Comparison of Validity of Mapping between Drug Indications and ICD-10

Direct and Indirect Terminology Based Approaches

1. Introduction

The need for mapping terminology in the medical domain is increasing. Mapping an interface terminology to a reference terminology can be used in clinical decision support systems (CDSS), and a mapping reference terminology to a classification such as ICD-10 can be used in health insurance claims [1, 2].

In general, terminology mapping is divided into lexical mapping and semantic mapping. Whilst the lexical mapping relies on the same lexical structure of terms, the semantic mapping method focuses on the conceptual connection between terms [1, 3–6].

There are two approaches to semantic mapping. One is a direct approach to map the terms to corresponding concepts directly utilizing the coder’s knowledge of medical and pharmaceutical terminologies. It is relatively time consuming and expensive [3–5], and the consistency of the data is difficult to guarantee [7]. The other is an indirect terminology based approach that sets a reference terminology for the terms, and maps the reference terminology to the target terminology or a classification, such as ICD-10. Less time and labor is needed because a significant portion of the manual work may be replaced by computer applications; therefore better consistency and accuracy may be expected because synonyms and homonyms are processed through the mapping process with the reference terminology [1, 6].

A systematic appraisal of the manual review process was carried out by the investi-
gators, which had not been previously reported by other authors; however, direct approaches have commonly been used to validate the results of automatic mappings [3, 5 – 8] as gold standards. However, direct mappings as gold standards in previous studies were just a premise to accomplish the studies.

Fee for service has been the primary payment system for health insurance in Korea. All performed medical services and assigned ICD-10 codes for associated diagnosed diseases must be stated on the submitted claims forms in order to be reimbursed. The health insurance reviewers at the Health Insurance Review and Assessment Service (HIRA) review all insurance claims to check the relevance between the stated services and diseases on behalf of the payer. The same process is applied to the review of drug prescriptions. The National Health Insurance Service (NHIS), the payer, reimburses the health institutions only when the stated diseases and the indication of prescribed drugs are met with relevance [9].

However, some physicians overlook to write prescriptions according to the indications. For example, when treating patients with multiple diseases that require several drugs, they may be completely unaware of the indications for all drugs and the reimbursement policy implications. Also, each year numerous new drugs are being introduced to the market, and therefore it is not easy for them to remember the contents of the approved drug labels for every drug precisely. According to a study targeting medical specialists in the U.S., only 55% of all correspondents correctly identified disease information service company tried to build up the mapping between drug indications and ICD-10 in 2011. The former HIRA, built the mapping by indirect terminology based approach and tried to apply it to an electronic claims review system. The first priority of the application domain was the cardiovascular drugs. The latter built the mapping by direct approach and tried to provide it to their customers for the preparation of insurance claims as well as the patients’ safety. They used two different approaches to achieve the same goal, providing a good opportunity to compare the validity of the two mapping approaches.

2. Objectives

This study was to compare the validity of a direct approach and an indirect terminology based approach of mapping indications in the approved cardiovascular drug labels to ICD-10 against the gold standard drawn from two mapping results.

3. Methods

3.1 Selection of Study Drugs

The number of listed drugs as products was 14,247 as of October 1, 2011. Among them 1,991 products belonged to the cardiovascular system of Anatomic Therapeutic and Chemical (ATC) classification. Of the 1,991 cardiovascular drugs, there were 375 reference listed drugs which were included in the databases of both institutions, which were subsequently selected for the study.

3.2 Mapping

In the direct approach, the indications were directly mapped to the Korean version of ICD-10 using coder’s knowledge. In the indirect terminology based approach of HIRA, a reference terminology was utilized in the mapping between the indications and Korean ICD-10.

Too coarse granular indications such as ‘diseases of cardiovascular system’ were excluded to avoid inappropriate mapping results and improved the mapping quality. Signs and symptoms were excluded from mapping too if they were not in the ICD-10.

3.2.1 Direct Approach

Two skilled nurse coders who had experience in the billing process reviewed the indications of each drug product by reading labels and searched for them in the Korean version of ICD-10. When corresponding ICD-10 codes to the indications of each drug product were found, they were entered into Microsoft Excel and then exported to the SQL server for the information service (Figure 1). A medical dictionary [15] and the index of the Korean version of ICD-10 were used as a reference to identify the meaning of the indications.

3.2.2 Indirect Terminology Based Approach

A health information manager extracted the indications from the approved drug labels. A nurse and two health information managers mapped all the indications to the Korean Standard Terminology of Medicine (KOSTOM) which is a UMLS [16] based medical terminology developed in Korea. The preferred terms (PT) are the representative terms for specific concepts, and were previously assigned to all terms in the KOSTOM. Terms that had the same concept were considered as synonyms [17]. When two words were spelt the same way, but had different meanings, they were given separate PTs as homonyms. If a term was not listed in the KOSTOM, it was added to the KOSTOM with an appropriate PT. Then, they mapped the PT terms in KOSTOM to the ICD-10 classification (Figure 2).

For convenience purposes, Microsoft OneNote was used during the mapping process to share various kinds of notes and MS Excel was used for the preliminary data input.

3.3 Construction of Gold Standard

Because an appropriate standard mapping table for indications and ICD-10 codes was not found, a gold standard was constructed by two pharmacists after comparing and evaluating the results of the two mappings.
processes that were carried out in this study. If they could decide on a set of ICD-10 codes for the indications of a drug, the review process for the drug was completed regardless of the discrepancy between the mapping results. If they failed to draw a consensus on the gold standard, or if clinical judgments or decisions related to the claims review were needed, a doctor, a nurse, four pharmacists and a health information manager met three times for the resolution. Specialists and associates of HIRA resolved the remainder of the discrepancies that were not able to be resolved during these consultation meetings.

3.4 Evaluation of Mapping Accuracy

The unit of analysis was a drug, because its use in the electronic claims review system was of primary interest to both institutions.

The number of ICD-10 codes that we matched between the study results and the gold standard was counted. When the number of codes mapped in the gold standard and each approach was the same, it was classified as ‘equal mapping’. When the number of codes mapped was higher in each approach than the gold standard, it was classified as ‘over mapping’, and the opposite was applied to the other cases which were classified as ‘under mapping’. A kappa statistic was calculated to see the compatibility of the two mapping approaches based on the classifications as above.

Precision, recall and F1 scores of each drug were calculated. The mean precision, recall and F1 scores were compared according to the type of mapping. The F1 score is a harmonic mean of precision and recall. It is considered to compensate the under mapping or over mapping caused by precision and recall. Paired t-test was applied to compare the accuracy of the mapping using SAS 9.1. A qualitative analysis of mapping was done with the F1 scores of each drug to elucidate the difference between mapping results.

![Diagram](image-url)

Figure 1 The direct approach
4. Results

4.1 Mapping Results of the Direct Approach

The number of indications was not known in direct approach because the coders mapped indications to ICD-10 codes directly and no record was available. However, it was presumed to be 2,028 as in the indirect terminology based approach. A total of 1,433 ICD codes were used in direct approach. The average number of ICD codes mapped to each drug was 46.32 (Table 1).

4.2 Mapping Results of the Indirect Terminology Based Approach

The number of indications from all study drugs was 2,028 and the drugs had 5.42 indications on mean (Table 1). Terms that had the same meanings were sorted out by KOSTOM, leaving 369 concepts for the next step. One hundred and thirty one concepts were excluded because they were not disease names, but signs, symptoms or concepts from Korean traditional medicine which are not covered by ICD-10. Finally, 238 concepts were selected for mapping to ICD-10 and a total of 1,291 ICD-10 codes were used for mapping. The average number of ICD-10 codes mapped to each drug was 56.94 (Table 1).

4.3 The Gold Standard

A total of 1,426 ICD-10 codes were used in the gold standard. The mean number of ICD-10 codes mapped to each drug was 39.12 (Table 1).

4.4 The Agreement between the Two Approaches

The number of equal mapping in the indirect terminology based approach was 156 and that of the direct approach was 62. The kappa statistic between the two approaches was 0.19 with 95% confidence interval of 0.12 – 0.26 showing poor agreement between the two approaches (Table 2).

4.5 Evaluation of Mapping Accuracy

The recall of the indirect terminology based approach (86.78%) was higher (p < 0.001) than that of direct approach (79.73%). However, there were no significant differences in the precision and F1 scores.
scores between the mapping approaches (►Table 3).

4.6 Evaluation of Mapping Approaches Based on F1 Scores

The F1 scores of both approaches were the same in 63 products. The indications of those products were clear for the coders to map to ICD-10. ‘Essential hypertension’ and ‘acute congestive failure’ were included in this category. The direct approach showed better F1 score in 108 products where the descriptions of indications were usually long, for example, ‘improvement of dizziness caused by chronic circulatory failure after cerebral infarction’ (Korean to English translation). The indirect terminology based approach showed better F1 score in 204 products in which age restrictions or a special situation such as pregnancy were considered.

5. Discussion

This study was conducted to compare the validity of a direct approach to an indirect terminology based approach of mapping indications on the approved cardiovascular drug labels to ICD-10 against the gold standard drawn from the two mapping results. The kappa statistic (0.19) showed that the results of the two mapping approaches are poorly compatible.

5.1 Treatment of Granularity

Indications are lists of diseases and symptoms used for drug approvals. The ICD-10 is a classification used primarily for morbidity and mortality statistics and further for insurance claims processing. Because the two terminologies have different purposes, they have different granularities. A clear understanding of the concept is required to prevent wrong mapping caused by the difference of granularities [19]. Therefore, we went through a process of identifying the granularity of indications before mapping them to the ICD-10. For example, the indication, ‘unstable angina pectoris’ can be precisely mapped to I20.0. However, the indication, ‘circulatory diseases’ can be mapped to all codes from I10 to I99. In this case the mapped ICD-10 codes can exceed the range of the actual indication and become useless.

5.2 Comparison to other Studies

There is limited number of mapping studies between drug indications to ICD-10. A similar study that mapped prescription information to ICD-10 using MeSH (Medical Subject Headings) [20] showed precision of 0.4 and maximum recall of 0.68. Another study that used SNOMED-CT [21] showed precision of 0.25 and recall of 0.46. In comparison, this study showed precisions of 0.82 and 0.86 in the direct and indirect terminology based approaches respectively, and the recalls were 0.80 in both approaches. These numbers indicated better results than the other studies. However, the gold standard in this study was built based on the consensus of two mapping results and the other studies compared automatic mappings with manual mappings. The intention of this subject is to get better precision and recall.

There is another recent study of mapping drug indications in SPL (Structured Product Labeling) to UMLS concepts with automatic mapping tools [7]. When compared with manual mapping the precision was 0.77 and recall was 0.95. However, the study lacked mapping to ICD-10 codes.

5.3 Problems of Korean Medical Terms

Korean medical terms have the problems of too many synonyms. For example,
The main difference of the mapping approaches was how to manage those problems described above. In the direct approach, all indications in each drug label were mapped to ICD-10 codes though the same indications appeared on the drug labels spending more time and resources. However, a terminology was introduced for the efficient management of those problems in the indirect terminology based approach.

### 5.4 Pros and Cons of Each Approach

The indications on the drug labels were written in free text similar to many other data items in the medical domain. When documents are written in unfettered free-text, they are very context-dependent and their meanings can be obscure making it difficult to prepare appropriate data for application [22, 23].

Because the coders can understand the context of the indications based on their medical knowledge and experience, they can find more appropriate codes to the indications in the direct approach. However, more time and manpower is required for the large number of indications in this approach, and many studies have pointed out the problems of the direct approach [4, 5, 7]. This approach is based on human intuitions and is prone to lose consistency.

The indirect terminology based approach can be used through a simple computer operation, and it reduces the time and manpower needed for generating and maintaining the mapping. Through the indirect terminology based approach, we were able to maintain the consistency of the mapping results by assigning preferred terms to different indication terms with the same concept and applying the same standard. For example; when particular codes must be added to or deleted from the disease code group representing 'high blood pressure', it is not necessary to update drugs that have 'high blood pressure' as an indication one by one, but simply update the mapping table between major indications and disease codes and thus maintaining its consistency.

### 5.5 Failure Analysis

The types of failure in the direct approach included omission and inconsistency of mapping. For example, hypertension to O10 in Xipamide and myocardial infarction to I21 and I22 in a generic drug of dopamine hydrochloride were not mapped by the coders. Both Nitrendipine and Benidipine have the same ingredients and indication of hypertension. But, only Nitrendipine was mapped to I13 resulting in inconsistent mapping.

Extraction errors and granularity problems were found in the indirect terminology based approach. For example, only 'stroke' was extracted from an indication, 'cardiovascular diseases in a patient whom had a previous stroke', and mapped to I63. If the mapping mechanism could interpret the context, the error would not have happened. Indications such as renal disease or shock were excluded in the indirect terminology based approach because the granularity was too coarse. All the diseases under N00 for renal disease and R57 for shock were not able to be mapped in this approach.

### Table 3

Comparison of drug indications to ICD-10 mapping approaches by precision, recall and F1 score (N = 375)

<table>
<thead>
<tr>
<th>Validation measure</th>
<th>Mapping approach (Mean% ± SD)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Direct</td>
<td>Indirect terminology-based</td>
</tr>
<tr>
<td>Precision</td>
<td>82.02 ± 27.16</td>
<td>81.87 ± 30.04</td>
</tr>
<tr>
<td>Recall</td>
<td>79.73 ± 25.02</td>
<td>86.78 ± 21.72</td>
</tr>
<tr>
<td>F1 score</td>
<td>76.63 ± 26.29</td>
<td>80.05 ± 28.17</td>
</tr>
</tbody>
</table>

### 5.6 The Limitations of the Study

#### 5.6.1 Unit of Analysis

In the direct approach, the mapping was done in each of the drug labels because the unit of analysis was a drug, and there was no mapping table which relates each indication to the set of ICD-10 codes. However, in the indirect terminology based approach, an indication was mapped to a set of ICD-10 codes and the set was applied to each drug more than once. The basic principle was multiple use of the same indication to ICD-10 codes mapping and the final validity check done by a unit drug might be negatively influenced by these multiple use.

#### 5.6.2 Terminology Problem

The performance of the indirect terminology based approach heavily depends on the density and quality of the terminology [2, 6]. If a term is not listed in the terminology, it will be omitted in mapping entire indications to ICD-10. During the study, we found two terms which have the same meaning but have different concept codes that were wrongly assigned in KOSTOM. In those cases we assigned the same ICD-10 code to each term and reported them to the organization which manages KOSTOM. Although the correction did not take much time or resources, the quality of the terminology should be improved.

#### 5.6.3 Extraction Problem

The intention of the study is to compare two mapping approaches and the mapping began with the extraction of the indications. However, the extraction was performed by two different organizations that we could not control. The difference in the extraction of indications would have effect on the result of the study as already described in the failure analysis.

#### 5.6.4 Significance of Difference in Validity Measure

The indirect terminology based approach showed better recall than the direct approach. However, the difference of the F1 score was just 3.42% with no significance.
In the application level, the difference is so small that both approaches will meet the requirements of the mapping, though the resource consumption is less in indirect terminology based approach.

6. Conclusions

The indirect terminology based approach showed better agreement and recall than those of the direct approach for mapping drug indications to ICD-10. At the application level, both approaches may be used because there was no difference of the F1 score in both approaches. However, in regards to consistency, time, and manpower, better results are expected from the use of the indirect terminology based approach.

Acknowledgments

This study was partially supported by Health Insurance Review and Assessment Service

References

16. NLM. Unified Medical Language System®.
19. Richesson RL. A process for achieving comparable data from heterogeneous databases (Disertation). Houston: The University of Texas School of Health Information Sciences; 2003.