

医療記録における縮約表現の量的構造—医療用語との比較—

言語資源
ワークショップ
2022

山崎誠 (国立国語研究所), 黒田航 (杏林大学), 東条佳奈 (大阪大学), 西嶋佑太郎 (医師), 麻子軒 (関西大学), 相良かおる (西南女学院大学)

1.はじめに(概要)

- 本発表は「実践医療用語辞書ComeJisyoから作成された2つのデータ, 「実践医療用語_語構成要素語彙試案表 Ver.2.0」(2022年3月, 言語資源協会より公開。以下, 「**試案表**」)と「**医療縮約表現**」(5690語)の語彙を量的に比較したものである。

3.実践医療用語とは

- 専門分野で使われる「専門用語」には, ①学術上の専門用語(学術用語)と, ②それ以外の専門用語がある。本研究では, 医療施設で使われる医療記録に含まれる①と②を合わせた用語を「実践医療用語」という。(相良2021: 559)

5.医療縮約表現

- 「実践医療用語辞書ComeJisyoUtf8-3」から「医療記録文に1回以上出現した、後要素がサ変語幹・形容動詞語幹・副詞可能名詞である複合名詞より選定した5690種」(関連発表, p4-5, 東条他)
- いわゆる「臨時一語」との違いは, 関連発表p4-5を参照。

6.分析手順

- 医療縮約表現に形態素解析を施す(一部目視で修正)。
- 分解された各語構成要素が「試案表」に登録されているかどうかを確認。
- 「試案表」での有無, および, 意味ラベル列を集計(意味ラベルが付かなかった語構成要素はNULLと表記した)。

表1 分析結果(一部)

医療縮約表現	形態素解析結果	「試案表」での有無(■:有, □:無)	有無パタン	意味ラベル列
切除困難	切除 困難	■切除 ■困難	11	医療行為 状態
日常管理	日常 管理	□日常 ■管理	01	NULL 行為
入院継続	入院 継続	■入院 □継続	10	医療行為 NULL
透過性改善	透過性 改善	□透過性 □改善	00	NULL NULL

表2 「試案表」に有無のパタン(上位10個)

パタン(0=「試案表」に無し,1=あり)	頻度	割合(%)	語構成要素数
00	1416	24.89	2
01	882	15.50	2
11	515	9.05	2
10	466	8.19	2
001	298	5.24	3
000	264	4.64	3
111	236	4.15	3
011	227	3.99	3
010	216	3.80	3
100	214	3.76	3

表3 「試案表」にある語構成要素を含む医療縮約表現の割合

語構成要素数	「試案表」に有の語の割合
2	56.8
3	85.7
4	92.6
5	96.7
6-9	100
全体	69.9

- 「試案表」にある語構成要素を1つでも含む医療縮約表現が全体の約7割であった(表3の全体の行)。
- 医療縮約表現の構成要素数が多くなると, 「試案表」に含まれる語が現れる割合が高くなることが確認された(表3)。

2.関連発表

- p1-2: 「急性」を含む病名の語構成 (相良かおる, 西嶋佑太郎, 東条佳奈, 高崎智子, 山崎誠)
- p2-6: 実践医療用語_語構成要素語彙試案表 Ver.2.0の構築 (東条佳奈, 黒田航, 相良かおる, 高崎智子, 西嶋佑太郎, 麻子軒, 山崎誠)
- p4-5: 医療記録における縮約表現の分析 (東条佳奈, 黒田航, 相良かおる, 西嶋佑太郎, 麻子軒, 山崎誠)

4.実践医療用語_語構成要素語彙試案表 Ver.2.0

- 「ComeJisyoSjis-1」に含まれる合成語のうち『分類語彙表 増補改訂版』収録の語を含むもの7,192語について, その語構成要素を一覧したもの。
- 語構成要素数は6,633個。
- 各語構成要素には意味ラベルが付与されている。
- 「**語構成要素**」の考え方(相良2021: 559より引用)
- 合成語を構成する要素で, 本研究では「合成語を医療の観点から意味的にまたは統語的に分割可能なすべての部分文字列」と定義する。分割できない合成語については元の合成語を語構成要素とする。
- 合成語: 脳幹多発性硬化症
- 語構成要素: 脳幹, 多発性, 硬化症, 多発性硬化症

相良かおる, 「実践医療用語における語構成要素の意味ラベルについて」言語処理学会第27回年次大会発表論文集, p.559-562, 2021.

7.結果

表4 意味ラベルの頻度(語頭に偏るもの)

意味ラベル	全体	語頭	語末	語頭-語末
身体部位	984	658	1	657
NULL	8164	3590	2988	602
体内物質・体外物質	133	107	0	107
体内物質	101	77	1	76
病名	93	46	4	42
機器	78	35	2	33
方法	71	44	12	32

表5 意味ラベルの頻度(語末に偏るもの)

意味ラベル	全体	語頭	語末	語頭-語末
状態	814	150	552	-402
経過	390	39	325	-286
医療行為	1130	246	492	-246
行為	441	82	288	-206
状態・経過	252	24	190	-166
部分	244	19	181	-162

- 「試案表」の意味ラベルを使った分析では, 「身体部位」「体内物質・体外物質」「病名」等は見出し語の語頭に偏って出現し, 「状態」「経過」「医療行為」「行為」「部分」等は語末に偏って出現することが分かった(表4, 表5)。