

A UMLS-based Investigation of Laterality in Biomedical Terminologies

Rashmie Abeysinghe, PhD¹, Xubing Hao², Licong Cui, PhD², Guo-Qiang Zhang, PhD^{1,2,*}

¹Department of Neurology, The University of Texas Health Science Center at Houston, Houston, TX

²School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX

Abstract

Laterality is an important anatomic directional property indicating the sidedness of body structures, diseases, and procedures. Errors in laterality could have catastrophic consequences in patient care. In this paper, we investigate how different biomedical terminologies organize terms indicating laterality. We leverage the Unified Medical Language System (UMLS) to identify lateral terms in different terminologies. For each lateral term, we attempt to obtain other matched lateral terms and further analyze how they are interrelated. Our results indicated that only 1.68% of the matched lateral term-pairs are hierarchically related. It was also seen that 44.24% of matched-pairs were siblings. We found that in SNOMED CT, bilateral concepts were hierarchically related to both left and right lateral concepts different to most other terminologies. Further investigation revealed that the likely causes for these relations are how the logical definitions of SNOMED CT concepts are arranged.

1 Introduction

“Digitization” is improving every facet of life, including healthcare. A side effect of digitization is the ever-increasing flow of data presenting challenges in management and analysis¹. Biomedical terminologies provide semantics and metadata standards to process information in biomedical research and in healthcare systems. They promote the integration of disparate data sources and enable data aggregation².

Different terminologies are used for disparate purposes facilitating various aspects of healthcare. Clinical terminologies like SNOMED CT are used in Electronic Health Records enabling semantic interoperability and supporting the exchange of data between different institutions³. On the other hand, classification systems like ICD-10-CM (International Classification of Diseases, 10th Revision, Clinical Modification) are used in medical claim reporting.

Laterality or sided-ness is an anatomic directional term used to indicate the left or right side of the body, as opposed to the medial⁴. It is usually indicated with terms such as “left,” “right,” “unilateral” (occurring on one side without specifying which), and “bilateral” (occurring on both left and right). However, synonyms such as “both sides” for bilateral and “one-sided” for unilateral may be used. Laterality is often found not only in terms related to body structures, but also in diseases and procedures. For instance, “*Left eye*” is a term denoting a body structure while “*Glaucoma of left eye*” is a disease, and “*Excision of left vas deferens*” is a procedure all with laterality of “left.”

Unfortunately, laterality errors represent some of the most catastrophic errors in medicine. Especially, when a body part is found on both sides of the body, there is an inherent risk of a procedure being performed on the wrong side. Even though efforts have been undertaken by many organization to reduce such misjudgments, they continue to occur⁵⁻⁷. In addition, laterality needs to be properly coded in patient data for effective downstream analysis. For instance, recent data has shown right-sided colon cancer has less survival rates than left-sided tumors⁸. Improper coding of laterality would affect such analyses.

Many biomedical terminology systems have introduced laterality in their terms. For instance, ICD-10-CM has been introduced with laterality directly built into the codes which the previous version ICD-9-CM did not have⁹. Therefore, it is worthwhile to investigate the terms indicating laterality in different biomedical terminologies. In this paper, we perform an analysis on how the terms indicating laterality are arranged in different terminologies based on the Unified Medical Language System (UMLS).

UMLS aggregates more than 16 million terms from 218 different source vocabularies including SNOMED CT, ICD-10-CM, Gene Ontology, and Human Phenotype Ontology^{10,11}. Terms from different terminologies conveying the

*Corresponding author. Email: guo-qiang.zhang@uth.tmc.edu

same meaning are grouped under 4 million UMLS concepts. A term from a source terminology is known as an atom with an Atom Unique Identifier (AUI) in the UMLS. A UMLS concept with a Concept Unique Identifier (CUI) is formed by grouping synonymous atoms. For instance, the UMLS concept “*Bilateral deprivation amblyopia*” (with CUI C2881375) groups atom “Deprivation amblyopia of bilateral eyes” (with AUI A32322707) from SNOMED CT and atom “*amblyopia deprivation both eyes*” (with AUI A20027683) from MEDCIN. Note that, in the rest of the manuscript, unless otherwise specified, we only include the AUI within parenthesis after the term.

According to our knowledge, this is the first analysis of the organization of terms indicating laterality across different terminological systems.

2 Methods

To perform our organizational analysis of terms indicating laterality, we first identify and extract such terms from the UMLS. Then, we perform several normalization steps and further identify matching term-pairs that we further analyze. In the rest of the paper, all terms indicating laterality are referred to as “lateral terms” in general. The terms indicating left side are referred to as “left lateral terms” and similarly those indicating right side are referred to as “right lateral terms.” The terms indicating one side not specifying left or right are referred to as “unilateral terms” while those indicating both sides are referred to as “bilateral terms.”

2.1 Identifying lateral terms

We perform a regular expression based search on all UMLS atoms to extract terms containing certain keywords as lateral terms. Table 1 contains keywords used for each lateral term type. Note that the search is case-insensitive.

Table 1: Keywords searched for identifying lateral terms.

Lateral term type	Keywords searched for
Left lateral terms	“left”
Right lateral terms	“right”
Unilateral terms	“unilateral,” “one-sided,” “one sided,” “one side”
Bilateral terms	“bilateral,” “left and right,” “right and left,” “both sides”

For instance, if a term contains “both sides,” it would be categorized as a bilateral term. Note that the keywords for bilateral terms are based on the atoms under UMLS concept “*Bilateral*” (with the CUI C0238767). We considered these as synonyms to the word “bilateral.” Similarly, keywords for unilateral terms are based on atoms under UMLS concept “*Unilateral*” (with CUI C0205092).

If a certain term was identified as belonging to more than one lateral term type, then it was disregarded. For instance, the atom “*Bilateral left bronchi*” (A19403752) was categorized as both a left lateral term and a bilateral term since it contains both “left” and “bilateral.” Hence, this term was discarded.

2.2 Normalization

To ensure that we capture the different lexical variations with the same meaning, we normalize the identified terms as follows. We first convert the term to lower case. Next, the term is tokenized into words by using the word tokenizer in the Natural Language Toolkit (NLTK)¹². Then the punctuation and stop words are removed and the remaining words are stemmed using the SnowballStemmer in NLTK. The duplicated stemmed words are removed and the resultant is considered as an unordered set of words. For instance, the term “*Primary iridocyclitis of right eye*” (A32320378) would be normalized to: { “primari,” “iridocycl,” “right,” “eye”}. Note that the term “*iridocyclitis primary right eye*” (A33582445), which is synonymous to the earlier term, is normalized to the same.

For bilateral and unilateral terms, an additional step is performed before normalization. If a bilateral term contains “left and right,” “right and left,” or “both sides,” these will be replaced with “bilateral.” Similarly, if a unilateral term

contains “one-sided,” “one sided,” “one side,” these will be replaced with “unilateral.” In addition, the words “left,” “right,” “bilateral,” “unilateral” are not stemmed.

2.3 Obtaining matching term-pairs

Matching lateral term-pairs are identified as follows. The normalization of each lateral term would contain a word indicating laterality which could be either of the four words “left,” “right,” “unilateral,” or “bilateral.” Given the normalization of a lateral term, we replace this word indicating laterality with the other three laterality indicating words to generate three sets-of-words. If there exists a lateral term from the same source terminology with the normalization equivalent to one of these sets-of-words, then, a match is found.

Table 2 shows an example of three matching left lateral, right lateral, and unilateral terms that were found for a given bilateral term “*absent radial pulses on both sides*” (A17270411). For instance, replacing the word “bilateral” in the normalization {“absent,” “radial,” “puls,” “bilateral”} with the word “left” would result in a set-of-words {“absent,” “radial,” “puls,” “left”} which is found to be the normalization of the left lateral concept “*absent left radial pulse*” (A17170173). Hence, the terms “*absent radial pulses on both sides*” (A17270411) and “*absent left radial pulse*” (A17170173) would form a matched lateral term-pair. Note that three matches may not be found for each lateral term.

Table 2: Matching term identification.

Bilateral term	<i>absent radial pulses on both sides</i> (A17270411)
Normalization	{“absent,” “radial,” “puls,” “bilateral”}
Replacement with “left”	{“absent,” “radial,” “puls,” “left”}
Matching left lateral term	<i>absent left radial pulse</i> (A17170173)
Replacement with “right”	{“absent,” “radial,” “puls,” “right”}
Matching right lateral term	<i>absent right radial pulse</i> (A17203609)
Replacement with “unilateral”	{“absent,” “radial,” “puls,” “unilateral”}
Matching unilateral term	<i>absent radial pulse on one side</i> (A17153629)

In some instance, replacement of a laterality indicating word by a different laterality indicating word would result in multiple matches. For instance, Table 3 contains three matches obtained by replacing “bilateral” in “*Imaging @ Upper Arteries @ Fluoroscopy @ Vertebral Arteries, Bilateral @ None @ None*” (A19891162) with “left.”

In such instances, we compute the edit distance (Levenshtein distance) between the original lateral term and the matched terms to identify the best match. The edit distance is a method to quantify the similarity between two strings which is computed by counting the minimum number of operations required to convert one string to the other¹³. In our adaptation, we consider a word-level edit distance where the operations are counted at word-level, not character level. Before computing the edit distance, we obtain a sequence of stemmed words from the two terms. Since, “*Imaging @ Upper Arteries @ Fluoroscopy @ Vertebral Artery, Left @ None @ None*” is the term with the least edit distance, this is considered as the best match.

3 Results

In this work, we used the 2022AA release of UMLS. Concept information was parsed from the MRCONSO.RRF file and the hierarchical relation information was parsed from MRHIER.RRF file of the UMLS release. In total, we obtained 177,874 left lateral concepts, 176,088 right lateral concepts, 14,320 bilateral concepts, and 1,869 unilateral concepts from this UMLS release. In total, we found 189,419 matched lateral term-pairs out of which only 3,186 (1.68%) were hierarchically related. Among the matches, 90.17% (170,800/189,419) were left lateral & right lateral matched term-pairs. Table 4 shows information regarding the matches obtained, the source terminologies from which the matches were obtained, situations where there exists a hierarchical relation between the matched term-pair, and the

Table 3: Edit distances between the bilateral term “*Imaging @ Upper Arteries @ Fluoroscopy @ Vertebral Arteries, Bilateral @ None @ None*” (A19891162) and three of its matched left lateral terms.

Bilateral term	<i>Imaging @ Upper Arteries @ Fluoroscopy @ Vertebral Arteries, Bilateral @ None @ None</i> (A19891162)
Stemmed sequence-of-words	[“imag,” “@,” “upper,” “arteri,” “@,” “fluoroscopi,” “@,” “vertebr,” “arteri,” “,” “bilateral,” “@,” “none,” “@,” “none”]
Match 1	<i>Imaging @ Upper Arteries @ Fluoroscopy @ Vertebral Artery, Left @ None</i> (A19891161)
Stemmed sequence-of-words	[“imag,” “@,” “upper,” “arteri,” “@,” “fluoroscopi,” “@,” “vertebr,” “arteri,” “,” “left,” “@,” “none”]
Edit distance	3
Match 2	<i>Imaging @ Upper Arteries @ Fluoroscopy @ Vertebral Artery, Left @ None @ None @ None</i> (A16250989)
Stemmed sequence-of-words	[“imag,” “@,” “upper,” “arteri,” “@,” “fluoroscopi,” “@,” “vertebr,” “arteri,” “,” “left,” “@,” “none,” “@,” “none,” “@,” “none”]
Edit distance	3
Match 3	<i>Imaging @ Upper Arteries @ Fluoroscopy @ Vertebral Artery, Left @ None @ None</i> (A19799708)
Stemmed sequence-of-words	[“imag,” “@,” “upper,” “arteri,” “@,” “fluoroscopi,” “@,” “vertebr,” “arteri,” “,” “left,” “@,” “none,” “@,” “none”]
Edit distance	1

terminologies containing related matched term-pairs for different types of matches. For instance, 8,554 matched-pairs were obtained between bilateral and left lateral terms which were observed in 18 source terminologies. Out of these, 1,502 were found to be related through hierarchical relations. These hierarchical relations were observed in only 2 source terminologies. Upon further investigation, it was found that out of these 1,502 related matched term-pairs, all but 1 were in SNOMED CT. It was also observed, while the 8,519 matched term-pairs between bilateral and right lateral terms were observed in 17 source terminologies, the 1,491 related matched term-pairs were only observed in SNOMED CT.

Table 4: Matched term-pairs identified and relations observed among them.

Matched term-pair type	Num. matched term-pairs	Num. terminologies containing matched term-pairs	Num. related matched term-pairs	Num. terminologies containing related term-pairs
Bilateral & left lateral	8,554	18	1,502	2
Bilateral & right lateral	8,519	17	1,491	1
Bilateral & unilateral	1,138	20	2	2
Left lateral & right lateral	170,800	31	5	4
Left lateral & unilateral	204	7	93	3
Right lateral & unilateral	204	7	93	3

A term-pair is considered to be siblings if they are hierarchically unrelated and share a common parent. Among the 189,419 matched lateral term-pairs obtained, we noticed that 83,804 (44.24%) are siblings. Table 5 shows the breakdown of siblings obtained for different types of matches. For instance, bilateral & left lateral matched term-

pairs are not hierarchically related in 7,052 instances (out of 8,554 as given in Table 4) out of which 3,108 instances represent sibling cases.

Table 5: Unrelated matched term-pairs and siblings identified among them.

Matched term-pair type	Num. unrelated matched term-pairs	Num. terminologies containing unrelated matched term-pairs	Num. of matched term-pairs that are siblings	Num. terminologies containing sibling term-pairs
Bilateral & left lateral	7,052	18	3,108	15
Bilateral & right lateral	7,028	17	3,109	15
Bilateral & unilateral	1,136	20	853	20
Left lateral & right lateral	170,795	31	76,661	30
Left lateral & unilateral	111	7	37	3
Right lateral & unilateral	111	7	36	3

4 Discussion

In this paper, we investigated how lateral terms are hierarchically organized in different terminologies. Our results indicated that only 1.68% of matching lateral term-pairs are hierarchically related. However, 44.24% of matched term-pairs are actually siblings. Note that, out of the sibling cases, 91.48% (76,661/83,804) are left lateral & right lateral matched term-pairs. From Table 5, it can also be seen that the sibling cases are dispersed across many source terminologies indicating their commonality.

4.1 Analyzing matched lateral term-pairs in SNOMED CT

As seen in Table 4, there existed a considerable number of bilateral & left lateral; bilateral & right lateral related matched term-pairs. However, these cases are not dispersed among many source terminologies. All of these cases other than 1, were found in SNOMED CT, which warrants a further SNOMED CT specific analysis to investigate the reasons behind this potential anomaly. For instance, “*Lesion of left external auditory canal (finding)*” in Figure 2 is actually a parent of “*Lesion of bilateral external auditory canals (finding)*” in Figure 1.

Therefore, we performed a concept-level analysis of these cases based on their logical definitions. Each SNOMED CT concept is logically defined with attribute groups which contain one or more attribute-value pairs. Attribute-value pairs are grouped to avoid ambiguity in how they apply^{14,15}. For example, Figure 1 shows the logical definition of the SNOMED CT concept “*Lesion of bilateral external auditory canals (finding)*” which contains two attribute groups:

- (Finding site (attribute), “Left external auditory canal structure (body structure)”) (Associated morphology (attribute), “Lesion (morphologic abnormality)”))
- (Finding site (attribute), “Right external auditory canal structure (body structure)”) (Associated morphology (attribute), “Lesion (morphologic abnormality)”))

There were 1,608 bilateral & left SNOMED CT matched term-pairs of which 1,501 were related and 107 were unrelated. There were 1,604 bilateral & right SNOMED CT matched term-pairs of which 1,491 were related and 113 were unrelated. Table 6 contains 10 examples of such cases in SNOMED CT. Note that we have used fully specified names of SNOMED CT concepts in this table. For instance, there exists a synonym “*Prosthetic arthroplasty of bilateral hips*” for the SNOMED CT bilateral term “*Replacement of bilateral hip joints (procedure)*” which matches with the term “*Prosthetic arthroplasty of left hip*.”

We considered a hierarchical relation to be potentially inferable between concepts *A* and *B* if they satisfy two conditions:

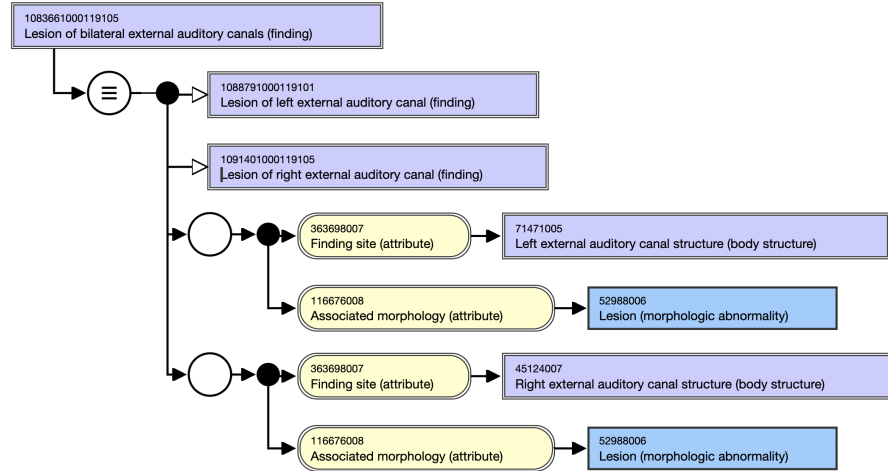


Figure 1: The logical definition of the SNOMED CT bilateral term: “*Lesion of bilateral external auditory canals (finding)*”. This figure has been adapted from the SNOMED International SNOMED CT Browser¹⁶.

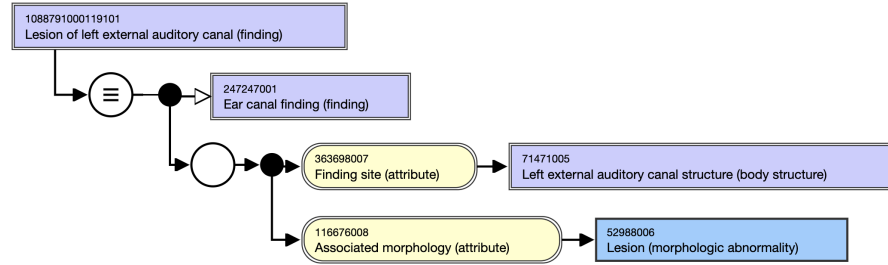


Figure 2: The logical definition of the SNOMED CT left lateral term: “*Lesion of left external auditory canal (finding)*”. This figure has been adapted from the SNOMED International SNOMED CT Browser¹⁶.

1. if the potential parent B is fully-defined,
2. if the attribute groups of the potential parent B is subsumed by potential child A

A SNOMED CT concept is fully-defined if its logical definition can distinguish it from other non-equivalent non-subtype concepts¹⁷. Otherwise, the concept is primitive¹⁸. If a concept is fully-defined, it is possible to automatically infer all its subtype concepts¹⁹.

We compared the attribute groups of these matched SNOMED CT term-pairs, particularly looking at whether one set of attribute groups is a subsumed by the other. If all the attribute groups of a concept B is found in concept A , then we consider the attribute groups of B are subsumed by A . For example, the attribute group of concept “*Lesion of left external auditory canal (finding)*” in Figure 2 can be found in the attribute groups of the concept “*Lesion of bilateral external auditory canals (finding)*”.

Table 7 shows the results for this analysis. For example, out of the 1,501 related bilateral & left lateral matched term-pairs, 1,381 (92%) were considered as inferable. As can be seen from the results, it is evident that a vast majority of related cases are potentially inferable while most unrelated cases are not potentially inferable. So, inferability is able to accurately identify an existence of or lack of a hierarchical relation among these cases. Therefore, root causes for such hierarchical relations may lie in the status (fully-defined or primitive) or logical definitions of the concepts. It is possible that these relations were inferred by a reasoner and not manually stated by SNOMED CT curators.

Table 6: Ten example related and unrelated matched lateral term-pairs in SNOMED CT.

Bilateral concept	Left/Right lateral concept	Related/Unrelated
<i>Astigmatism of bilateral eyes (disorder)</i>	<i>Astigmatism of right eye (disorder)</i>	related
<i>Angiography of bilateral adrenal arteries (procedure)</i>	<i>Angiography of right adrenal artery (procedure)</i>	related
<i>Polycystic bilateral ovaries (disorder)</i>	<i>Polycystic right ovary (disorder)</i>	related
<i>Amputation of bilateral hands (procedure)</i>	<i>Amputation of left hand (procedure)</i>	related
<i>Venography of bilateral lower limbs (procedure)</i>	<i>Venography of right lower extremity (procedure)</i>	related
<i>History of amputation of bilateral lesser toes (situation)</i>	<i>History of amputation of left lesser toe (situation)</i>	unrelated
<i>Bilateral fixed dilatation of pupil (finding)</i>	<i>Fixed dilatation of right pupil (finding)</i>	unrelated
<i>Both epididymides (body structure)</i>	<i>Structure of left epididymis (body structure)</i>	unrelated
<i>Excess skin of bilateral eyelids (finding)</i>	<i>Dermatochalasis of left eyelid (finding)</i>	unrelated
<i>Replacement of bilateral hip joints (procedure)</i>	<i>Prosthetic arthroplasty of left hip (procedure)</i>	unrelated

Table 7: Subsumption of the attribute groups of the bilateral term by left lateral or right lateral terms in SNOMED CT.

	Total cases	Potentially inferable
Related bilateral & left lateral	1,501	1,381 (92%)
Related bilateral & right lateral	1,491	1,372 (92%)
Unrelated bilateral & left lateral	107	3 (2.8%)
Unrelated bilateral & right lateral	113	2 (1.8%)

In some instances, non-inferability among an unrelated matched lateral term-pair could be tracked down to inconsistent modeling among the corresponding bilateral and left/right lateral concepts. For instance, consider the logical definition of the SNOMED CT bilateral term “*Sudden visual loss of bilateral eyes (disorder)*” in Figure 3 and corresponding matched right lateral term “*Sudden visual loss of right eye (disorder)*” in Figure 4. One of the “*Finding site (attribute)*” attribute relations in the bilateral term is linked to the term “*Structure of right visual system (body system)*.” However, in the right lateral term, this relation is linked to a different term: “*Right eye structure (body structure)*.” It is likely that if these bilateral and right lateral terms were consistently modeled, a hierarchical relation would be inferred between them.

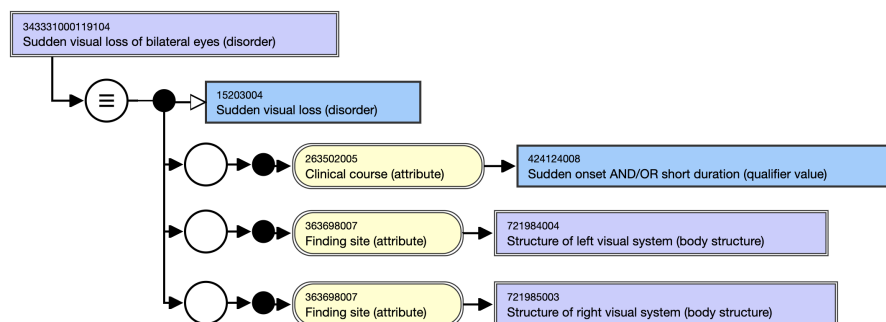


Figure 3: The logical definition of the SNOMED CT bilateral term: “*Sudden visual loss of bilateral eyes (disorder)*.” This figure has been adapted from the SNOMED International SNOMED CT Browser¹⁶.

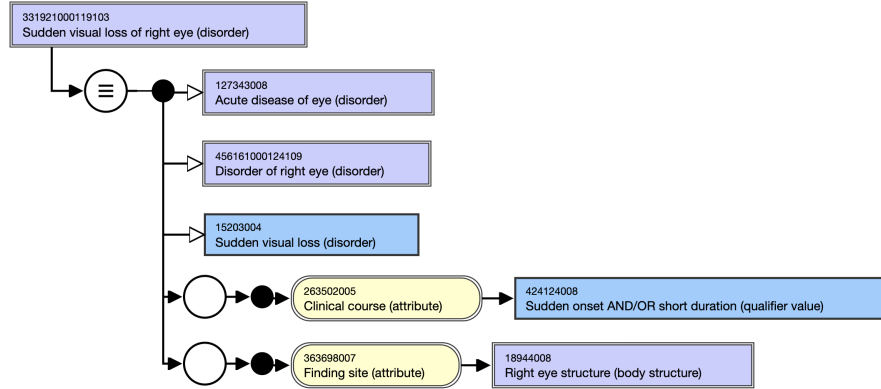


Figure 4: The logical definition of the SNOMED CT right lateral term: “*Sudden visual loss of right eye (disorder)*.” This figure has been adapted from the SNOMED International SNOMED CT Browser¹⁶.

4.2 Limitations and Future work

Our analysis has some limitations. First, certain lateral concepts are not considered in our method. For example, “*bunion of both feet*” (A24190341) is a bilateral term that is missed by our approach since it does not contain any of the keywords that were searched. In addition, we discarded concepts that were identified as belonging to multiple lateral term categories. For instance, the term “*Glaucoma, unilateral or bilateral*” (A27815785) was not considered as it was categorized as both unilateral and bilateral. It would be interesting to further investigate methods that could incorporate such cases.

Among the unrelated matched term-pairs, we only investigated siblings. There might be other interesting scenarios like uncle-nephew pairs that may be worthwhile investigating. Therefore, we will investigate a strategy that could identify all such types of relationships between unrelated matched term-pairs, not limiting to siblings.

Our analysis on SNOMED CT revealed that logical definitions of SNOMED CT concepts may have caused SNOMED CT bilateral & left; bilateral & right terms to be differently organized than other terminologies. However, we would like to further investigate these through a manual evaluation by domain experts to identify and validate potential inconsistencies that may exist among these concepts. If any inconsistencies are identified, how they affect downstream systems that utilize these terms could be determined and the impact of such shortcomings could be properly quantified.

5 Conclusion

In this paper, we investigated how lateral concepts are arranged in different biomedical terminology systems. We identified four kinds of lateral terms: left lateral terms, right lateral terms, unilateral terms, and bilateral terms. Then, we obtained matching lateral-pairs and further investigated how they are interrelated. From our results, it was seen that only 1.68% of matched lateral term-pairs are hierarchically related. We also identified a much larger 44.24% of cases as being siblings. It was also seen that, SNOMED CT handles bilateral & left; bilateral & right matched term-pairs differently to most other terminologies: a majority of such cases in SNOMED CT are hierarchically related. Further investigations revealed that the likely root cause for these relations might be how SNOMED CT arranges the logical definitions of bilateral concepts.

Acknowledgment

This work was supported in part by the National Science Foundation (NSF) under grant 2047001, as well as National Institutes of Health (NIH) under grants R01LM013335, R01NS116287, R01NS126690, R21AG068994, and U24MH130988. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NSF or NIH.

References

1. Konopka BM. Biomedical ontologies—A review. *Biocybernetics and Biomedical Engineering*. 2015 ;35(2):75-86.
2. Musen MA, Noy NF, Shah NH, Whetzel PL, Chute CG, Story MA, Smith B, NCBO team. The national center for biomedical ontology. *Journal of the American Medical Informatics Association*. 2012;19(2):190-5.
3. National Library of Medicine. Overview of SNOMED CT [Internet]. 2022 [cited 12 September 2022]. Available from: https://www.nlm.nih.gov/healthit/snomedct/snomed_overview.html.
4. ThoughtCo. Anatomical Directional Terms and Body Planes [Internet]. 2022 [cited 13 September 2022]. Available from: <https://www.thoughtco.com/anatomical-directional-terms-and-body-planes-373204>.
5. Gormley GJ, Dempster M, Corry R, Brennan C. ‘When right could be so wrong’. *Laterality errors in healthcare. The Ulster medical journal*. 2018;87(1):3.
6. Neily J, Mills PD, Eldridge N, Dunn EJ, Samples C, Turner JR, Revere A, DePalma RG, Bagian JP. Incorrect surgical procedures within and outside of the operating room. *Archives of Surgery*. 2009;144(11):1028-34.
7. Mitchell P, Dale T. Side errors in neurosurgery and human factors training. *Acta neurochirurgica*. 2015;157(3):487-91.
8. Bazarbashi S, Omar AT, Raddaoui L, Badran A, Alzahrani A, Aljubran A, Elhassan T. The Significance of Sidedness in Patients with Metastatic Colorectal Cancer Treated with Triplet First-Line Chemotherapy. *Asian Pacific Journal of Cancer Prevention: APJCP*. 2019;20(10):2891.
9. Cartwright DJ. ICD-9-CM to ICD-10-CM codes: what? why? how? *Advances in Wound Care (New Rochelle)*. 2013;2(10):588-92.
10. Bodenreider O. The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic acids research*. 2004;32(suppl_1):D267-70.
11. National Library of Medicine. Statistics - 2022AA Release [Internet]. 2022 [cited 13 September 2022]. Available from: https://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/release/statistics.html.
12. Loper E, Bird S. Nltk: The natural language toolkit. *arXiv preprint cs/0205028*. 2002.
13. Rane S, Sun W. Privacy preserving string comparisons based on Levenshtein distance. In: 2010 IEEE international workshop on information forensics and security 2010:1-6. IEEE.
14. SNOMED International. High Level Logical Model of SNOMED CT [Internet]. 2022 [cited 14 September 2022]. Available from: <https://confluence.ihtsdotools.org/display/DOCRELFMT/2.1+High+Level+Logical+Model+of+SNOMED+CT>.
15. SNOMED International. Expressions With Attribute Groups [Internet]. 2022 [cited 14 September 2022]. Available from: <https://confluence.ihtsdotools.org/display/DOCSCG/6.4+Expressions+With+Attribute+Groups>.
16. SNOMED International. SNOMED International SNOMED CT Browser [Internet]. 2022 [cited 5 January 2023]. Available from: <https://browser.ihtsdotools.org/?>.
17. SNOMED International. sufficiently defined concept [Internet]. 2022 [cited 14 September 2022]. Available from: <https://confluence.ihtsdotools.org/display/DOCGLOSS/Sufficiently+defined+concept>.
18. SNOMED International. primitive concept Internet. 2022 [cited 14 September 2022]. Available from: <https://confluence.ihtsdotools.org/display/DOCGLOSS/primitive+concept>.
19. SNOMED International. What does it mean if a concept is fully-defined or primitive and how do I tell the difference? [Internet]. 2022 [cited 14 September 2022]. Available from: <https://ihtsdo.freshdesk.com/support/solutions/articles/4000050378>.