

# Leveraging Lexical Matching and Ontological Alignment to Map SNOMED CT Surgical Procedures to ICD-10-PCS

Kin Wah Fung<sup>1</sup>, MD, MS, MA; Julia Xu<sup>1</sup>, MD, PhD; Filip Ameye, MD, FRCS(Eng), FACS; Arturo Romero Gutiérrez<sup>2</sup>, MD; Arabella D'Havé, MSc, BSc<sup>3</sup>

<sup>1</sup>National Library of Medicine, Bethesda, MD; <sup>2</sup>Ministry of Health, Social Services and Equality, Spain;

<sup>3</sup>Federal Public Service of Health, Food Chain Safety and Environment, Belgium

kwfung@nlm.nih.gov|xujc@mail.nlm.nih.gov|filip\_ameye@telenet.be|aromerog@msssi.es|arabella.dhave@gezondheid.belgie.be

## Abstract

*In 2015, ICD-10-PCS replaced ICD-9-CM for coding medical procedures in the U.S. We explored two methods to automatically map SNOMED CT surgical procedures to ICD-10-PCS. First, we used MetaMap to lexically map ICD-10-PCS index terms to SNOMED CT. Second, we made use of the axial structure of ICD-10-PCS and aligned them to defining attributes in SNOMED CT. Lexical mapping produced 45% of correct maps and 44% of broader maps. Ontological mappings were 40% correct and 5% broader. Both correct and broader maps will be useful in assisting mappers to create the map. When the two mapping methods agreed, the accuracy increased to 93%. Reviewing the MetaMap generated body part mappings and using additional information in the SNOMED CT names and definitions can lead to better results for the ontological map.*

## Introduction

After being used for over 30 years for reimbursement and other purposes in the U.S., the ICD-9-CM coding system was finally replaced by ICD-10-CM in October 2015. While procedures were included as part of ICD-9-CM (volume 3), they are not part of ICD-10-CM. To replace the ICD-9-CM procedure codes, a brand-new coding system called 'The International Classification of Diseases 10<sup>th</sup> Revision Procedure Coding System' (ICD-10-PCS) was created by the U.S. Centers for Medicare and Medicaid Services (CMS) through a contract with 3M Health Information Systems<sup>1</sup>. ICD-10-PCS is now used in the U.S. for the reporting of inpatient procedures. Other countries that have used ICD-9-CM are also transitioning to ICD-10-CM and ICD-10-PCS. Belgium made the transition in early 2015, and similar changes will happen in Spain and Portugal in the near future. In the realm of electronic health records (EHR), SNOMED CT has been designated as the clinical terminology standard. The Meaningful Use of EHR incentive of CMS specifies SNOMED CT as the terminology for the encoding of problem lists and procedures<sup>3</sup>. Many of the member countries of the International Health Terminology Standards Development Organisation (IHTSDO) are using, or have plans to use, SNOMED CT to document medical procedures. To facilitate the interoperability and integration of clinical and administrative data, it is evident that a map between SNOMED CT procedure codes and ICD-10-PCS will be necessary.

A project group was formed under IHTSDO in early 2015 with the goal to create a map between SNOMED CT and ICD-10-PCS.<sup>4</sup> The group includes representatives from Belgium, Spain, Portugal and U.S., together with medical terminology and informatics vendors. The map is expected to provide the following benefits:

- Improve ICD-10-PCS coding efficiency, consistency and accuracy
- Promote re-use of clinical data for administrative, epidemiologic and statistical purposes
- Facilitate integration of clinical and administrative data to support data analytics

Another potential benefit of the map is related to CPT (Current Procedure Terminology), another procedure terminology maintained by the American Medical Association (AMA). CPT is used in the U.S. for the billing of medical procedures and physician services. There is ongoing negotiation between IHTSDO and AMA to build a map between CPT and SNOMED CT. If both CPT and ICD-10-PCS are linked to SNOMED CT, SNOMED CT can become a lingua franca to facilitate interoperability between the two procedure coding systems.

Among the topics discussed by the project group are automated ways to map between the two terminologies. It is understood that the mappings generated by algorithmic methods will still need human validation, but reviewing candidate maps that are of reasonably high accuracy is easier and faster than creating a map from scratch. Among the methods considered, the group identified lexical matching of the ICD-10-PCS index terms and ontological alignment as among the most promising approaches of automatic mapping. This paper describes the two mapping methods and results of their evaluation.

### Lexical Mapping

Various methods of automatic mapping have been studied.<sup>5-10</sup> Generally, these are either lexical or semantic methods. Lexical methods rely on terms matching, while semantic methods depend on relationships. The axial structure of ICD-10-PCS poses special challenges and opportunities to automatic mapping.

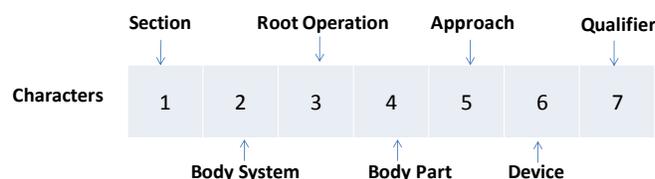
Traditional lexical mapping depends on finding matches between names of codes from two terminologies. String normalization can improve matching rate by reducing variations due to punctuations, upper/lower cases and lexical variants such as conjugation. However, direct lexical mapping between SNOMED CT and ICD-10-PCS names does not work well because of two reasons. First, ICD-10-PCS names are composite terms created by combining the values in the various axes, which are often different from terms used in clinical discourse. For example, the ICD-10-PCS name for laparoscopic appendectomy is *Resection (or Excision) of appendix, percutaneous endoscopic approach*, which will not match the common clinical term *Laparoscopic appendectomy* used in SNOMED CT. Second, there is mismatch in granularity between SNOMED CT and ICD-10-PCS. A full ICD-10-PCS code often includes information about approach (e.g., open approach, via natural or artificial opening), device (e.g., synthetic substitute, metallic joint prosthesis) and intent (e.g., diagnostic). In SNOMED CT, such additional information is only included in a minority of surgical procedures.

On the other hand, the ICD-10-PCS index is more promising for lexical matching because it contains many ‘clinician-friendly’ terms, such as:

- Duodenectomy**  
see Excision, Duodenum 0DB9  
see Resection, Duodenum 0DT9
- Duodenocholedochotomy** see Drainage, Gallbladder 0F94
- Duodenocystostomy**  
see Bypass, Gallbladder 0F14  
see Drainage, Gallbladder 0F94
- Duodenoenterostomy**  
see Bypass, Gastrointestinal System 0D1

### Ontological mapping

ICD-10-PCS is built on a multi-axial structure. The seven characters in the code correspond to the seven axes, each describing a particular aspect of the procedure.



This structure makes ICD-10-PCS expandable because new procedures can be easily incorporated as unique codes. In SNOMED CT, concepts are logically defined by attributes,<sup>11</sup> which is in some ways similar to an axial structure. SNOMED CT descends from predecessors that are based on an axial structure. The use of attribute-value pairs in definitions liberates SNOMED CT from the limitations of earlier axis-based definitions, but the compositional nature

of the terminology remains.<sup>12, 13</sup> By aligning the ICD-10-PCS axes and SNOMED CT attributes, it is possible to identify equivalence between the two terminologies (table 1). [ENREF 5](#)For example, laparoscopic appendectomy is defined in SNOMED CT by the attribute-value pairs: *Method = Excision, Procedure site – direct = Appendix structure* and *Using access device = Laparoscope*. By aligning the three attributes *Method, Procedure site - direct* and *Using access device* to *Root operation (axis-3), Body part (axis-4)* and *Approach (axis-5)* respectively in ICD-10-PCS, and matching the values for these attributes/axes, one can find the equivalent ICD-10-PCS codes *0DTJ4ZZ* or *0DBJ4ZZ* (figure 1).

SNOMED CT attribute	ICD-10-PCS axis	Axis no.
Procedure site	Body system	2
Method	Operation	3
Procedure site - Direct	Body part	4
Access	Approach	5
Surgical Approach	Approach	5
Using access device	Approach	5
Direct device	Device	6
Indirect device	Device	6
Procedure device	Device	6
Using device	Device	6
Procedure site - Indirect	Qualifier (Site)	7
Has intent	Qualifier (Intent)	7
Using substance	Qualifier (Substance)	7
Direct substance	Qualifier (Substance)	7

Table 1. Matching SNOMED CT procedure attributes to ICD-10-PCS axes

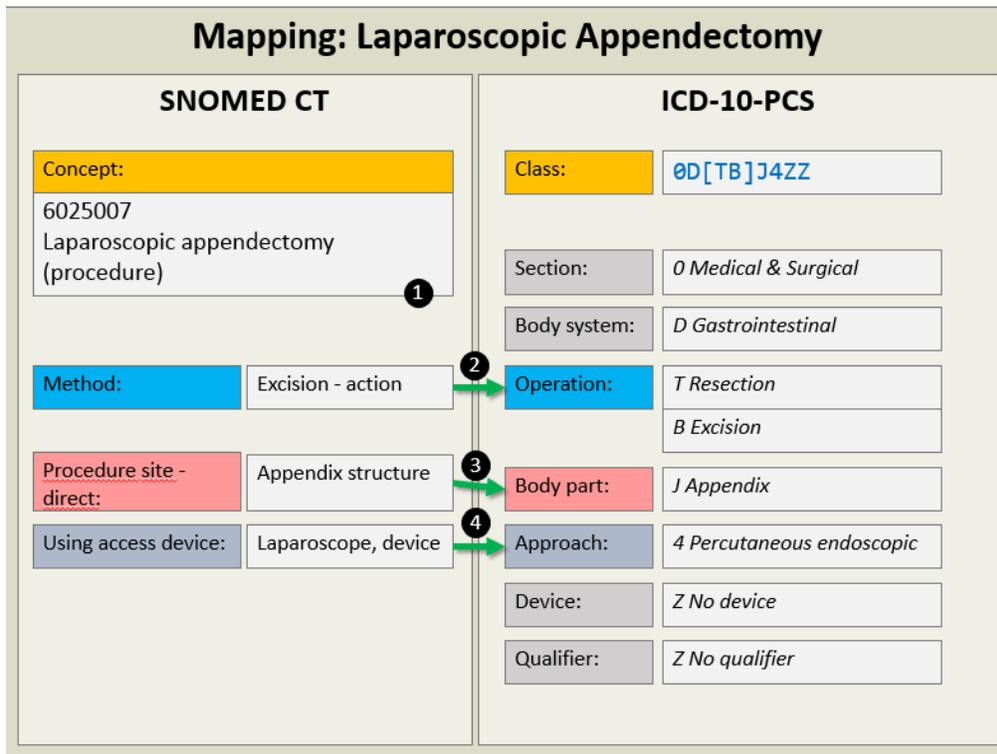


Figure 1. An example of ontological mapping, based on matching attributes and values at source and target systems. A source SNOMED CT procedure concept (1) has some attributes and values defining the concept; SNOMED CT attribute *Method* can be aligned to PCS target axis *Operation* (2); *Procedure site - direct* can be aligned with *Body part*(3); *Using access device* can be aligned to *Approach* (4)

## Methods

### *Creating the lexical and ontological maps*

From the 2016 version of the ICD-10-PCS Index xml file, we extracted main index terms and the associated codes listed with the main entries and sub-entries. We mapped the index terms to SNOMED CT concepts using NLM's MetaMap program (2015 version).<sup>14, 15</sup> Since the ICD-10-PCS index also included terms that were not procedures (e.g., body parts and devices), we restricted the SNOMED CT concepts to surgical procedures, defined as descendants of *Surgical procedure (387713003)*. We considered all ICD-10-PCS codes included under a main index term as candidate map targets for the SNOMED CT concept mapped to that index term.

For the ontological map, we first collected all the body parts values used in axis-4 of ICD-10-PCS. To this list, we added the inclusion terms, which were finer-grained terms included by the body part values (e.g., auditory ossicle includes incus, malleus and stapes), as defined by the ICD-10-PCS Definitions xml file. We mapped all body part terms to SNOMED CT using MetaMap, restricting the output to SNOMED CT body structure concepts. In this study, we focused on the ICD-10-PCS codes with excision or resection as the root operation. For all surgical procedures in SNOMED CT, we identified those with *Method = Excision*, and matched the *Procedure site – direct* value to body part terms in ICD-10-PCS. If there was more than one relationship group in the SNOMED CT definition, the body part match had to come from the same group as the excision method. We considered all ICD-10-PCS codes with root operation of resection or excision, and a matching body part value, to be candidate map targets for that SNOMED CT procedure.

### *Evaluation of the maps*

We selected a random sample of 100 concepts from each map, among which 50 concepts were common to both maps, to be reviewed by two of the authors (JX and FA). Each individual mapping was rated independently by each reviewer as:

- a. Correct, no better codes available
- b. Too specific, should use more general codes
- c. Too broad, should use more specific codes
- d. Related somewhat, not broader/narrower, better codes available
- e. Target codes unrelated to source concept

If the mapping was correct, the reviewer would also determine if all correct map targets had been identified. For each incorrect mapping, the reviewer would assign a reason for failure. Results from the reviewers were compared, and the differences were discussed until a consensus was obtained. If a consensus could not be reached, a third reviewer (KWF) would cast the deciding vote.

## Results

### *Lexical mapping*

A total of 2,557 unique main entries were extracted from the ICD-10-PCS index. MetaMap mapped the index entries to 569 unique SNOMED CT surgical procedures, which were in turn mapped to 5,679 unique ICD-10-PCS codes. The majority of the SNOMED CT concepts (95%) were mapped to four or less ICD-10-PCS codes, mostly at the 3 or 4 character level. A random subset of 100 SNOMED CT concepts was selected for review, excluding concepts with maps to over 20 ICD-10-PCS codes. In the review subset, 100 SNOMED CT concepts mapped to a total of 141 unique ICD-10-PCS codes (54 4-character and 87 3-character codes). Among them, 33 concepts mapped to only one ICD-10-PCS code, 52 concepts to two codes, with the maximum of nine codes (one concept). There were altogether 197 mappings (unique pairs of SNOMED CT and ICD-10-PCS codes) for review. Table 2 summarizes the review results for lexical mapping.

Map category	Mappings (%)	Reason of failure		
		Index granularity	MetaMap error	Index questionable
Correct	89 (45%)	NA	NA	NA
Too broad	87 (44%)	87	0	0
Related	12 (6%)	0	4	8
Unrelated	9 (5%)	0	9	0
Total	197 (100%)	87	13	8

Table 2. Results of lexical mapping

Overall, 45% of the maps were correct. Among the incorrect maps, the following reasons of failure were identified:

1. Index not sufficiently granular

The index pointed to a 3- or 4-character code, but the correct map was more specific. For example, the index for choledochectomy led to 0F[B/T], but the correct map should be 0F[B/T]9.

**Choledochectomy**

see Excision, Hepatobiliary System and Pancreas **0FB**

see Resection, Hepatobiliary System and Pancreas **0FT**

**0FB**

<i>Section</i>	0	Medical and Surgical		
<i>Body System</i>	F	Hepatobiliary System and Pancreas		
<i>Operation</i>	B	Excision: Cutting out or off, without replacement, a portion of a body part		
<i>Body Part</i>		<i>Approach</i>	<i>Device</i>	<i>Qualifier</i>
0 Liver		0 Open	Z No Device	X Diagnostic
1 Liver, Right Lobe		3 Percutaneous		Z No Qualifier
2 Liver, Left Lobe		4 Percutaneous Endoscopic		
4 Gallbladder				
G Pancreas				
5 Hepatic Duct, Right		0 Open	Z No Device	X Diagnostic
6 Hepatic Duct, Left		3 Percutaneous		Z No Qualifier
8 Cystic Duct		4 Percutaneous Endoscopic		
9 <b>Common Bile Duct</b>		7 Via Natural or Artificial		
C Ampulla of Vater		Opening		
D Pancreatic Duct		8 Via Natural or Artificial		
F Pancreatic Duct, Accessory		Opening Endoscopic		

2. MetaMap error

MetaMap mapped to the wrong SNOMED CT concept. For example, Osteoplasty was mapped to Bone fusion.

3. ICD-10-PCS index entry was questionable

In some cases, we think that the ICD-10-PCS index was not correct. For example, Colpocentesis pointed to drainage of the vagina (0U9G) while the correct map should be drainage of the cul-de-sac (0U9F). Another example is Gastroenterostomy, which pointed to either bypass or drainage. We think that drainage is not the correct root operation in this case since the procedure aims at altering the route of passage of the contents of the stomach to the small bowel, rather than letting out fluids from the stomach.

Looking at the results from the perspective of SNOMED CT concepts, 44% of the concepts had at least one correct map, and 96% had at least one correct or broader map. For those concepts with correct mappings, there were no missing map targets that the map failed to identify.

## Ontological mapping

There were altogether 19,658 surgical procedures in SNOMED CT, and each procedure was defined by at least one *Method* attribute, whose value was restricted to the *Surgical action* (129284003) sub-hierarchy. *Excision* was the most common method, followed by *Surgical action* and *Repair* (table 3).

Method attribute values	Number of concepts (% of all surgical procedures)
Excision - action (qualifier value)	3788 (19%)
Surgical action (qualifier value)	2356 (12%)
Repair - action (qualifier value)	2253 (11%)
Incision - action (qualifier value)	1467 (7%)
Closure - action (qualifier value)	1118 (6%)
Reconstruction - action (qualifier value)	726 (4%)
Surgical removal - action (qualifier value)	690 (4%)
Fixation - action (qualifier value)	680 (3%)
Destruction - action (qualifier value)	672 (3%)
Grafting - action (qualifier value)	483 (2%)
Other values	5425 (28%)
Total	19658 (100%)

Table 3. Distribution of methods in the definitions of SNOMED CT surgical procedures

Ontological mapping found maps for 1,746 SNOMED CT surgical procedures, mapping to 689 unique ICD-10-PCS codes (all 4-character). 95% of the concepts were each mapped to 10 or less ICD-10-PCS codes. The review subset consisted of 100 randomly selected concepts, excluding those mapping to over 20 ICD-10-PCS codes. Altogether, the 100 concepts mapped to 237 unique ICD-10-PCS codes. Most of the concepts (82 concepts) mapped to one to four ICD-10-PCS codes. There were a total of 362 mappings for review. The results are summarized in table 4.

Map category	Mappings (%)	Reason of failure					
		Body part mismatch	Missing approach	Resection vs excision	Operation mismatch	ICD-10-PCS inconsistent	Not classifiable
Correct	146 (40%)	NA	NA	NA	NA	NA	NA
Too specific	83 (23%)	83	0	0	0	0	0
Too broad	19 (5%)	0	19	0	0	0	0
Related	61 (17%)	15	0	25	17	2	2
Unrelated	53 (15%)	28	0	0	0	0	25
Total	362 (100%)	126	19	25	17	2	27

Table 4. Results of ontological mapping

Overall, 40% of the maps were correct. Among the incorrect maps, the following reasons of failure were revealed:

### 1. Body part mismatch

The map was incorrect due to the body part term in ICD-10-PCS being mapped to an incorrect body structure in SNOMED CT concept. In most cases, the SNOMED CT body structure was too broad, which resulted in the mapping of a broad SNOMED CT procedure to an overly-specific ICD-10-PCS code. For example, MetaMap mapped *Face artery*, *Thyroid artery* and *Hand artery* in ICD-10-PCS to *Arterial structure* in SNOMED CT. As a result, *Excision of artery* in SNOMED CT was mapped to *Excision of face artery*, *Excision of thyroid artery* and *Excision of hand artery* in ICD-10-PCS.

## 2. Missing approach information

In this study, we did not make use of information about surgical approach for mapping. Some SNOMED CT concepts containing approach information were therefore mapped to ICD-10-PCS codes that were too broad. For example, *Laparoscopic adrenalectomy* in SNOMED CT was mapped to excision/resection of left/right adrenal gland (e.g., 0GB2 – *Excision of left adrenal gland*), while the correct map target should include the approach (e.g., 0GB24 – *Excision, left adrenal gland, percutaneous endoscopic approach*). Other examples included *Transurethral prostatectomy* and *Thoracoscopic excision of neoplasm of pericardium*.

## 3. Excision vs. resection

While SNOMED CT considered excision and resection as synonymous, they were considered different root operations in ICD-10-PCS. In ICD-10-PCS, excision was defined as “the cutting out or off, without replacement, **a portion of** a body part”; while resection was “the cutting out or off, without replacement, **all of** a body part”. Generally, the correct mapping of a SNOMED CT excisional procedure should include both the excision and resection options in ICD-10-PCS. However, in cases where the SNOMED CT concept referred specifically to a partial or complete removal, a more specific ICD-10-PCS code should be used. For example, *Excision of focal lesion of larynx* should only map to *Excision of larynx (0CBS)* and not *Resection of larynx (0CTS)*. On the other hand, *Total thymectomy* should map to *Resection of thymus (07TM)* but not *Excision of thymus (07BM)*.

## 4. Root operation mismatch

In SNOMED CT, excision is used broadly to define procedures that involve removal or destruction of a body part, regardless of the method used. In ICD-10-PCS, the codes may be different depending on the method and what is being removed. For example, *Chemothalamectomy* was a destruction operation in ICD-10-PCS. *Vasectomy using silicon plug* would be coded as occlusion. *Removal of sequestra from bone* was an extirpation procedure.

## 5. ICD-10-PCS inconsistent

There was one case in which the reviewers found conflicting information in the ICD-10-PCS index. While the entry Cricoidectomy pointed to the excision of larynx, the entry for Cricoid cartilage suggested to use trachea as the body part.

**Cricoidectomy** see Excision, Larynx 0CBS

**Cricoid cartilage** use Trachea

## 6. Unclassifiable

There were three SNOMED CT concepts that we considered not classifiable in ICD-10-PCS. *Excision of accessory or ectopic lung tissue* was not classifiable because the location of the ectopic lung tissue was not known (could be in the thorax, abdomen or subcutaneous). *Excision of benign lesion of trunk* was not classifiable because trunk was not a defined body region in ICD-10-PCS. *Laparoscopy with excision of lesion* was too broad to be meaningfully classified.

Looking at the results from the perspective of SNOMED CT concepts, 68% of SNOMED CT concepts had at least one correct map, and 78% had at least one correct or broader map. Among the SNOMED CT concepts with correct maps, the map missed some target codes in four concepts because the SNOMED CT concept contained more than one procedure. One example is *Anterior resection of rectum with colostomy* in which the target code for colostomy was missed.

### Comparing the two maps

Half of the reviewed concepts were common to both maps by design, so that we could compare the two mapping methods. For the 50 concepts in common, there were 101 lexical and 182 ontological mappings. We analyzed the mappings by whether they had: a. exact same target in the other map b. matching map target in the other map at the 3-character level c. no matching map targets (table 5).

		Number of mappings	Correct map %	Broader map %	Correct or broader map %
Lexical Map (LM)	Same target as OM	29	93%	7%	100%
	Match OM target at 3-character level	53	19%	79%	98%
	No match	19	58%	37%	95%
	Total	101	50%	49%	98%
Ontological Map (OM)	Same target as LM	29	93%	7%	100%
	Match LM target at 3-character level	116	50%	0%	50%
	No match	37	14%	0%	14%
	Total	182	50%	1%	51%

Table 5. Accuracy of maps for concepts common to both mapping methods (OM: ontological map, LM: lexical map)

Overall, the two reviewers agreed in 50% of the mappings after the first round of individual assessment. Most differences were resolved after discussion, and only 10% of mappings required the third reviewer's vote to decide.

### Discussion

Creating a map between two terminologies is labor intensive and time consuming. Mapping is not an exact science. Due to the differences in scope, granularity and organizing principles between the source and target terminologies, the mapping experts sometimes have to rely on their judgment, which can result in considerable variability. Common methods to ensure consistency include dual independent mapping (each map created by two mappers independently) and sequential review (review of some or all maps by a second mapper). Automated methods to identify candidate maps can help by suggesting map targets to mappers, which can save time and improve consistency. The candidate maps can also act as an independent source to corroborate the manually-created maps. Our previous study shows that algorithmically created maps can lead to considerable saving in mapping time.<sup>16</sup>

In our study, 45% of the maps generated by lexical mapping of the ICD-10-PCS index are correct. While this level of precision is not particularly impressive on its own, an additional 44% of the maps are broader maps at the 3- or 4-character level. In the use case of assisting manual mapping, the broader maps are also useful because they lead to the correct table, from which it is relatively simple to pick the correct target code. So we can say that 89% of the maps are useful in this context. Overall, 96% of SNOMED CT concepts which can be mapped in this way will have a useful map. However, the main drawback of this method is the limited coverage, as the index of ICD-10-PCS is not very extensive. Only 569 SNOMED CT surgical procedures can be mapped by this method.

On the other hand, ontological mapping can potentially cover many more SNOMED CT concepts. We are able to find maps for 46% (1,746 out of 3,788) of the SNOMED CT excision procedures. The precision of the ontological map (40%) is slightly lower than the lexical map, and the proportion of broader maps is also lower (5%). Overall, the ontological map is able to find at least one correct or broader map for 78% of SNOMED CT concepts.

Analyzing the reasons of failure enables us to discover ways of improving the performance of the maps. The main reason of failure in the lexical map is insufficient granularity of the codes listed in the index, most of them only at the table (3-character) level. There is no way to get around this at present, unless the ICD-10-PCS developers decide to add new entries or use more granular codes in the index in future. For the ontological map, there are some possibilities for improvement. Body part mismatch makes up the majority of failures. Our method depends solely on the output of MetaMap without manual review. Since the total number of body part terms in ICD-10-PCS is relatively small (less than 800), it is feasible to manually review and correct the maps found by MetaMap, which could improve the overall performance. Another possible improvement is to make use of the approach attributes in SNOMED CT (*Access*, *Surgical approach* and *Using access device*) to match to axis-5 (*approach*) in ICD-10-PCS. This will lead to better maps when the SNOMED CT concept includes approach information. For example, *Laparoscopic adrenalectomy* in SNOMED CT has attribute *Use access device = Laparoscope*, which can be matched to *Percutaneous Endoscopic Approach* in ICD-10-PCS. Furthermore, the mismatch of excision (removal of part of a body part) and resection (removal of all of a body part) between SNOMED CT and ICD-10-PCS can potentially be resolved by clues in the name of the SNOMED CT concept. For example, SNOMED CT procedures with ‘complete’, ‘total’, ‘entire’ or ‘radical’ in their names can be mapped only to resection, while those with ‘partial’, ‘segment’, ‘wedge’, ‘biopsy’ or the pattern ‘excision of x of y’ (e.g., *Excision of cyst of vulva*) can be mapped only to excision. In the subset that we reviewed, many of the incorrect mappings can be avoided in this way.

Combining the two maps offers an additional way to use the maps. Maps that agree exactly are much more likely to be correct (93%) than each map alone. The overall performance of the ontological map is just fair because only 45% of the mappings are potentially useful (correct or broader maps). If we restrict to the subset of the ontological map that is corroborated by the lexical map (same map target or agreeing at the 3-character level), the proportion of useful maps will rise to 60%.

In future, we will explore ways to improve the ontological map by curating the body part matches and making use of other ICD-10-PCS axes and SNOMED CT defining attributes. We will also study the possibility of first mapping ICD-9-CM procedure codes to SNOMED CT, and then using the General Equivalence Maps (GEM) published by CMS<sup>1</sup> [ENREF 15](#) to map from ICD-9-CM to ICD-10-PCS. The ICD-9-CM volume 3 index is another potential resource that can enhance lexical mapping, since the ICD-9-CM index is more comprehensive than the ICD-10-PCS index.

We recognize the following limitations in our study. We only focused on surgical operations but the scope of ICD-10-PCS includes other types of medical procedures (e.g., obstetrical procedures, imaging studies). For the ontological map, we only studied excision procedures, and the results may not be generalizable to other types of operations.

## **Conclusion**

We explored two automatic ways of mapping SNOMED CT surgical procedures to ICD-10-PCS: lexical mapping of the ICD-10-PCS index, and ontological mapping by aligning SNOMED CT attributes and ICD-10-PCS axes. Lexical mapping yielded 89% of useful maps (correct or broader) but the coverage of SNOMED CT was limited. Ontological mapping had higher coverage but only 45% of the maps were useful. The performance of the ontological maps can potentially be improved by refining the body part matches and making use of additional information in the SNOMED CT names and defining attributes.

## **Acknowledgements**

This work was supported in part by the Intramural Research Program of the National Institutes of Health and the National Library of Medicine.

## **References**

Fung K, Xu J, D'Havé A, Ameye F, Gutiérrez A. Leveraging lexical matching and ontological alignment to map SNOMED CT surgical procedures to ICD-10-PCS. *AMIA Annu Symp Proc* 2016. 2016:570-9.

1. US Centers for Medicare & Medicaid Services ICD-10-PCS web page. Available from: <https://www.cms.gov/Medicare/Coding/ICD10/2016-ICD-10-PCS-and-GEMs.html>.
2. Averill RF, Mullin RL, Steinbeck BA, Goldfield NI, Grant TM. Development of the ICD-10 procedure coding system (ICD-10-PCS). *Topics in health information management*. 2001;21(3):54-88. Epub 2001/03/10.
3. Health Information Technology Certification Criteria, U.S. Health and Human Services Department on 10/16/2015. Available from: <https://www.federalregister.gov/articles/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base>.
4. Fung K, Xu J, D'Havé A, Ameye F, Gutiérrez A, Santos A. Developing a Map from SNOMED CT Procedure Concepts to ICD-10-PCS. *SNOMED CT Expo 2015*; Montevideo, Uruguay.
5. Cimino JJ, Barnett GO. Automated translation between medical terminologies using semantic definitions. *MD Comput*. 1990;7(2):104-9.
6. Dolin RH, Huff SM, Rocha RA, Spackman KA, Campbell KE. Evaluation of a "lexically assign, logically refine" strategy for semi-automated integration of overlapping terminologies. *J Am Med Inform Assoc*. 1998;5(2):203-13. Epub 1998/04/03.
7. Fung KW, Bodenreider O. Utilizing the UMLS for semantic mapping between terminologies. *AMIA Annu Symp Proc*. 2005:266-70.
8. Fung KW, Bodenreider O, Aronson AR, Hole WT, Srinivasan S. Combining lexical and semantic methods of inter-terminology mapping using the UMLS. *Medinfo*. 2007;12(Pt 1):605-9. Epub 2007/10/04.
9. Rocha RA, Huff SM. Using digrams to map controlled medical vocabularies. *Proc Annu Symp Comput Appl Med Care*. 1994:172-6. Epub 1994/01/01.
10. Rocha RA, Rocha BH, Huff SM. Automated translation between medical vocabularies using a frame-based interlingua. *Proc Annu Symp Comput Appl Med Care*. 1993:690-4.
11. SNOMED CT Editorial Guide, IHTSDO. Available from: [http://ihtsdo.org/fileadmin/user\\_upload/doc/en\\_us/eg.html](http://ihtsdo.org/fileadmin/user_upload/doc/en_us/eg.html).
12. Spackman KA, Campbell KE. Compositional concept representation using SNOMED: towards further convergence of clinical terminologies. *Proc AMIA Symp*. 1998:740-4.
13. Dolin RH, Spackman K, Abilla A, Correia C, Goldberg B, Konicek D, et al. The SNOMED RT Procedure Model. *Proc AMIA Symp*. 2001:139-43.
14. Aronson AR, Lang FM. An overview of MetaMap: historical perspective and recent advances. *J Am Med Inform Assoc*. 2010;17(3):229-36. Epub 2010/05/06.
15. MetaMap - A Tool For Recognizing UMLS Concepts in Text. Available from: <https://metamap.nlm.nih.gov/>.
16. Fung KW, Xu J. Synergism between the mapping projects from SNOMED CT to ICD-10 and ICD-10-CM. *AMIA Annu Symp Proc*. 2012;2012:218-27. Epub 2013/01/11.