Building Continents of Knowledge in Oceans of Data: The Future of Co-Created eHealth
A. Ugon et al. (Eds.)

doi:10.3233/978-1-61499-852-5-900

Clarifying Diagnoses to Laymen by Employing the SNOMED CT Hierarchy

Hugo J. Th. VAN MENS[a,b,1], Nicolette F. DE KEIZER[a], Remko NIENHUIS[b], and Ronald CORNET[a]
[a]*Department of Medical Informatics, Amsterdam Public Health research institute, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands.*
[b]*Research and Development, ChipSoft B.V., Amsterdam, The Netherlands*

Abstract. Patient access to electronic health records (EHRs) is associated with improved efficiency, self-management, and patient engagement. However, the EHR contains medical language that can be difficult to comprehend by patients. In Dutch hospitals, the Diagnosethesaurus (DT) is used as an interface terminology to register diagnoses, but it does not contain patient-friendly terms. Fortunately, the DT is partly mapped to SNOMED CT and there is a proportionately small set of patient-friendly terms available in the Dutch SNOMED CT release. The purpose of this study was, therefore, to investigate if SNOMED CT can be used to generate clarifications of diagnoses for patients. Only 1.2% of the DT diagnoses that were already mapped to SNOMED CT had patient-friendly synonyms that were different from the diagnoses descriptions. However, by generalizing diagnoses to SNOMED CT concepts with patient-friendly terms, this number could be increased to 71%. In conclusion, we showed that a high percentage of diagnoses could be clarified to at least some extent with the relatively small set of patient-friendly terms. Future research will involve the further optimization of the clarifications, and evaluation with clinicians and patients.

Keywords. Diagnoses, health literacy, patient access to records, patient-friendly terminology, personal health records, SNOMED CT

1. Introduction

There is an increased attention to patient access to electronic health records (EHRs), because of associated benefits such as improved patient satisfaction, patient-provider communication, patient engagement, self-management, efficiency, and patient safety. [1,2] However, EHRs contain medical language that can be difficult to comprehend by patients. [3-5] A patient-friendly terminology could help patients to better understand their medical records. [4,5]

Zeng et al. [6,7] showed that the Unified Medical Language System (UMLS) and the Consumer Health Vocabulary (CHV) could be used to find synonyms, and generate explanations of medical concepts for laymen. However, the UMLS is not available for the Dutch medical language. Fortunately, the National Release Center of SNOMED CT in the Netherlands published the SNOMED CT Patient-Friendly Extension Release [8] (PFE) with 301 patient-friendly terms to describe 288 SNOMED CT concepts. This

[1] Corresponding author: Hugo J. Th. van Mens, J1B-109, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands; E-mail: h.j.vanmens@amc.uva.nl

small initial set was based on the terms that can be found on a website of the Dutch patient federation where patient experiences about healthcare provider can be shared. [9] SNOMED CT and the PFE can be used to clarify diagnoses in the same way as the UMLS and the CHV. Furthermore, the Diagnosethesaurus (DT) [10] is becoming the standard interface terminology to register diagnoses at the point of care in Dutch hospitals, and this interface terminology is currently partly mapped to SNOMED CT (the full mapping was not yet finished at the time of writing). Therefore this provides opportunities to make descriptions of diagnoses more comprehensible for patients.

A diagnosis can be clarified by providing a patient-friendly synonym or by generalizing it to one or more generic concepts that can be described by a patient-friendly term. [6,7] Table 1 illustrates these clarifications methods. For example, "agranulocytosis is a type of immune system disorder" is a generalization of a diagnosis to a more general concept, where the diagnosis is called the subtype and the more general concept the supertype. UMLS and SNOMED CT both contain hierarchical relationships that can be used for this purpose. There are also other relationships in these terminology systems that can be used to clarify medical concepts, such as the "finding site" and "part of" relationship. This enables clarifications such as "aortic valve is a part of the heart". These latter relationships were also used in the studies with the UMLS, [6,7] but were out of the scope of the current study, where we focused on the hierarchical subtype-supertype relationship.

The purpose of this study was thus to investigate if the PFE and the SNOMED CT hierarchy can be used to generate Dutch clarifications of diagnoses for patients.

Table 1. Three methods to clarify diagnoses illustrated by an example (freely translated)

Clarification method	Example of the clarification method
Synonym	Trigeminal neuralgia is another word for facial pain.
Generalization to one concept	Agranulocytosis is an immune system disorder.
Generalization to multiple concepts	Papillon-Lefèvre syndrome is a heritable disorder of bone, tooth, and skin.

2. Methods

We combined the DT (version of June 2017) with the SNOMED CT Netherlands edition and the PFE (version of March 2017) and investigated how many of the diagnoses can be clarified with how many patient-friendly terms. When a diagnosis was mapped to a SNOMED CT concept that had a patient-friendly term in the PFE, this term was provided as the synonym of the diagnosis. We verified whether this term was actually different from the diagnosis description, using text matching.

The SNOMED CT hierarchy was used to determine if a diagnosis is a subtype of one or more concepts that can be described by a patient-friendly term (see Figure 1). In order to find the supertypes of a concept in SNOMED CT, the transitive closure table was used. [11] This table contains all subtype-supertype relationships. This way all supertypes with patient-friendly terms could be found for each diagnosis. However, this might provide redundant explanatory terms, because supertypes of the supertypes would be found as well. For example, "Aortic valve stenosis", is a "Heart valve disorder", "Heart disease", and "Disorder of cardiovascular system". In these cases, we only used the most specific generalizations. This would simplify the clarification to "Aortic valve

stenosis is a heart valve disorder". In case a diagnosis had multiple supertypes, as is illustrated with "Telangiectasia macularis eruptiva perstans" which is an "immune system disorder", "skin disorder" as well as a "disorder of cardiovascular system", the most specific generalizations were used in combination. We counted the number of unique combinations of these generalizations, i.e., the number of unique clarifications that could be generated with subtype-supertype relationships.

When a supertype was contained in the diagnosis description, it was not regarded to add any new information. For example, that "Salmonella infection" is an "infectious disease" is already implied by the word "infection" in "Salmonella infection". For this reason, we disregarded supertypes that are not informative and filtered them out by text matching, e.g. matching on "infect" to match both "infection" as well as "infectious" when the supertype was "infectious disease".

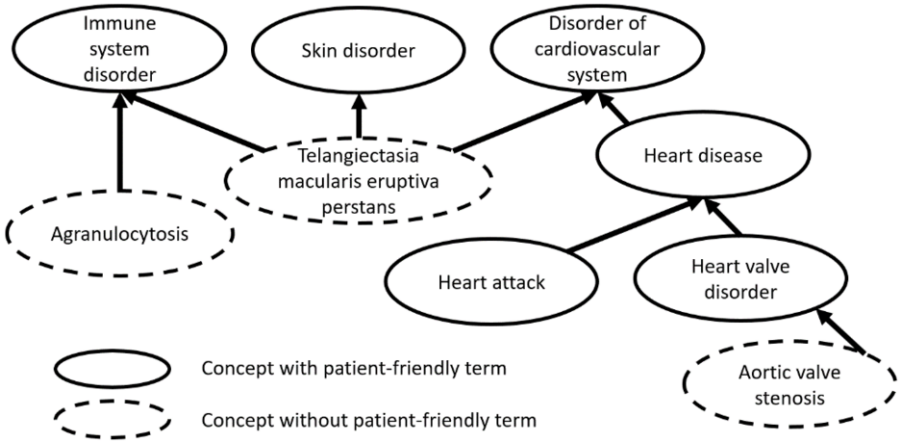

Figure 1. Simplified example of SNOMED CT concepts in the SNOMED CT hierarchy

3. Results

Of the 21,426 diagnoses in the DT, only 12,453 diagnoses (58.1%) had already been mapped to SNOMED CT. In total 288 SNOMED CT concepts had a patient-friendly term in the PFE, and these described 225 of the diagnoses with a SNOMED CT mapping. 75 diagnoses descriptions were already equal to the patient-friendly terms, hence, the diagnosis description was actually different from the patient-friendly term for 150 diagnoses. Therefore the patient-friendly terms could be used to clarify 1.20% of all diagnoses that have a SNOMED CT mapping with a synonym.

By generalizing diagnoses to supertypes with patient-friendly terms, 8,797 diagnoses (70.6% of all that are mapped to SNOMED CT) could potentially be clarified with 211 concepts that have patient-friendly terms. These diagnoses were described by 1 to 6 different supertypes; as shown in Table 2. The 211 supertypes formed 735 unique combinations of supertypes that can be used for clarifications.

There were 110 diagnoses with a patient-friendly term that could also be clarified using supertypes. Analogously, 128 patient-friendly terms were also used as a supertype. See Figure 2. The total numbers are thus 40 + 110 + 8,687 = 8,837 diagnoses and 22 +

128 + 83 = 233 concepts with patient-friendly terms. As a result, 8,837 diagnoses (71.0% of all mapped diagnoses) could be clarified with a synonym and/or supertypes using 233 concepts with patient-friendly terms and 735 unique combinations these terms.

Table 2. Number of diagnoses that have a certain number of patient-friendly supertypes to clarify the diagnoses (excluding the redundant ones) and the number of unique clarifications with a certain number of supertypes

Supertypes	Diagnoses	Unique clarifications
1	6,329	192
2	2,050	376
3	366	132
4	49	32
5	2	2
6	1	1
Total	**8,797**	**735**

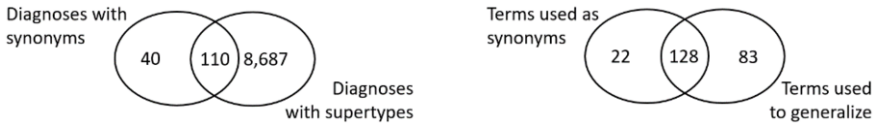

Figure 2. The overlap between diagnoses that can be clarified with synonyms and supertypes, and the overlap between patient-friendly terms used as synonyms or to generalize, and their numbers.

4. Discussion

With the relatively small set of SNOMED CT concepts that have one or more patient-friendly terms in the PFE, a high percentage of diagnoses could be clarified to at least some extent, by using hierarchical subtype-supertype relationships of SNOMED CT, and by providing synonyms.

The DT was not completely mapped to SNOMED CT yet, but this is expected to be completed in coming releases. [10] We will have to repeat the analysis from this paper after this is finished. The earlier studies [6,7] did not mention how much content of the terminology systems used could be clarified with their system, which makes it difficult to compare their results with those of our study.

The PFE is quite small compared to the number of diagnoses in the DT, and, particularly, compared to the number of concepts in SNOMED CT. Perhaps some concepts that could be informative to patients, or could clarify the 29% of the diagnoses that were not clarified by the method presented in this study, could relatively easily be translated to patient-friendly terms to obtain even better results. However, finding patient-friendly synonyms for *all* diagnoses and other types of medical concepts is a costly process and might not be possible for some complex diagnoses. Compared to English language, only limited tools are available to perform natural language processing for the Dutch language as well as many other languages. [12] Therefore our result is of particular interest, showing that text simplification might be feasible with a relatively small set of patient-friendly terms. We believe that this method could be applied to other languages.

Whether the clarifications will increase the comprehension of patients and other users, such as caregivers, remains to be evaluated. The clarifications are not yet validated by clinicians, although it can be assumed the hierarchical relationships are correct, because SNOMED CT is clinically validated. We want to first improve the clarification method by utilizing other types of relationships in SNOMED CT, such as the "finding site" relationship or "associated morphology" relationship. We expect this might be more useful for certain diagnoses, such as malignant neoplasms and infections, where the clarification could be e.g. that it is a form of cancer (associated morphology) with a certain finding site. The method presented in this paper can also be used to determine which concepts could be translated to patient-friendly terms to result in a maximum increase of the number of concepts that could potentially be clarified.

5. Conclusion

We showed that a relatively small patient-friendly terminology and the SNOMED CT hierarchy can be used to generate clarifications of a large proportion of diagnoses for patients. Future research will involve the further optimization of the clarifications, the utilization of other types of relationships, and evaluation with clinicians and patients. Additionally, research should focus on which parts of SNOMED CT could be translated to patient-friendly terms in order to clarify the highest number of concepts in the most comprehensive manner.

References

[1] C.S. Kruse, K. Bolton, and G. Freriks, The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review, *J Med Internet Res* **17** (2015), e44.
[2] F. Mold, S. de Lusignan, A. Sheikh, A. Majeed, J.C. Wyatt, T. Quinn, M. Cavill, C. Franco, U. Chauhan, and H. Blakey, Patients' online access to their electronic health records and linked online services: a systematic review in primary care, *Br J Gen Pract* **65** (2015), e141-e151.
[3] T. Irizarry, A. DeVito Dabbs, and C.R. Curran, Patient Portals and Patient Engagement: A State of the Science Review, *J Med Internet Res* **17** (2015), e148.
[4] R.A. Bush, C.D. Connelly, M. Fuller, and A. Perez, Implementation of the Integrated Electronic Patient Portal in the Pediatric Population: A Systematic Review, *Telemed J E Health* (2015).
[5] B. Hemsley, M. Rollo, A. Georgiou, S. Balandin, and S. Hill, The health literacy demands of electronic personal health records (e-PHRs): An integrative review to inform future inclusive research, *Patient Educ Couns* (2017).
[6] Q. Zeng-Treitler, S. Goryachev, H. Kim, A. Keselman, and D. Rosendale, Making texts in electronic health records comprehensible to consumers: a prototype translator, *AMIA Annu Symp Proc* (2007), 846-850.
[7] S. Kandula, D. Curtis, and Q. Zeng-Treitler, A Semantic and Syntactic Text Simplification Tool for Health Content, *AMIA Annu Symp Proc* (2010), 366-370.
[8] SNOMED CT Netherlands/Nederlands Edition, *SNOMED CT National Release Center Netherlands*, March 31, 2017, https://mlds.ihtsdotools.org/.
[9] Zorgkaart Nederland, *Patiëntenfederatie*, https://www.zorgkaartnederland.nl/.
[10] Diagnosethesaurus, Version 2.5, June 22, 2017, *Dutch Hospital Data*. https://www.dhd.nl/.
[11] SNOMED International, Terminology Services Guide, *SNOMED International*, https://confluence.ihtsdotools.org/display/DOCTSG/.
[12] R. Cornet, A. Van Eldik, and N. De Keizer, Inventory of tools for Dutch clinical language processing, *Stud Health Technol Inform* **180** (2012), 245-249.