HIR
Healthcare Informatics Research

Exploring the Possibility of Information Sharing between the Medical and Nursing Domains by Mapping Medical Records to SNOMED CT and ICNP

Eun-Young So, MS, Hyeoun-Ae Park, PhD
College of Nursing, Seoul National University, Seoul, Korea

Objectives: The purpose of this study is to explore possibility of information sharing between the medical and nursing domains. **Methods:** Narrative medical records of 281 hospitalization days of 36 gastrectomy patients were decomposed into single-meaning statements. These single-meaning statements were combined into unique statements by removing semantically redundant statements. Concepts from the statements describing patients' problem and medical procedures were mapped to Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) and International Classification for Nursing Practice (ICNP) concepts. **Results:** A total 4,717 single-meaning statements were collected and these single-meaning statements were combined into 858 unique statements. Out of 677 unique statements describing patients' problems and medical procedures, about 85.5% statements were fully mapped to SNOMED CT. The remaining statements were partially mapped. In the mapping to the ICNP concepts, 17.4% of unique statements were fully mapped, 62.8% were partially mapped, and 19.8% were not mapped. About 32.3% of 705 concepts extracted from the statements were mapped to both SNOMED CT and ICNP concepts. **Conclusions:** These mapping results suggest that physicians' narrative medical records can be structured and can be used for electronic medical record system, and also it is possible for medicine and nursing to share patient care information.

Keywords: Information Management, Medical Records, Controlled Vocabulary, Systematized Nomenclature of Medicine, International Classification for Nursing Practice

I. Introduction

Throughout the healthcare sector, introduction and utilization of information systems is becoming widespread. Electronic medical records, which are the most crucial component of hospital information systems, improve the accessibility of medical information and contribute to the readability and completeness of records, allowing users to search for and use information with more ease through greater integration of information [1,2].

However, in order to use such electronic medical records more efficiently, and to facilitate the smooth sharing and exchange of information between systems and medical institutions, it is imperative for an electronic medical record system to be based on a controlled vocabulary [3]. In nursing,

an electronic nursing records system based on International Classification for Nursing Practice (ICNP), a controlled nursing vocabulary, was introduced in early 2003 in Korea [4], and went so far as to use the data gathered using this system in decision-making and research [5]. But in the case of physicians' records, only fragmentary information such as chief complaints [6], decision-making rules [7], discharge summaries, diagnoses, and operation names [3] has been mapped to Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT). Records that compose a great part of all medical records, such as admission notes, progress notes, and summary discharge notes are still left in unstructured free text format.

Meanwhile, records regarding the condition and care process of the patient are made based on the professional knowledge of a multidisciplinary nature, only confined to a specific specialty. Therefore, the smooth sharing of information among different specialties is necessary in order to receive comprehensive and complete information about the patient. Specifically, it might be useful for doctors and nurses - the two professions most directly involved in patient care - to share and use each other's information when appropriate. However, fluid communication is difficult due to the fact that medicine uses a medical terminology system and nursing uses its own to express information. Thus, examining how well information expressed in different terminology systems can be shared may be a meaningful endeavor for the future of information use in medical, nursing practices, and research.

Such research on the possibilities for exchange and sharing of medical and nursing information is still in its early stages, however in a study that mapped ICNP Version 1 catalogue concepts to SNOMED CT, Park et al. [8] found that 92.5% of ICNP catalogue concepts were expressed by SNOMED CT. Based on such comparisons between terminology systems, it is necessary to go a step further and confirm whether the same results can be obtained in clinical medical records. Therefore, in the present study, we map doctors' electronic medical records documented in free-text form such as admission notes, progress notes, summary discharge notes to SNOMED CT concepts in order to explore the possibility of structured data input; how much the medical records can be expressed in ICNP concepts, a nursing terminology system; and how much information can be shared between the medical and nursing domains.

II. Methods

1. Data Collection
We analyzed the free-text admission, progress and summary discharge notes of gastrectomy patients who were admitted to the Department of General Surgery in a tertiary hospital in Korea. Medical records of the gastrectomy patients were chosen for analysis because gastrectomy is one of the most frequently performed surgeries in Korea with a relatively well defined care procedure. In order to limit the medical records of the gastrectomy patients, we eliminated the records of patients who were transferred to other departments before or after the surgery, or who had other operations performed on them simultaneously.

Medical records of the gastrectomy patients admitted before September 30, 2009 were collected in reverse chronological order. Taking into consideration of monthly rotation of the residents at the study hospital, we only included three patients' medical records per month in the pool for analysis. We collected the free-text portions of the patients' medical records, dissected them into single statements by meaning, and continued the process until there were three patients who no longer yielded statements with new meanings (saturation sampling). As a result, we collected 4,717 single statements from the medical records of 36 patients, documented by 19 doctors over a period of 281 days.

Research approval was obtained through the research review committee of the study hospital.

2. Analysis of Data
The collected statements often overlapped in meaning, although they were expressed differently by different doctors. By combining the statements by meaning, a total of 858 unique statements were extracted. We classified the extracted unique statements into those that describe the "medical condition of the patient" (symptoms, test results, diagnosis and etc.), those that describe "medical procedures performed on the patient" (treatment, medication, care plans and etc.), and "other statements" (patient's habits and other administrative information).

Out of 858 unique statements, 677 statements describing the "medical condition of the patient" and the "medical procedures performed on the patient" were the target of analysis in this study. First, we dissected each unique statement into concepts and mapped them to SNOMED CT (2009-07-31 international edition) concepts using the CliniClue Xplore browser. If a pre-coordinated concept existed, it was mapped to that concept first. Results of the mapping were classified into "fully mapped", "partially mapped", and "not mapped". Next, using the web browser from ICNP C-space (http://icnp.clinicaltemplates.org/), we checked to see how these unique statements would be mapped with the ICNP, a nursing terminology. If the ICNP catalogue concept coexisted

with the 7-axis model concept, we mapped to the catalogue concept first. Results of the mapping to ICNP were also classified into "fully mapped", "partially mapped", and "not mapped".

Lastly, in order to explore the possibility of information sharing between the nursing and medical domains, we compared the concepts mapped to SNOMED CT with those mapped to the ICNP. Results of the comparison were classified into "can be mapped to both SNOMED CT and ICNP", "can be only mapped to SNOMED CT", "can be only mapped to ICNP", and "cannot be mapped to either". In the case of statements that were "able to be mapped to both SNOMED CT and the ICNP", the type of mapping between the two terminology systems was further classified ("lexically mapped", "semantically mapped", "mapped to a broader concept", "mapped to a narrower concept", and "mapped to more than one concept"). A diagram of the analysis process is presented in Figure 1.
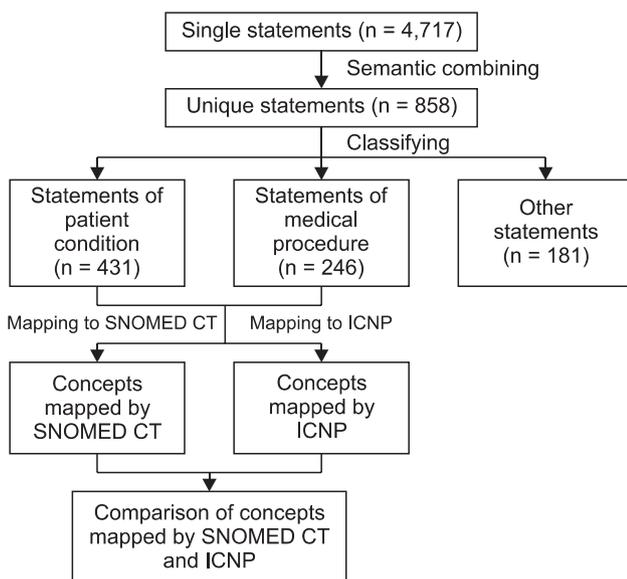
Figure 1. Research process.
SNOMED CT: Systematized Nomenclature of Medicine Clinical Terms, ICNP: International Classification for Nursing Practice.

3. Validation
Results of extracting the concepts from the statements and mapping them to SNOMED CT and ICNP concepts were verified by a group of experts. The experts consisted of a surgeon of who performs gastrectomies at the study hospital, a nurse with PhD degree in nursing informatics with experience in SNOMED CT mapping research, a doctoral student with experience in ICNP mapping research and nursing informatics research, and a student with master's degree who maintains electronic nursing records using the ICNP at the study hospital. The experts were presented with the mapping results along with possible replacement concepts, and asked for their opinions. The mapping results were finally modified based on their verification.

III. Results

1. Mapping to SNOMED CT
When 677 unique statements describing the "medical condition of the patient" and the "medical procedures performed on the patient" were dissected into concepts and mapped to SNOMED CT concepts, 579 unique statements - 85.5% of the total - were fully mapped and the remaining 14.5% were partially mapped. There were no statements that were not mapped to SNOMED CT concepts at all. Out of 4008 statements before removing redundancy in meaning, 3,740 single statements - 93.3% of the total - were fully mapped to SNOMED CT concepts. Statements that described "the medical condition of the patient" (91.9%) showed a higher rate of being fully mapped than statements that described "the medical procedures performed on the patient" (74.4%) (Table 1). A total of 705 concepts were extracted during the course of the mapping.

2. Mapping to the ICNP
When the unique statements were mapped to the ICNP, 17.4% of all statements were fully mapped, 62.8% were partially mapped, and 19.8% were not mapped. In statements before removing redundancy in meaning, the rate of being

Table 1. Mapping of statements to Systematized Nomenclature of Medicine Clinical Terms

	Patient conditions		Procedures performed		Total	
	Total statements	Unique statements	Total statements	Unique statements	Total statements	Unique statements
Fully-mapped	3,071 (96.8)	396 (91.9)	669 (80.0)	183 (74.4)	3,740 (93.3)	579 (85.5)
Partially-mapped	101 (3.2)	35 (8.1)	167 (20.0)	63 (25.6)	268 (6.7)	98 (14.5)
Total	3,172 (100.0)	431 (100.0)	836 (100.0)	246 (100.0)	4,008 (100.0)	677 (100.0)

Values are presented as number (%).

fully mapped was slightly higher with 34.8% of 4,008 single statements fully mapped to the ICNP (Table 2).

3. Comparison of Concepts Mapped to SNOMED CT and the ICNP

When the agreement between concepts mapped to SNOMED CT and the ICNP were compared, 230 of the 705 concepts - a rate of 32.6% - could be mapped to both SNOMED CT and the ICNP. In terms of the types of mapping, 10.5% were lexically mapped, 7.5% were semantically mapped, 4.8% were mapped to a broader concept, 0.4% were mapped to a narrower concept, and 9.4% were mapped to more than one concept. Meanwhile, 61.0% of all concepts were mapped to SNOMED CT but not to the ICNP; 0.6% were mapped to the ICNP but not to SNOMED CT, and 5.8% were not mapped to either terminology systems. Taking the frequency of concept appeared in the statements into consideration, 705 concepts were appeared a total of 9,415 times, and 51.6% were mapped to both SNOMED CT and the ICNP (Table 3).

IV. Discussion

The present study mapped doctors' free-text electronic medical records of gastrectomy patients to SNOMED CT and the ICNP, and compared the concepts mapped to the two terminology systems in order to explore the possibility of structured data entry of free text medical record and mutual exchange and sharing between medical and nursing domains. In this section, we aim to analyze the results of the study as well as propose some suggestions we came up with during the course of the study.

We found that more than 85% of free-text gastrectomy patients' medical records documented by doctors in admission, progress and discharge notes were mapped to SNOMED CT. This is similar to the content coverage of SNOMED CT to represent the most common nonduplicated patient problems seen at the Mayo Clinic [9]. In the Mayo Clinic's research, SNOMED CT, when used as a compositional terminology, can represent 92.3% of the terms used commonly in medical problem lists. This implies that that SNOMED CT can

Table 2. Mapping of statements to International Classification for Nursing Practice

	Patient conditions		Procedures performed		Total	
	Total statements	Unique statements	Total statements	Unique statements	Total statements	Unique statements
Fully-mapped	1,273 (40.1)	90 (20.9)	121 (14.5)	28 (11.4)	1,394 (34.8)	118 (17.4)
Partially-mapped	1,553 (49.0)	255 (59.2)	608 (72.7)	170 (69.1)	2,161 (53.9)	425 (62.8)
Not-mapped	346 (10.9)	86 (19.9)	107 (12.8)	48 (19.5)	453 (11.3)	134 (19.8)
Total	3,172 (100.0)	431 (100.0)	836 (100.0)	246 (100.0)	4,008 (100.0)	677 (100.0)

Values are presented as number (%).

Table 3. Comparison of Concepts Mapped to SNOMED–CT and ICNP

	Unique concept	Total concept
Can bemapped to both SNOMED-CT and ICNP	230 (32.6)	4,862 (51.6)
Lexically mapped	74 (10.5)	856 (9.1)
Semantically mapped	53 (7.5)	2,339 (24.8)
Mapped to a broader concept	34 (4.8)	638 (6.8)
Mapped to a narrower concept	3 (0.4)	28 (0.3)
One-to-many mapping	66 (9.4)	1,001 (10.6)
Can be mapped only to SNOMED CT	430 (61.0)	4,273 (45.4)
Can be mapped only to ICNP	4 (0.6)	118 (1.3)
Can not be mapped to either	41 (5.8)	162 (1.7)
	705 (100.0)	9,415 (100.0)

Values are presented as number (%).
SNOMED-CT: Systematized Nomenclature of Medicine Clinical Terms, ICNP: International Classification for Nursing Practice.

be used to structure free-text doctors' medical records. The unique statements used more frequently in medical records are mapped even better to SNOMED CT. This is much greater compared to the studies in which free-text nursing records were mapped to the ICNP in order to structuralize nursing records [10,11]. If the ICNP - with its relatively low coverage - could be used to structure and utilize an electronic nursing record system, the possibility of structuring free-text doctors' medical records seems very high.

In addition, the mapping rate to SNOMED CT was higher with the statements that described the "medical condition of the patient" than the statements that described the "medical procedures performed on the patient". In the current electronic medical record system, information on "medical procedures performed on the patient" is relatively easy to use, due to procedures were coded because they are used in doctors' orders and reimbursement. However, the medical condition of the patient - especially the patient's symptoms or the doctor's judgments and opinions - usually remains unstructured as free-text and is therefore difficult to search for later. If such records become structured based on SNOMED CT, the information will become very useful.

In mapping to SNOMED CT concepts, statements describing test results such as "platelet: */mm" or "total calcium: *mg/dl" were imbued with value judgments regarding the results. Thus the appropriate concepts were first searched in the "clinical finding" hierarchy, then in the "observable entity" hierarchy. However, concepts describing some clinical laboratory tests could not be found in the abovementioned hierarchies, and existed only as a concept in the "procedures" hierarchy. In these cases, the concepts were considered as not mapped. An example is the hepatic enzyme "GPT (Glutamic Pyruvic Transaminase)". GPT only existed as "alanine aminotransferase measurement (procedure)" and was thus classified as not mapped.

Such issues of inconsistency were present not only in clinical laboratory tests, but also in some pre-coordinated concepts. For example, "no sputum (finding)" or "not hoarse (finding)" are pre-coordinated concepts in SNOMED CT, but "no dyspnea" is not a pre-coordinated concept in SNOMED CT and needed to be post-coordinated with "dyspnea (finding)" and "absent (qualifier)" concepts. Not all concepts must be expressed as pre-coordinated concepts, but issues of inconsistency may arise when similar types of concepts are expressed partly through pre-coordinated and partly through post-coordinated concepts. Post-coordination may prove to be useful in terms of data utilization, so certain principles regarding these situations must be established.

Meanwhile, mapping free text medical record to the ICNP yielded lower mapping rate compared to mapping to SNOMED CT. Only 17.4% of all statements were fully mapped, and even when partially mapped statements were included, 80.2% could be expressed through the ICNP. This is perhaps an expected result, since the ICNP is a terminology system created to express content exclusive to nursing - such as nursing diagnoses, nursing interventions, and nursing outcomes - and content exclusive to the medical domain was excluded from consideration since its very developing stages.

Despite such differences in inherent purpose and scope between the two terminology systems, 32.6% of the 705 concepts extracted from free text medical records of the gastrectomy patients could be expressed in both SNOMED CT and the ICNP. If the repetitive use of concepts is taken into account, 51.6% of all concepts can be shared with the nursing terminology system. Although modification and supplementation through revision is necessary with respect to the 67.4% of information that could not be shared, it is still notable that 32.6% of information can be shared and exchanged between the two domains.

The present study are limited to gastrectomy patients at the Department of Surgery at a tertiary hospital in Korea, and continued further research into the possibility of structuring doctors' records and the sharing of information between the medical and nursing domains is necessary through the analysis of medical records in other areas.

In this study we explored possibility of structured data entry of free text medical records by mapping free text statements documented in the admission notes, progress notes and discharge notes of gastrectomy patients to SNOMED CT. More than 85% of free-text statements were fully mapped to SNOMED CT. This result implies that physician's free text medical records can be structured and thus structured data entry is possible.

We also explored the possibility of data exchange between medicine and nursing by mapping the same free text statements to ICNP, a nursing terminology. More than 32% of concepts extracted from the free text statements were mapped to both SNOMED and ICNP. This implies that the medical and nursing domain can share patient data to some extent. Information shared in this manner can be used not only in research, but in practice. For example, data may be shared in order to minimize overlapping collection of information, or a clinical decision support system may be created in which one type of profession is informed of the clinical content recorded by the other.

Conflict of Interest

Acknowledgements

References

1. Dick RS, Steen EB, Detmer DE. The computer-based patient record: an essential technology for health care. Washington, D.C: National Academy Press; 1997.
2. van Ginneken AM. The computerized patient record: balancing effort and benefit. Int J Med Inform 2002; 65: 97-119.
3. Kim SH, Han SB, Choi J. The expressive power of SNOMED-CT compared with the discharge summaries. J Korean Soc Med Inform 2005; 11: 265-272.
4. Cho IS, Park HA, Chung EJ, Lee HS. Formative evaluation of standard terminology-based electronic nursing record system in clinical setting. J Korean Soc Med Inform 2003; 9: 413-421.
5. Park HA, Cho I, Chung E. Exploring use of a clinical data repository containing international classification for nursing practice-based nursing practice data. Comput Inform Nurs 2011; 29: 419-426.
6. Chin HJ, Kim S. Standardization of main concept in chief complaint based on SNOMED-CT for utilization in electronic medical record. J Korean Soc Med Inform 2003; 9: 235-247.
7. Kim HY, Cho IS, Lee JH, Kim JH, Kim Y. Concept representation of decision logic for hypertension management using SNOMED CT. J Korean Soc Med Inform 2008; 14: 395-403.
8. Park HA, Lundberg CB, Coenen A, Konicek DJ. Evaluation of the content coverage of SNOMED-CT to represent ICNP Version 1 catalogues. Stud Health Technol Inform 2009; 146: 303-307.
9. Elkin PL, Brown SH, Husser CS, Bauer BA, Wahner-Roedler D, Rosenbloom ST, Speroff T. Evaluation of the content coverage of SNOMED CT: ability of SNOMED clinical terms to represent clinical problem lists. Mayo Clin Proc 2006; 81: 741-748.
10. Park IS, Shin HJ, Kim EM, Park HA, Kim YA, Jo EM. Mapping nursing statements with the ICNP and its practical use in electronic nursing records. Stud Health Technol Inform 2006; 122: 989-990.
11. Cho I, Park HA. Evaluation of the expressiveness of an ICNP-based nursing data dictionary in a computerized nursing record system. J Am Med Inform Assoc 2006; 13: 456-464.