

Algorithmic Approach to Mapping SNOMED CT to ICD-10

Kathy Giannangelo, IHTSDO

Brian Carlsen, West Coast Informatics, LLC

Delivering

SNOMED CT

The global
language of
healthcare

Objectives

- Describe the characteristics of a mapping algorithm
- State the different types of mapping algorithms
- Explain how algorithms may be used to create a map from SNOMED CT to ICD-10
- Identify the benefits and challenges to the algorithmic approach

Background

- Over 66,000 maps to create within a short period of time
- Availability of third party knowledge resources
- Manual process had resulted in the development of mapping principles

Process Employed

- Concepts
 - Mapped by the algorithm
 - Matching
 - Batching
 - Targeted
 - Mapped by a human map specialist with access to
 - Information about all candidate codes and their ranking produced by the algorithm
 - Mapping principle applied
 - IHTSDO Mapping Service Team

Types of Algorithms

MATCHING
BATCHING
TARGETED

Matching

- Evidence and threshold based approach
 - Gather evidence from a number of sources to determine a list of potential candidate codes

#	Description	Strength
1	<u>SNOMED CT</u> to ICD10CM <u>are</u> connected via the mappings developed by <u>NLM</u> and the ICD10CM code can be algorithmically matched to an ICD10 code.	VERY STRONG
2a	<u>SNOMED CT</u> and ICD10 share the same concept in <u>UMLS/NCI-META</u>	STRONG
2b	<u>SNOMED CT</u> and ICD10CM share the same concept in <u>UMLS/NCI-META</u> or <u>UMLS</u> and the ICD10CM code can be <u>algorithmically matched</u> to an ICD10 code.	STRONG

- Scoring
 - Weigh candidate target codes based on strength of evidence and assign “points”

Matching

- Lexical approach
 - Compare descriptions in SNOMED CT to various lexical constructions from ICD-10
 - Preferred names
 - Inclusions
 - Index entries
 - Cleaning of ICD-10 data
 - Creating fully qualified values for index and inclusion terms
 - Development of a synonym file
 - Translation between expressions in SNOMED CT and those in ICD-10
 - Abrasion => friction burn
 - Lucene scoring method

Matching

- Scoring of all candidates generated from both approaches
- Apply threshold conditions to determine which candidate had sufficient strength to be the suggested candidate target code
- Apply other rules

Batching

- Goal: Present map specialists with similar content grouped together
 - Similar mapping principles/coding rules
- Early approach based around analysis of the strength and focus of available evidence
 - Easy/Medium/Hard
- Revised approach based on the algorithmic analysis of the remaining concepts using four key principles

Batching

- Pool of concepts placed into 14 categories/subcategories

#	Name	Description
01	Easy	First pass through algorithm and its <u>judgement</u> that there was very strong evidence for only a single code. This was the initial batch that was edited.
02	Not Classifiable	<p>Based list of <u>SNOMED CT</u> concepts for which the algorithm was unable to determine any possible mapping to ICD10 based on rules of evidence. Final list of concepts to include was determined by human lexical review.</p> <p>This analysis determined that certain lexical patterns were good clues to likely <u>unmappable</u> cases. For example, “able to” or “normal” are things that do not exist in ICD10 because it codes for deficiencies, inabilities, and abnormalities (e.g. problems). So a finding that was a lack of a problem was a likely <u>unmappable</u> case. Other similar approaches to lexical analysis were used to keep cases from winding up in other categories.</p>

Batching

- Effective way to partition the concepts
 - Focus on the same types of concepts
 - Able to keep the same mapping rules in mind
- Enabled assigning of content to specialist with knowledge in a particular content area
 - Intended to increase productivity and consistency of mapping

Targeted

- Identify rules
 - Identify candidate map principles
- Modify rules
 - Modify target codes and advice after editing
- Flag for QA rules
 - Flag for QA review concepts that violated known conditions

Identify Rules

- Used to categorize concepts into different groups and/or predict whether a particular mapping principle applies
- Applied when concept assigned to a specialist
- Example:
 - Description: Identify allergy and sensitivity cases
 - Predicate:
 - Lexical (case insensitive non-word-boundary) match to SNOMED CT PT or FN on the following words:
 - Allergy
 - Allergic
 - Allergies
 - Sensitivity
 - Anaphylaxis
 - Anaphylactic
 - AND not a (case insensitive non-word-boundary) match to “pseudo”
 - Action: Label/categorize concept as a potential PRINCIPLE 19 case

Modify Rules

- Used to add (or remove) advice to map record entries after completion of editing or changing a map target code
- Applied when a specialist completes editing
- Example:
 - Description: When SNOMED fracture concept does not indicate open or closed, WHO guidance is to code to “closed” with the advice “MAPPED FOLLOWING WHO GUIDANCE”.
 - Predicate:
 - Map entry has S code target id and the SNOMED PT or FN contain the (case insensitive) word “fracture” and the SNOMED PT and FN do not have a (case-insensitive, word boundary) match to either the word “open” or the word “closed”
 - Action: Add “MAPPED FOLLOWING WHO GUIDANCE” advice to the map entry

Flag for QA

- Used as real-time validation checks to
 - To prevent completion of a map record or
 - As post-editing QA checks to identify conditions among the pool of ready for publication concepts
- Example
 - Description: When SNOMED does not indicate open or closed, map records should not be coded to an “open” 5th digit code.
 - Predicate:
 - Map entry has S code target id and the SNOMED PT and FN do not have a (case-insensitive, word boundary) match to the word “open” and the 5th digit is coded to 1 (open)
 - Action: Validation rule – prevent finishing of this mapping

Challenges

- Mapping Principles
 - Development of predicate identified areas in need of clarification
- Classification rules
 - Algorithm not possible due to incomplete/vague guidance
- Concepts requiring multiple target codes
 - Under certain circumstances, effective dual coding could be algorithmically done (e.g. poisoning cases).

Benefits

- Leverage work
 - Enhance IHTSDO mapping tool functionality
 - Batching
 - Initial algorithmic map target generation
 - Targeted algorithms
 - Identify
 - Modify
 - Flag for QA
 - Applicable to new map projects involving other terminologies

Exploration of Impact

- Quality improvements to the SNOMED CT to ICD-10 map
 - Consistent application of mapping principles
 - Enforce map principles at the time of mapping
 - Find and re-map older maps that predate current principles.
 - Increased reproducibility of target code

The logo for ihtsdo, featuring the lowercase letters 'ihtsdo' in a white, sans-serif font centered within a solid blue square.

Leading healthcare
terminology, worldwide

Contact Information

KGI@IHTSDO.ORG

BCARLSEN@WESTCOASTINFORMATICS.COM