, the United Kingdom-based National Pathology Exchange and Snowflake Browsers, the Spain/IBM-based Integrated Terminology System browser and the Denmark-based SNOMED CT Browser©, which was developed at the University of Copenhagen Faculty of Health Sciences (7,19–23). During the six-week study period, each of the SNOMED CT browsers was unavailable for one or more weeks. This sporadic unavailability of browsers in some cases necessitated lookups for a single concept to be conducted in more than one SNOMED CT browser. Finally, students compared the manual lookups method results to the SNOMED CT concept identifier in the January 2013 Omaha System-SNOMED CT terminology cross-map (Figure 1). The SNOMED CT-Omaha System terminology cross-map provided a complete map of SNOMED CT concept identifiers with SNOMED CT descriptions that have been mapped to every Omaha System term by terminology experts (see excerpt, Table 2) (7,11,12). Success was defined as an exact match between the lookups result and the SNOMED CT concept identifier corresponding to the selected term in the SNOMED CT-Omaha System cross-map. Student narratives provided insights into the manual lookups method experience (24). Qualitative analysis of themes from the narratives was conducted by the research team (authors). The first author read each narrative three times, abstracting common themes with supporting quotes until saturation of the data was achieved. The other team members reviewed and critiqued this analysis. Differences were resolved by consensus. Themes relevant to the success of the lookups method are reported in this article.

## RESULTS

Of the 42 problem terms in the Omaha System, 16 were selected by students. Table 3 describes the selected terms with definitions. Within the Omaha System-SNOMED CT terminology cross-map, the 16 Omaha System terms expanded to 18 SNOMED CT concept identifiers due to the splitting of Omaha System terms for *Caretaking/parenting* and *Neuro-musculo-skeletal function*. *Caretaking/parenting* mapped to *Finding related to dependent adult care*



**Figure 1.** Educational exercise process for evaluating success of manual lookups method.

(*finding*) 129880000 and *Parenting finding (finding)* 129879003; and *Neuro-musculo-skeletal function* mapped to *Musculoskeletal finding* 106028002 and *Neurological finding (finding)* 102957003. Some student pairs selected the same term. Therefore, for each of the following terms, there were two student pairs conducting manual lookups: *Family planning*, *Pregnancy*, and *Respiration*. In total, 19 student pairs completed 21 lookups. Of 21 total lookups, 12 (57.1%) were successful (Table 4) and nine were unsuccessful (Table 5).

Initial lookup results reported by students varied, with some Omaha System terms resulting in no SNOMED CT concept identifiers (*Caretaking/parenting*, *Neuro-musculo-skeletal function* and *Spirituality*); and other Omaha System terms resulting in over 1000 related SNOMED CT concept identifiers (*Pain* and *Pregnancy*) (Table 3). The number of SNOMED CT concept identifiers returned when multiple browsers were used for a single concept varied by the SNOMED CT browser. In five instances, students used more than one browser to conduct the lookup; and four out of the five lookups (80%) resulted in different numbers of SNOMED CT concept identifiers returned for the same Omaha System term (differences ranged from 4 to 319) (Table 3).

Student narratives regarding use of the lookups method described issues related to facilitators (use of exact language) and challenges (synonymy or partial term matching, semantic gaps and lack of granularity). Students noted facilitation of the lookups method specifically for Communication with community resources and Sanitation problems:

The Omaha System term ‘Communication with community resources’ was found as an exact match [literally: Finding related to communication with community resources (*finding*)] in the SNOMED CT browser (Concept ID 129686009). The Omaha System factors that could affect communication

**Table 3.** Number of SNOMED CT concept identifiers returned from (21) Omaha System term lookups (browser if specified).

Omaha System term and definition	Number of SNOMED CT concept identifiers
Pain: unpleasant sensory and emotional experience associated with actual or potential tissue damage. Martin, 2005, p. 367.	1333
Pregnancy: period from conception to childbirth. Martin, 2005, p. 370.	1126
Pregnancy (see above)	not reported
Mental health: development and use of mental/emotional abilities to adjust to life situations, interact with others, and engage in activities. Martin, 2005, p. 363.	155, 474 (NPEx)
Respiration: inhaling and exhaling air into the body and exchanging oxygen. Martin, 2005, p. 368.	90 (UMLS)
Respiration (see above)	90 (NPEx)
Neglect: child or adult deprived of minimally accepted standards of food, shelter, clothing, or care. Martin, 2005, p. 365.	73
Family planning: practices designed to plan and space pregnancy within the context of values, attitudes, and beliefs. Martin, 2005, p. 372.	56, 19 (NPEx)
Family planning (see above)	25 (NPEx), 41 (ITS)
Income: money from wages, pensions, subsidies, interest, dividends, or other sources available for living and health care expenses. Martin, 2005, p. 361.	42, 46 (NPEx)
Grief: suffering and distress associated with loss. Practices designed to plan and space pregnancy within the context of values, attitudes, and beliefs. Martin, 2005, p. 363.	29
Physical activity: state or quality of body movements during daily living. Martin, 2005, p. 371.	16
Interpersonal relationship: associations or bonds between the individual/family/community and others. Martin, 2005, p. 363.	13
Social contact: interaction between the individual/family/community and others outside the immediate living area. Martin, 2005, p. 362.	4
Communication with community resources: Interaction between the individual/family/community and social service organizations, schools, and businesses in regard to services, information and goods/supplies. Martin, 2005, p. 362.	2
Sanitation: environmental cleanliness and precautions against infection and disease. Martin, 2005, p. 361.	1
Caretaking/parenting: providing support, nurturance, stimulation, and physical care to dependent child or adult. Martin, 2005, p. 364.	0
Caretaking/parenting (see above)	0
Neuro-musculo-skeletal function: Ability of nerves, muscles, and bones to perform or coordinate specific movement, sensation, or regulation. Martin, 2005, p. 368.	0
Neuro-musculo-skeletal function (see above)	0
Spirituality: beliefs and practices that involve faith, religion, values, the spirit, and/or the soul. Martin, 2005, p. 363.	0

UMLS, United States-based unified medical language system; NPEx, national nathology exchange; and ITS, integrated terminology system.



**Table 4.** Successful manual lookups for Omaha System terms in SNOMED CT browsers (12 of 21 Lookups).

Omaha System term	SNOMED CT concept identifier	SNOMED CT description
Caretaking/parenting	129 880 000	Finding related to dependent adult care (finding)
Communication with community resources	129 686 009	Finding related to communication with community resources (finding)
Family planning	13 197 004	Contraception (finding)
Grief	224 965 009	Grief finding (finding)
Income	365 552 003	Finding of income details (finding)
Interpersonal relationship	225 706 007	Interpersonal relationship finding (finding)
Neuro-musculo-skeletal function	106 028 002	Musculoskeletal finding
Neuro-musculo-skeletal function	102 957 003	Neurological finding (finding)
Pain	22 253 000	Pain (finding)
Physical activity	681 300 003	Physical activity finding (finding)
Pregnancy	118 185 001	Finding related to pregnancy (finding)
Sanitation	129 683 001	Sanitation finding (finding)

with community resources as ‘signs/symptoms’, whereas the SNOMED CT browser returned ‘subtypes.’ The signs/symptoms of the Omaha System and the subtypes of the SNOMED CT were nearly identical, with the exception that the Omaha System identified more ‘signs/symptoms’ than did the SNOMED CT identify subtypes. (24, p. 107)

...the term ‘sanitation’, was arbitrarily selected from the Omaha System. Next, the term ‘sanitation’ was searched in SNOMED CT. A total of 6 results were returned in SNOMED CT. Duplicate results were eliminated, leaving a total of 4 results. Next, the exact Omaha System term was located in SNOMED CT. It was titled ‘Sanitation finding’. This term was a subclass of ‘Environmental finding’ in SNOMED CT. The SNOMED CT concept ID for ‘Sanitation finding’ was 129683001 and matched the Omaha System-SNOMED CT map. (24, p. 29)

Students noted challenges due to synonymy specifically for respiration and pain problems:

When the term ‘respiration’ was entered in the SNOMED CT browser, 90 various terms or concepts populate the browser regarding the term respiration. None of 90 terms or concepts exactly matched the word respiration as found in the Omaha System. The concept of ‘finding of respiration’ was then chosen. This concept has 27 sub-concepts related to the different aspects of breathing...Of those concepts, we next chose ‘spontaneous respiration’, as it again seemed to most closely relate to the definition of respiration. After choosing ‘spontaneous respiration’, only one sub-concept was generated: ‘normal spontaneous respiration’ with concept ID 276888009 (24, p. 153)

**Table 5.** Unsuccessful manual lookups for Omaha System terms in SNOMED CT browsers (9 of 21 lookups).

Omaha System term	Incorrect SNOMED CT concept identifier	Description	Correct SNOMED CT concept identifier	SNOMED CT description
Family planning	410290005	Family planning education, guidance and counseling (procedure)	13197004	Contraception (finding) 13197004
Mental health	720821000000105	Mental Health (qualifier value)	36456004	Mental state finding (finding) 36456004
Neglect	418496002	Child neglect (finding)	95930005	Victim of neglect (finding) 95930005
Neuro-musculo-skeletal function	22171002	Neuro-musculo fatigue (finding)	102957003	Neurological finding (finding) 102957003
Pregnancy	237288003	Abnormal weight gain in pregnancy (finding)	118185001	Finding related to pregnancy (finding) 118185001
Respiration	276888009	Normal spontaneous respiration	106048009	Respiration (finding) 106048009
Respiration	301282008	Finding of respiration (finding)	106048009	Respiration (finding) 106048009
Social contact	423315002	Limited social contact (finding)	129874008	Social interaction finding (finding) 129874008
Spirituality	385988005	Spiritual care (finding)	129886006	Spiritual finding (finding) 129886006

The intent of the initial search was to identify the Omaha System concept “pain” in SNOMED CT. Initial search for “pain” in a SNOMED CT browser resulted in 1333 different concepts related to the word or concept “pain”. (24, p. 2)

Students noted challenges due to semantic gaps specifically for family planning and social contact problems:

According to the Omaha System, family planning is defined as ‘activities that support consideration and use of methods to prepare for and space pregnancy; practices designed to plan and space pregnancy within the context of values, attitudes and beliefs’ (Martin, 2005, p. 372)...In searching SNOMED CT, the term that was linked with the Omaha System concept of Family planning was contraception... We found this interesting because these two terms can have two very different implications in the health care setting and especially amongst providers. In the healthcare setting, an important distinction is made between family planning, which is a very broad concept, and contraception, which implies a more limited range of methods to plan for a family. (24, p. 8)

...because the human brain and human language is so complex, we have not yet reached a point in standardized terminology systems to completely rely on one to capture the nuances and uniqueness of a patient's clinical situation. Such a simple term [Social contact] can have multiple meanings and words can be misconstrued to illustrate a completely different issue/event/situation. (24, p. 80)

Finally, students noted challenges due to semantic gaps specifically for Neglect and Communication with community resources problems:

The U.S. Department of Health & Human Services provided a definition that can be applied and understood by both the general public and health care providers... [including] multiple forms of neglect within the definition by identifying physical needs, medical needs, educational needs, emotional needs, psychosocial needs, and risk taking behavior. These multiple forms of neglect were not easily identifiable within the SNOMED CT system. (24, p. 40)

SNOMED CT failed to acknowledge the sign/symptom of 'unavailable or limited access to resources' [of the Communication with community resources problem] such might be the case in a rural/under populated areas, or impoverished settings. (24, p. 107)

These comments also demonstrate student understanding of the importance of rigor in knowledge representation as a necessary foundation for semantic interoperability.

## DISCUSSION

This educational exercise and formal study evaluated the rigor of the manual lookups method described by Lee et al. for assigning SNOMED CT concept identifiers to selected terms from a multidisciplinary interface terminology, the Omaha System. Experienced clinicians who were graduate nursing students conducted 21 lookups for Omaha System terms, of which 12 (57.1%) were successful. Student narratives indicated that success in the manual lookups method was more likely when there was an exact Omaha System term match with the SNOMED CT description, and additionally described challenges of semantic gaps, granularity and synonymy. This study may contribute insight into the challenges of clinicians, software vendors and others tasked with using the lookups method to encode clinical content. Findings suggest that more accurate SNOMED CT encoding of concepts in interface terminology terms would be achieved using terminology cross-maps.

This study is an exemplar of an experiential educational exercise that engaged students in learning about rigor in clinical knowledge representation. Use of on-line SNOMED CT search engines and a terminology cross-map created a hands-on experience that simulated the process that is occurring in hospitals and EHRs. Such experiences may be beneficial for graduate students in preparing for informatics and clinical leadership roles.

Success of the lookups method was facilitated when the SNOMED CT description was the exact Omaha System term. However, one student failed to recognize an exact description match in SNOMED CT due to the addition of

a SNOMED CT descriptor (finding). Although students received initial training and practice at the beginning of the study, further SNOMED CT training may have reduced student confusion related to SNOMED CT descriptions and may have resulted in more successful code assignments. These results reinforce the need for informaticians with specialized expertise in all encoding methods and suggest that it may be important to educate clinicians about use of standardized terminologies for clinical knowledge representation. It should be noted that this study evaluated the accuracy of the lookup by comparing results to a terminology cross-map obtained from the terminology organization. Actual inter-terminology mapping for clinical purposes should be performed in collaboration with clinical experts and terminology developers such as IHTSDO; not by individuals or groups working separately (3,4).

Errors were noted in failure to map to the Omaha System term (semantic gap) and insufficient granularity; and incorrect maps to synonyms (synonymy or partial term matching). Examples of insufficient granularity were noted regarding *Neglect*, and the absence of a sign/symptom of *unavailable or limited access to resources* for the Omaha System problem *Communication with community resources*. Implications of limited language to describe serious health and social care problems using standardized terms are that such structured data may relegate important information to non-retrievable free-text or complete omission and diminish the clinical picture portrayed by the data.

Examples of semantic gaps were noted in the return of no SNOMED CT concept identifiers from a lookup (Table 3). Such gaps require human-assisted searching by synonym or other word-association, which may result in errors. Three Omaha System problem terms resulted in initial lookups with no returns: *caretaking/parenting*, *neuro-musculo-skeletal function* and *spirituality*. Of these, both *caretaking/parenting* and *neuro-musculo-skeletal function* are complex concepts that were mapped to two separate SNOMED CT concept identifiers, making lookups success unlikely for the exact term. Another form of semantic gaps was found in which definitions of mapped terms differed. For example, the broad term *family planning* found in the Omaha System was mapped to the SNOMED CT concept identifier 13197004 with description of *contraception*; a much more narrowly defined concept relative to prevention of pregnancy. Student reflections described potentially controversial implications of this semantic gap. Such a potentially controversial gap may have ethical consequences, depending on the source and re-use of the data.

Examples of synonymy or partial term matching were noted in the return of more than one SNOMED CT concept identifier from a lookup (Table 3). For example, students failed to identify *respiration (finding)* 106048009 for which 90 SNOMED CT concept identifiers were returned from each of two browsers; and instead selected *finding of respiration (finding)* 301282008 and *normal spontaneous respiration* 276888009. This is one example of numerous SNOMED CT descriptions identified in this exercise that are nearly exact synonyms but have different concept identifiers and thus threaten the validity of SNOMED CT use. This synonymy or partial term matching threat suggests a need for building new terminology cross-maps to achieve internal

harmonization among SNOMED CT descriptions and concept identifiers. Narratives provided insights into the students' perceptions of successes and challenges using manual lookups for assigning SNOMED CT concept identifiers to Omaha System terms. The themes found in student reflections have been mentioned in the knowledge representation literature (17,18). Student narratives highlighted the complexity of the manual lookups approach and were consistent with the findings that semantic gaps and synonymy or partial term matching occurred.

There was considerable between-browser variation in results of lookups for the same terms. For example, one SNOMED CT concept identifier (129880000) for the Omaha System term *caretaking* was found to have different descriptions in two browsers: *finding related to dependent adult care (finding)* and *finding related to care and support circumstances and networks* (Table 4). This may be due to differences in release dates (SNOMED CT is updated semiannually), national SNOMED CT versions or extensions. SNOMED CT is continuously updated and developed in response to requests for changes from IHTSDO's Working Groups and Special Interest Groups. Extant terminology cross-maps must be reviewed and updated subsequent to each SNOMED CT release. These findings underscore the importance of defining and maintaining domain-specific subsets and cross maps that enable consistent use of terms and concepts regardless of nation of origin, SNOMED CT release and SNOMED CT browser (4,6).

## IMPLICATIONS

Globally, manual lookups continue to be used to access SNOMED CT concept identifiers in numerous browsers (15,19–23). In addition to internet browsers, mobile apps enable SNOMED CT lookups in iPhone and iPod devices (e.g. *Snomobile*) and Android devices (e.g. *MedTalk SNOMED CT Browser*) (1,21,27). The value of multiple publicly available SNOMED CT browsers is that these codes and their descriptions are widely accessible. It is imperative to address inconsistencies among SNOMED CT browsers in order to support clinical knowledge representation, interoperability and the success of the manual lookups method.

The lessons learned from this educational exercise and study are applicable in clinical settings globally. Numerous EHR vendors and clinicians must select methods of SNOMED CT concept ID assignment for clinical terms used in their respective clinical settings (1). While SNOMED CT offers a valuable resource that will enable interoperability for multidisciplinary knowledge representation within the greater healthcare context, the method selected for assigning SNOMED CT concept identifiers is critical. Use of the SNOMED CT concept identifiers in terminology cross-maps would prevent the errors identified in this study and such rigor could improve SNOMED CT data quality for use in research and program evaluation. Currently, interface terminologies have separate terminology cross-maps within SNOMED CT. Early attempts to establish inter-terminology terminology cross-maps among nursing interface terminologies have begun (25,26). Further research is needed to evaluate use of terminology cross-maps to encode SNOMED CT concepts based on interface terminologies.

## CONCLUSIONS

An educational exercise provided clinicians in a graduate informatics class with hands-on experience in clinical knowledge representation using SNOMED CT and the Omaha System. Narratives demonstrated student learning about principles and implications of rigorous clinical knowledge representation. Further research is needed to evaluate outcomes of using of terminology cross-maps to encode clinical terms with SNOMED CT concept identifiers based on interface terminologies.

## ACKNOWLEDGEMENTS

The authors thank the graduate students who participated in the SNOMED CT concept identifier study.

## DECLARATION OF INTEREST

This study was conducted through the Omaha System Partnership.

## REFERENCES

1. Scichilone, R. [Internet]. SNOMED CT basics for clinical staff. American Health Information Management Association, 2013. Available from: <http://www.ihtsdo.org/show13/slides63.pdf> [last accessed 8 May 2014].
2. Gray BH, Bowden T, Johansen I, Koch S. Electronic health records: an international perspective on “Meaningful Use”. In issues in international health policy. New York: The Commonwealth Fund; 2011.
3. HealthIT.gov [Internet]. Meaningful use stage 2. Standards hub. Available from: <http://www.healthit.gov/policy-researchers-implementers/meaningful-use-stage-2-0/standards-hub> [last accessed 8 May 2014].
4. International Health Terminology Standards Development Organisation (IHTSDO) [Internet]. SNOMED CT. Available from: <http://www.ihtsdo.org/snomed-ct/> [last accessed 8 May 2014].
5. Vizine-Goetz D, Hickey C, Houghton A, Thompson R. Vocabulary mapping for terminology services. *J Digit Inform* 2006;4. ISSN: 1368–7506. Available from: <https://journals.tdl.org/jodi/index.php/jodi/article/view/114/113> [last accessed 17 Sep 2014].
6. International Health Terminology Standards Development Organisation (IHTSDO) [Internet]. Terminology services guide. Representing SNOMED CT in an application. Available from: [http://www.ihtsdo.org/fileadmin/user\\_upload/doc/en\\_us/rf1.html?t=tsg](http://www.ihtsdo.org/fileadmin/user_upload/doc/en_us/rf1.html?t=tsg) [last accessed 8 May 2014].
7. National Library of Medicine [Internet]. Supporting interoperability – terminology, subsets and other resources from NLM. Available from: [http://www.nlm.nih.gov/hit\\_interoperability.html](http://www.nlm.nih.gov/hit_interoperability.html) [last accessed 8 May 2014].
8. Rosenbloom ST, Miller RA, Johnson KB, et al. Interface terminologies: facilitating direct entry of clinical data into electronic health record systems. *J Am Med Inform Assoc* 2006;13:277–88.
9. Park HA, Lundberg C, Coenen A, Konicek D. Mapping ICNP Version 1 concepts to SNOMED CT. *Stud Health Technol Inform* 2010;160:1109–13.
10. Park HT, Lu DF, Konicek D, Delaney C. Nursing interventions classification in systematized nomenclature of medicine clinical terms: a cross-mapping validation. *Comput Inform Nurs* 2007;25:198–208.
11. Omaha System [Internet]. The Omaha System: solving the clinical-information-data puzzle. Available from: <http://omahasystem.org> [last accessed 8 May 2014].

12. Martin KS. The Omaha System: a key to practice, documentation, and information management (Reprinted 2nd ed.). Omaha, NE: Health Connections Press; 2005.
13. Matney SA, Warren JJ, Evans JL, et al. Development of the nursing problem list subset of SNOMED CT®. *J Biomed Inform* 2012;45:683–8.
14. College of American Pathologists (CAP) [Internet]. Press release published on January 25, 2005. SNOMED CT® Mappings to NANDA, NIC, and NOC Now Licensed for Free Access Through National Library of Medicine. Available from: [http://www.cap.org/apps/cap.portal?\\_nfpb=true&cntvwrPtlActionOverride=%2Fportlets%2FcontentViewer%2Fshow&cntvwrPtl%7BactionForm.contentReference%7D=media\\_resources%2Fnewsrel\\_nlm\\_mappings.html&\\_pageLabel=cntvwr](http://www.cap.org/apps/cap.portal?_nfpb=true&cntvwrPtlActionOverride=%2Fportlets%2FcontentViewer%2Fshow&cntvwrPtl%7BactionForm.contentReference%7D=media_resources%2Fnewsrel_nlm_mappings.html&_pageLabel=cntvwr) [last accessed 8 May 2014].
15. IHTSDO. [Internet]. SNOMED CT concept identifier lookup service. Available from: <http://www.ihtsdo.org/snomed-ct/snomed-ct-lookup-service/> [last accessed 8 May 2014].
16. Lee DH, Lau FY, Quan H. A method for encoding clinical datasets with SNOMED CT. *BMC Med Inform Decis Mak* 2010;10:53.
17. Phansalkar S, Robinson G, Getty G, et al. Challenges in exchanging medication information: identifying gaps in clinical document exchange and terminology standards. *AMIA Annu Symp Proc* 2009;2009:526–30.
18. Osgood CE, Suci G, Tannenbaum P. The measurement of meaning. Urbana: University of Illinois Press; 1957.
19. UMLS [Internet]. SNOMED Clinical Terms® (SNOMED CT®). Available from: [http://www.nlm.nih.gov/research/umls/Snomed/snomed\\_main.html](http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html) [last accessed 8 May 2014].
20. National Pathology Exchange (NPEX) [Internet]. SNOMED-CT browser. Available from: <http://www.snomedbrowser.com/> [last accessed 8 May 2014].
21. Snoflake Browser [Internet]. The Snoflake browser. Available from: <http://snomed.dataline.co.uk/Account/LogOn?ReturnUrl=%2f> [last accessed 8 May 2014].
22. Integrated Terminology Server (ITS) [Internet]. Integrated terminology server: adding value to clinical information. Available from: <http://www.itserver.es/ITServer/common/index.faces> [last accessed 8 May 2014].
23. SNOMED CT Browser [Internet]. The SNOMED CT browser ©. Available from: <http://www.medicalclassifications.com/SNOMEDbrowser/> [last accessed 8 May 2014].
24. Minnesota Omaha System Users Group [Internet]. A compendium of papers evaluating concepts in SNOMED CT and the Omaha System. Available from: <https://dl.dropboxusercontent.com/u/16553143/5115/Evaluation%20of%20Concepts%20in%20SNOMED%20CT%20and%20Omaha%20System.pdf> [last accessed 8 May 2014].
25. International Council of Nurses. June 2013 ICN Bulletin: eHealth terminology harmonizing activities. Geneva, Switzerland: International Council of Nurses; 2013.
26. Kim TY, Coenen A. Toward harmonizing WHO International Classifications: a nursing perspective. *Inform Health Soc Ca* 2011;36:35–49.
27. MedTalk [Internet]. MedTalk™ SNOMED CT® Browser. Available from: <http://www.pyebrook.com/medtalk%E2%84%A2-snomed-ct%C2%AE-browser/> [last accessed 8 May 2014].