

## 診療情報管理士テキストにおける 収載専門用語の経版変化について

**成美大学** 経営情報学部  
 医療福祉マネジメント学科  
 ○森本 大志 福富 亮平 三宅 貴之  
 貝増 英貴 家原 葵 星 雅之

## はじめに

- 昨年度、「診療情報管理士テキストからの用語抽出の試み」という演題で報告
- 教科書用語解析の手法構築と、用語の出現頻度を分析した。

## 目的・方法

- 専門課程Ⅲについては、2年毎の診療報酬改定に合わせて改訂されているようだ
- 改訂を経て、テキスト収載用語(内容)がどの程度変化しているのかを知りたい！

↓

- 昨年度構築した手法を用いて、**2004・2007・2009・2012年度版テキストの本文を解析し、用語抽出を行った。**

## 方法) OCRによる文字データ取得

OCR → 文字データ

イメージ スキャナ

index\_php.txt \* - TeraPad

ファイル(F) 編集(E) 検索(S) 表示(V) ウィンドウ(W) ツール(T) ヘルプ(H)

1 | 1-0-0-0.本章の目的は、医療の成立における社会資源の必  
 2 | 1-1-1-0.近年、病院管理において情報管理が取り上げられ  
 3 | 1-1-2-0.診療情報管理士(HIM: Health Info  
 4 | 1-1-2-1.わが国では、1964年から実施されている日厚  
 5 | 1-1-2-2.諸外国においては、病院で診療情報管理を専門的  
 6 | 1-1-2-3.1975年(昭和50年)に日本診療情報管理学会が  
 7 | 1-1-2-4.診療情報管理士は医療改革が急速に進む中で専門  
 8 | 1-1-2-5.診療情報管理士の業務は、法的な位置づけを基盤  
 9 | 1-1-3-0.医師法第24条には、「医師は診療を行ったとき  
 10 | 1-1-4-0.プライマリ・ケアを担う医療施設においては、住  
 11 | 1-2-0-0.医療は医学の社会的適応であり、社会システムの  
 12 | 1-2-1-0.保健・医療・福祉・介護にかかわる施設として、  
 13 | 1-2-1-1.医療を提供する場である医療施設には、病院、診

### 方法) 茶釜 (Ver.2.1)による形態素解析

WinCha  
 ファイル(U) 編集(E) 解析(A) ヘルプ(H)

全文解析 1-0-0-0.本章の目的は、医療の成立における社会資源の必要性を理解し、医療関連  
 1-1-1-0.近年、病院において情報管理が取り上げられるようになった。人的・物的  
 部分解析 1-1-2-0.診療情報管理士 (HIM: Health Information Mana  
 1-1-2-1.わが国では、1964年から実施されている旧厚生省病院管理研究所の教  
 1-1-2-2.諸外国においては、病院で診療情報管理を専門的に行う職種やその教育制

クリア

表層語  基本形  読み  発音  品詞  活用  区切り

オプション -F "%m,%U(%P-%)T,%F %n"

1-0-0-0.未知語,,  
 本章\_名詞一般,,  
 の\_助前-連体化,,  
 目的\_名詞一般,,  
 は\_助前-係助詞,,  
 記号\_読点,,  
 医療\_名詞一般,,  
 の\_助前-連体化,,  
 成立\_名詞-サ変接続,,  
 における\_助前-格助詞-連語,,  
 社会\_名詞一般,,  
 医療\_名詞一般,,  
 の\_助前-連体化,,  
 必要\_名詞-形容動詞語幹,,  
 性\_名前-接尾一般,,  
 を\_助前-格助詞一般,,

### 方法) Excelによる用語の抽出

A	B	C	D	E	F	G	H	I	J	K	L	M
1		単語	名詞分類	記号分類	未知語	非自立	品詞xの変種	用語連	品詞a	品詞b		
2	11-0-0-0	1-0-0-0	0	0	0	0	0	0	0	0	0	未知語
3	21-0-0-0	本章	1	0	0	0	1	1	1	1	1	名詞 一般
4	31-0-0-0	の	0	0	0	0	1	1	0	0	0	助詞 連体化
5	41-0-0-0	目的	1	0	0	0	1	1	1	1	1	名詞 一般
6	51-0-0-0	は	0	0	0	0	1	1	0	0	0	助詞 係助詞
7	61-0-0-0	記号	0	0	0	0	1	1	1	0	0	記号 読点
8	71-0-0-0	医療	1	0	0	0	1	1	1	1	1	名詞 一般
9	81-0-0-0	の	0	0	0	0	1	1	0	0	0	助詞 連体化
10	91-0-0-0	成立	1	0	0	0	1	1	1	1	1	名詞 サ変接続
11	101-0-0-0	における	0	0	0	0	1	1	0	0	0	助詞 格助詞
12	111-0-0-0	社会	1	0	0	0	1	1	1	1	1	名詞 一般
13	121-0-0-0	資源	1	0	0	0	1	1	1	1	1	名詞 一般
14	131-0-0-0	の	0	0	0	0	1	1	0	0	0	助詞 連体化
15	141-0-0-0	必要	1	0	0	0	1	1	1	1	1	名詞 形容動詞
16	151-0-0-0	性	1	0	0	0	1	1	1	1	1	名詞 接尾
17	161-0-0-0	を	0	0	0	0	1	1	0	0	0	助詞 格助詞
18	171-0-0-0	理解	1	0	0	0	1	1	1	1	1	名詞 サ変接続
19	181-0-0-0	し	0	0	0	0	1	1	0	0	0	助詞 自立
20	191-0-0-0	し	0	0	0	0	1	1	0	0	0	記号 読点

### 方法) Accessによるデータ集計

年	集計
2004	2
2003	2
2002	2
2001	2
2000	2
1999	2
1998	2
1997	2
1996	2
1995	2
1994	2
1993	2
1992	2
1991	2
1990	2
1989	2
1988	2
1987	2
1986	2
1985	2
1984	2
1983	2
1982	2
1981	2
1980	2
1979	2
1978	2
1977	2
1976	2
1975	2
1974	2
1973	2
1972	2
1971	2
1970	2
1969	2
1968	2
1967	2
1966	2
1965	2
1964	2
1963	2
1962	2
1961	2
1960	2
1959	2
1958	2
1957	2
1956	2
1955	2
1954	2
1953	2
1952	2
1951	2
1950	2
1949	2
1948	2
1947	2
1946	2

### 収載用語の変化の計測

- テキスト間の収載用語の変化の度合いは用語群の類似度によって測ることにした。

↓

**Jaccard係数**を評価指標に

### Jaccard係数 (Jaccard Similarity Coefficient)

- 集合の類似度 (similarity) を測る指標  
⇒今回は2つの用語群の**種類**に注目した
  - 文章中の用語の出現頻度は考慮せず、用語の種類数によって類似度を測った。
- Jaccard係数 ( $J$ ) の計算式
  - 用語群A・Bに含まれる用語の種類数を $a \cdot b$ 、A・Bに共通に含まれる用語の種類数を $c$ とすると、

$$J = \frac{c}{a + b - c} \quad 0 \leq J \leq 1$$

参考:大垣俊一. 多様性と類似度、分類学的新指標. Argonauta 2008; 15: 10-22.

### 章別各年度版テキストの用語類似度 (数値: Jaccard係数)

第1章	2004年版	2007年版	2009年版	2012年版
2004年版		0.6793	0.5355	0.5000
2007年版			0.6783	0.6052
2009年版				0.6165
2012年版				

  

第2章	2004年版	2007年版	2009年版	2012年版
2004年版		0.7176	0.5092	0.4203
2007年版			0.6368	0.5093
2009年版				0.6440
2012年版				

### 章別各年度版テキストの用語類似度 (数値: Jaccard係数)

第1章	2004年版	2007年版	2009年版	2012年版
2004年版		0.6793	0.5355	0.5000
2007年版			0.6783	0.6052
2009年版				0.6165
2012年版				

  

第3章	2004年版	2007年版	2009年版	2012年版
2004年版		0.6872	0.1917	0.1924
2007年版			0.2274	0.2173
2009年版				0.7136
2012年版				

### 章別用語の平均類似度 (数値: Jaccard係数)

章	平均類似度	章題目
第1章	0.6025	医療管理総論
第2章	0.5729	医療管理各論 I
第3章	0.3716	医療管理各論 II
第4章	0.5165	医療情報学
第5章	0.4911	医療統計学
第6章	0.4792	診療情報管理論 I
第7章	0.4677	診療情報管理論 II

## 結果のまとめ

- 第1章から7章の類似度の平均  
⇒3章を除いて0.5前後を示した



- 我々が知るべき用語は、改訂によってもそれほど大きくは変化していない可能性がある

## 考察(今後の課題)

- 専門用語の変化までは追えていない  
⇒専門用語を定義できなかった  
(使える辞書がなかった)
- 用語の正確な抽出には、辞書の充実が必要である。  
⇒今回得られたデータから独自辞書の作成が可能である