

令和 5 年 6 月 14 日現在

機関番号：34419

研究種目：基盤研究(C) (一般)

研究期間：2020～2022

課題番号：20K07213

研究課題名(和文)患者向け文書における医療用語の難易度判定と平易化システムの構築

研究課題名(英文) Construction of a system for determining the difficulty of medical terms in patient documents

研究代表者

松野 純男 (Matzno, Sumio)

近畿大学・薬学部・教授

研究者番号：30299094

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：患者から「医師や薬剤師が使う言葉が難しい」「説明文を読んでも理解できない」といった訴えを聞くことがある。そこで、難易度評価試験を実施し、データを分析することで、患者向け文書で使われる「医療用語を含む文章」を、医療知識のない人でも理解できる「わかりやすい文章」に変更するシステムを作成することを目的にした。

165単語中無作為に20単語を抽出して医療者、薬学生、一般に行ったテストデータを解析し、特異的パターンを示す14単語を抽出した。この14単語についてフォーカスグループによるインタビュー調査を実施したところ、単語単独の意味ではなく、複合語に関する理解しづらさを感じていることが明らかとなった。

研究成果の学術的意義や社会的意義

今回の研究で、難解な医療用語を抽出してたところ、「患者さん向け文書」として説明に用いられている多くの文書に専門用語が使用されている実態が明らかとなった。また難解と思われる単語へのヒアリングから、それぞれの単語の意味は理解できても、組み合わせによる複合語は何を意味しているか、専門家しか理解できていないことが明らかとなった。現在、この複合語に関するクイズを作成中であり、再度オンライン調査を実施することで、複合語に関する説明文書の作成を行う。最終的にはこれらの複合語を検索できるwebサイトを構築し、医療者と患者との齟齬から生じる医療に対する不信を払拭するシステムの構築を目指す。

研究成果の概要(英文)：We sometimes hear complaints from patients that the language used by doctors and pharmacists is difficult or that they cannot understand the explanatory text. Therefore, we conducted a difficulty evaluation test and analyzed the data to create a system to change "sentences containing medical terms" used in documents for patients into "easy-to-understand sentences" that can be understood even by users without medical knowledge.

Quiz response data from medical professionals, pharmacy students, and the general public were analyzed to identify 14 words that exhibit specific patterns. Focus group interviews were conducted with these 14 words, which revealed that respondents had more difficulty understanding compound words than the meaning of each word alone.

研究分野：医療統計学, データサイエンス

キーワード：医学専門単語 平易化 対応分析 ヒートマップ分析 複合語

1. 研究開始当初の背景

近年、医療においては患者の自己決定権を尊重するのが基本であり、インフォームド・コンセントに基づく治療の選択は必要不可欠となっている。しかし、実際の治療現場では「医療に関する説明をなかなか理解できない」という患者の不満も見受けられる。実際に「患者向け文書」として用いられている説明用のパンフレットを確認すると、その中には医療従事者しか用いないような難解な単語が多く存在している。また、薬剤師が患者に服薬指導を行う際の指導記録を確認すると、気づかないうちに専門単語を用いて説明を行い、患者がその内容を理解できていない可能性が認められた。

以上のことから、これらの文書や口頭で用いられる単語を見直し、非医療関係者である一般の患者に視点を置いた文書への改訂することが、円滑なコミュニケーションを図る上で重要と考えられた。

2. 研究の目的

上記の背景からまず、「医療用語を含む文章」をテキストマイニングなどの手法で解析し、医療知識のない者でも理解できるよう文章をわかりやすくし、医療従事者と患者のコミュニケーションを円滑にすることを目的として、医療用語の難易度判別を試みた。

次にこのデータを解析し、患者向け文書での専門単語の使用状況について評価した。

3. 研究の方法

医療専門単語に対するクイズを作成し、医療関係者と一般の方の認知度の違いを判定する。上記1の調査で抽出される単語数は数百レベルと予想されるため、クイズに関しては単語辞書ファイルからランダムに20問程度を出題するWebベースのシステムを構築した。出題にあたっては「～の意味を表す単語」を選ばせる問題と「～の単語の正しい意味」を選ばせる問題をそれぞれ構築する。正答率を正確に測定するために、誤答となる選択肢はランダムではなく、正答と類似したものを作成する必要があり、この部分はシステム的に行うのではなく、人間の関与が必要となるため、時間を要することが予想される。問題作成後、選択肢をシャッフルしたのちに、ランダムに20問程度を出題するWebシステムを構築し(図1)医療関係者、薬学生および一般の方のボランティアを募って回答を収集し、正答率の判定を行った。

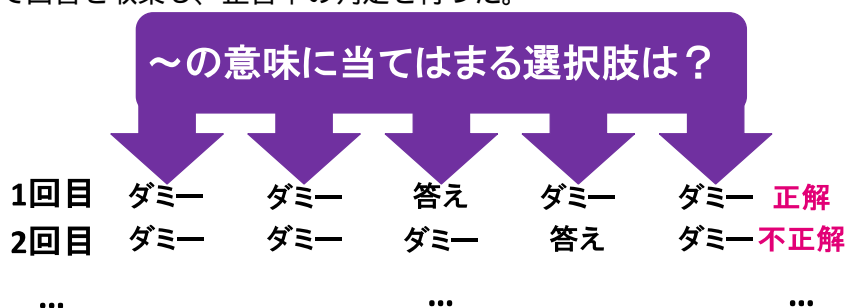


図1 Webクイズの構築

このクイズの回答結果に関して、ヒートマップ分析と対応分析を実施し、回答傾向に関して専門知識の有無がどのように関与するかを確認した。

4. 研究成果

単語辞書の作成

患者に向けて使われる医療関連用語に特化した辞典が流通していなかったため、薬局の服薬指導記録やがん治療の患者向けパンフレットなどの文章を、形態素に分割し、医療に関連する単語を抽出した。これらの単語に対して、医療専門辞典(保健同人: 家庭の医学、PDQ®がん用語辞書、JAIH国際保健用語集)と一般的な国語辞典(小学館: デジタル大辞泉、三省堂: 国語辞典)を参照し、その意味を付け加えた簡易辞書を作成した。この辞書をベースに、意味に対して理解度をテストする問題を、ランダ

ムなパターンで自動生成し難易度判定に用い、文章レベルの改善に繋げることを試みた。

形態素分割において、医薬品名などのカタカナ用語が細切れになり、医療に関連する単語抽出の大きな妨げとなった(表1)。そのため「錠」、「0.5mg」や「OD」などの修飾語に着目し、特定の修飾語が後に続くカタカナ語をPython3で抽出し、形態素から排除した。

表1 医薬品名に続く修飾語の例

外用	貼付剤	抽出	配合	皮	注	塩化
経口	軟膏	徐放	遺伝子	眼	点	臭化
テープ	ローション	腸溶	含嗽	筋	散	塩酸塩
ゼリー	カプセル	坐剤	細粒	鼻	液	硫酸塩
クリーム	ドライシロップ	浣腸液	顆粒	腔	錠	
エキス	シロップ	吸入	舌下		口腔	

### クイズ出題サイトの構築

簡易辞書CSVをベースに、与えられた意味に該当する単語を選択させる五択問題をWeb上に公開<sup>1)</sup>し、一人当たり20問をランダムに出題し、難易度判定を行った。医療専門職、医療系学生、一般の方の順で7.3%、9.5%、11.1%と総誤答率が高いが、誤答率の高い単語は、医療専門職、医療系学生、一般の方で概ね一致している。全回答者に共通して誤答の多かった「局所」「乾癬」「制御」などは、本来の定義を考慮すると、簡易辞書を作成した際の言い回しに問題があった可能性が考えられ、辞書CSVの質を上げ、同時にデータ数と語彙量を増やしていく必要性が示された。

表2 5択問題での誤答率

医療専門職		医療系学生		一般の方	
局所	0.438	局所	0.833	浸潤	0.697
制御	0.409	スケジュール	0.625	局所	0.641
喉頭	0.400	乾癬	0.579	乾癬	0.563
乾癬	0.368	軽減	0.550	喉頭	0.541
咽頭	0.333	喉頭	0.500	制御	0.500
スケジュール	0.313	制御	0.500	予後	0.429
上顎	0.300	改善	0.500	軽減	0.412
発疹	0.267	異食	0.471	経過	0.355
寛解	0.263	浸潤	0.438	スケジュール	0.346
症候群	0.250	リンパ腫	0.389	潰瘍	0.340

上記の点を踏まえ、次に単語の意味を3択で問う形でのクイズを作成し公開した<sup>2)</sup>。その結果、誤答率一覧とその内訳は下記の通りとなった。

表3 3択問題での誤答率まとめ

	総データ数	総誤答数	総誤答率
医療専門職	2340	191	0.0816
医療系学生 (無回答3人)	1520	183	0.1253

一般の方	2480 (無回答9人)	342	0.1487
------	-----------------	-----	--------

表4 専門性の異なる集団での誤答率の高い単語の抽出

誤答率が高い順 (Med:医療従事者,Stu:医療系学生,Oth:その他の方,数字は各々の誤答率)			
Med	Stu	Oth	
パンデミック	0.600	発疹	0.917
発疹	0.571	温存	0.750
オーバーシュート	0.538	パンデミック	0.700
クラスター	0.529	オーバーシュート	0.667
上皮	0.462	上皮	0.667
動悸	0.417	緩下剤	0.636
疾患	0.364	阻害	0.556
浸潤	0.353	アナムネーゼ	0.500
評価	0.333	リンパ腫	0.500
残存	0.333	クラスター	0.500
公衆衛生	0.308	公衆衛生	0.500
口蓋	0.300	鼻腔	0.500
外耳	0.286	服用	0.500
プライマリケア	0.250	疾患	0.455
進展	0.250	末梢	0.444
ソーシャルディスタンス	0.250	動悸	0.417
発作	0.231	早期	0.385
組織	0.231	プライマリケア	0.375
乾癬	0.222	方針	0.375
上顎	0.222	寛解	0.333
		咽頭	0.385
		緩下剤	0.889
		パンデミック	0.867
		上皮	0.857
		発疹	0.778
		放射	0.750
		動悸	0.667
		アナムネーゼ	0.667
		療法	0.556
		口蓋	0.538
		プライマリケア	0.533
		クラスター	0.500
		離開	0.471
		疾患	0.467
		主訴	0.455
		方針	0.438
		潰瘍	0.429
		原発	0.429
		浸潤	0.421
		リンパ腫	0.417
		咽頭	0.385

この回答結果を対応分析(図2)およびヒートマップ分析(図3)にかけたところ、特に2グループ以上で難解と判定される14単語が抽出された(図4)

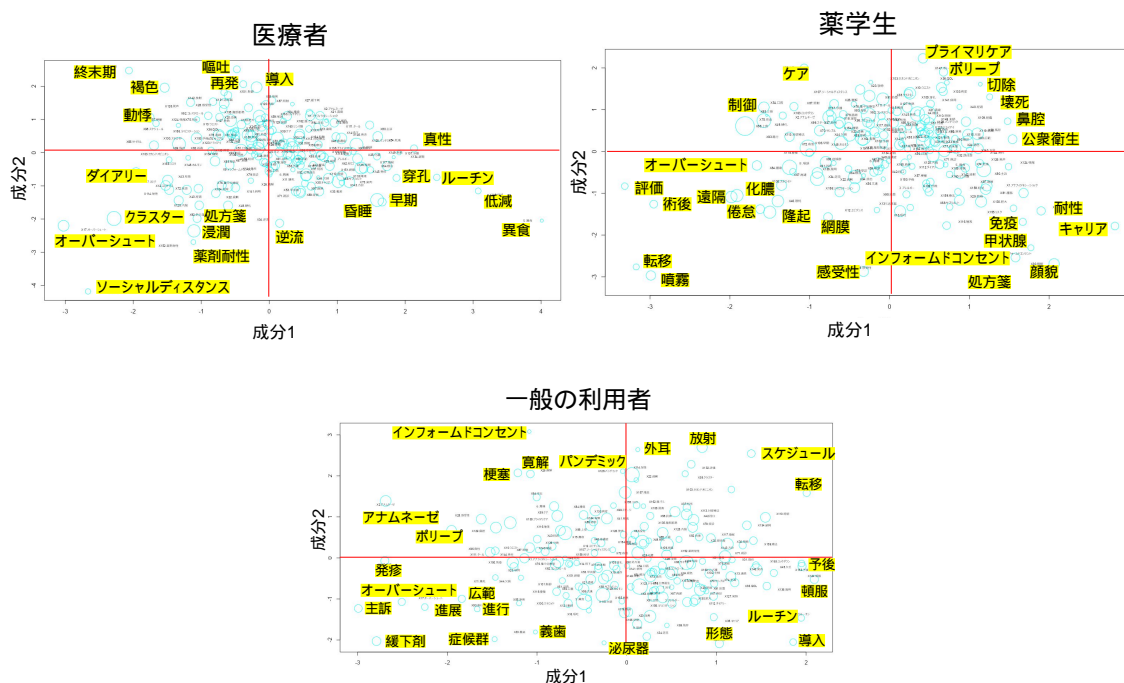


図2 医療知識の違うグループでの対応分析の比較

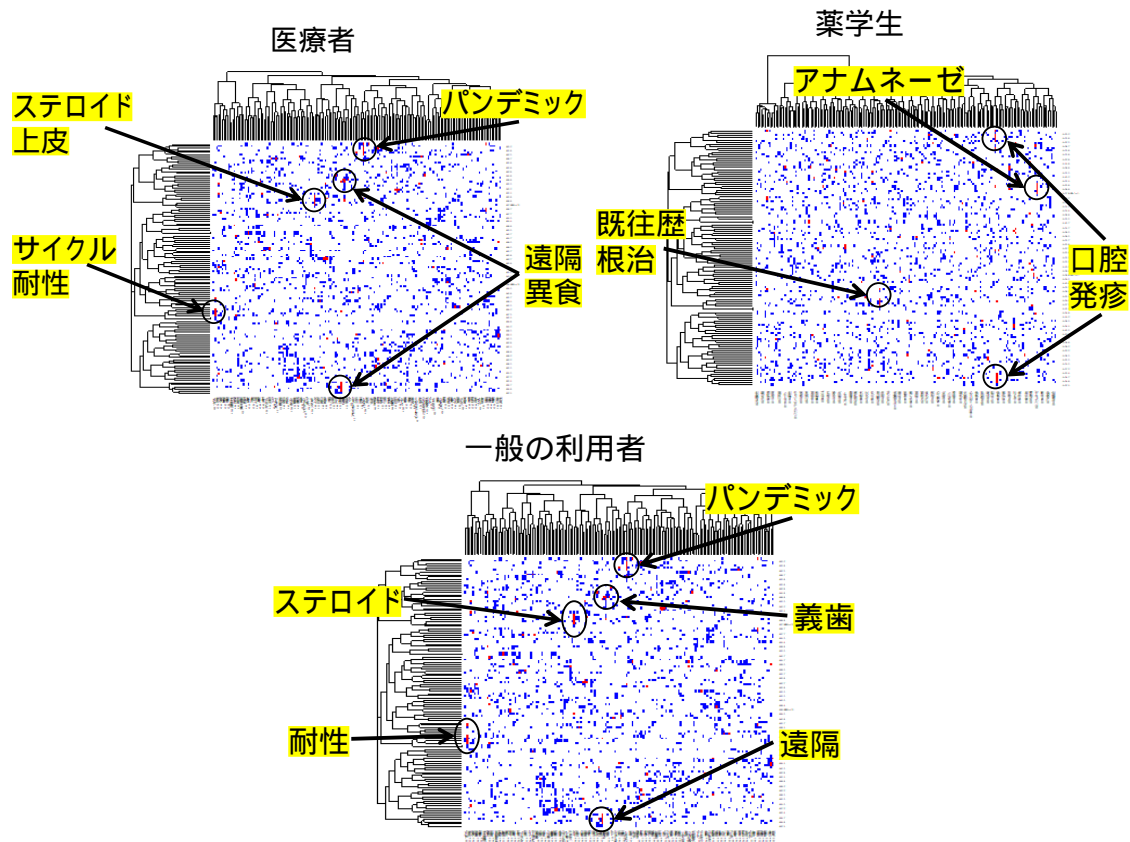


図3 医療知識の違うグループでのヒートマップ分析の比較

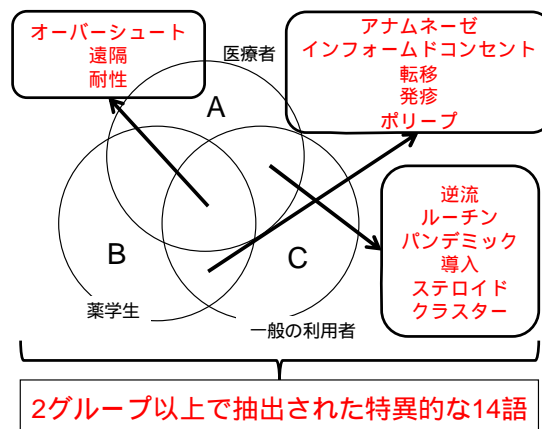


図4 2グループ以上で難解と判定される14単語

これら14単語について「患者向け文書」での使用状況を確認したところ、単語単体ではなく複合語で用いられていることが確認された。複合語になっている場合、単語単体よりもさらに理解が困難になることが予測されたため、今後は複合語での難易度について検討を行う予定である。

発表：難易度判定テストおよび患者向け文書の解析による難解医療用語の抽出：橋佑一郎、大鳥徹、大内秀一、和田哲幸、船上仁範、八軒浩子、大星直樹、松野純男。日本薬学会第134年会（2023年3月：札幌）

- 1) <https://pharmedu.phar.kindai.ac.jp/quiz/>
- 2) <https://pharmedu.phar.kindai.ac.jp/quiz2/>

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計3件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 橋佑一郎、大鳥徹、大内秀一、和田哲幸、船上仁範、八軒浩子、大星直樹、松野純男
2. 発表標題 難易度判定テストおよび患者向け文書の解析による難解医療用語の抽出
3. 学会等名 日本薬学会第143年会
4. 発表年 2023年

1. 発表者名 上海道毅、北小路学、松野純男、大鳥 徹
2. 発表標題 在宅医療における薬剤師の役割 医療と介護の連携を中心に経時的に考察する
3. 学会等名 第71回日本薬学会関西支部大会
4. 発表年 2021年

1. 発表者名 大和幹枝、松野純男、土肥弘久、長南謙一
2. 発表標題 統計教育に関する研究動向および課題点の抽出
3. 学会等名 日本薬学会第142年会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------