

令和 6 年 6 月 5 日現在

機関番号：16401

研究種目：基盤研究(B)（一般）

研究期間：2021～2023

課題番号：21H03164

研究課題名（和文）RWDに基づく臨床推論過程の定式化と集団知としての臨床推論システムの実現

研究課題名（英文）Formalization of clinical reasoning process based on RWD and realization of clinical reasoning system as collective intelligence

研究代表者

奥原 義保（Okuhara, Yoshiyasu）

高知大学・医学部・名誉教授

研究者番号：40233473

交付決定額（研究期間全体）：（直接経費） 13,900,000円

研究成果の概要（和文）：電子カルテの初診時記録を対象に、臨床推論システムへの活用可能性を検討した。まず、分類各クラスで、クラス間の差が多い頻出単語を選択、頻出単語が否定か肯定かについても評価した。この特徴語に患者の性別、年齢を加えた説明変数と電子カルテの病名から分類した2値クラスの間関係を、Naive Bayes, Exact Bayes, logistic回帰分析、決定木によって解析、推論精度や推論システムとしての実装可能性について検討した。その結果、特徴語と疾患の感度、特異度を求めて知識データベース化し、Naive Bayesを用いて推論システムとする方法が最も適していると考えられた。

研究成果の学術的意義や社会的意義

診療における推論過程は、臨床推論という形で様々な定式化が行われ、その大まかな道筋は広いコンセンサスを得ている。しかし、この過程は、医師が経験を積むほど頭の中で自動化され意識され難くなり、大規模なエビデンスに基づく検証は行われていない。しかし、病院情報システムの普及による電子化データの蓄積により、この推論過程を、大規模データから、誰もが理解できるよう論理的に見える化できる可能性が生まれている。これをエビデンスによって定量的に確認し、論理的に再現可能な形にすることにより、今後の医学教育の効果や質を飛躍的に高め、コンピュータによる、診断理由が説明可能な診療支援システムを可能にするはずである。

研究成果の概要（英文）：We investigated the feasibility of using electronic medical records at first consultations in a clinical inference system.

First, for each classification class, we selected frequently occurring words with the greatest difference between classes, and also evaluated whether the frequently occurring words were negative or positive. We analyzed the relationship between these feature words, plus the patient's gender and age as explanatory variables, and the binary classes classified from the disease names in the electronic medical record using Naive Bayes, Exact Bayes, logistic regression analysis, and decision trees, and examined the inference accuracy and feasibility of implementation as an inference system.

As a result, we found that the most suitable method was to obtain the sensitivity and specificity of the feature words and diseases, create a knowledge database, and use Naive Bayes to create an inference system.

研究分野：医療情報学

キーワード：臨床推論 real-world data 知識データベース 推論システム

1. 研究開始当初の背景

診療における推論過程は、臨床推論という形で様々な定式化が行われ、その大まかな道筋は広いコンセンサスを得ている^[1]。しかし、この過程は、医師としての経験を積むほど頭の中で自動化され意識され難くなり、大規模なエビデンスに基づく検証はこれまで行われていない。このような状況の中で、最近では、病院情報システム(HIS)の普及による電子化診療データの蓄積により、医師が疾患の診断や治療方法選択のために頭脳内で行う推論の過程を、大規模診療データから、誰もが理解できるように論理的に定式化(見える化)できる可能性が生まれている。とりわけ電子カルテの初診時記録には、臨床推論に有用な多くの情報を含んでいるはずであるが、構造化されていない様式のゆえに活用が難しいと考えられている。

2. 研究の目的

電子カルテの初診時記録を対象に、どのような形で臨床推論に活用できるかを検討した。また、推論システムとしての実装可能性についても検討した。

3. 研究の方法

対象としたのは高知大学医学部附属病院の電子カルテシステムに 2008 年から 2022 年に記載された初診時記録であり、2 種類のコホートに対して解析を行った。コホートの 1 つは、主訴として「めまい」が記載された患者であり、アウトカムとしての診断が「めまい疾患病名」(良性発作性頭位めまい症(BPPV)、突発性難聴、メニエール病、前庭神経炎、末梢性めまい、脳梗塞、脳出血)であるかそれ以外であるかの 2 クラス分類で評価した。もう一つのコホートは、脳神経外科を受診し初診時記録に主訴として「頭痛」が記載された患者で、診断が「脳外科領域疾患」の病名(脳出血、脳梗塞、脳腫瘍に関連)かそれ以外であるかの 2 クラス分類で評価した。

電子カルテの初診時記録に対し、まず、日本語自然言語処理オープンソースライブラリ「GiNZA(Python 3.7.15)」^[2]と、症状や病名に関連する語を広く抽出したデータである「万病辞書 Ver.202106」^[3]を用いてテキスト解析を実施、2 値分類各クラスでの名詞出現比率を求め、クラス間の差と件数が多い頻出単語を選択した。さらに、患者ごとの初診時記録について、頻出単語が否定か肯定かについて GiNZA の出力結果によって評価した。

このように、肯定、否定の有無も考慮した頻出単語(以降「特徴語」と表記)に患者の性別、カテゴリ変数化した年齢(40 歳未満、40 歳以上から 80 歳未満まで 10 歳間隔、80 歳以上)を加えた説明変数と電子カルテの病名から分類した 2 値クラスの間接関係を、Naive Bayes, Exact Bayes(条件付独立を仮定しない厳密な Bayes 推定を以降このように表記), logistic 回帰分析、決定木によって解析、推論精度や推論システムとしての実装可能性について検討した。

各方法の推論精度を比較するための指標としては、それぞれの方法によって推定されたクラス分類と実際のデータのクラス分類を比較し、正解の割合 = 「一致率」を考えた。

なお、Naive Bayes では、クラス分類を条件とする「特徴語」間の条件付独立に加え、「特徴語」間の独立も仮定した。

GiNZA は Python 3.7.15 で実行し、logistic 回帰分析、決定木解析は、それぞれ R4.1.3 上の glm 関数、rpart ライブラリを用いて実施、Naive Bayes, Exact Bayes は R4.1.3 上で作成したスクリプトを用いて実施した。

4. 研究成果

主訴として「めまい」が記載された患者は 249 人、主訴「頭痛」は 616 人であった。

(1) Bayes 推定：

主訴「めまい」の場合、

診断が「めまい疾患病名」である事前確率は 71.1%であった。特徴語 77 語(表 1)を用いた Naive Bayes の場合、一致率 76.7%、logistic 回帰の stepwise 法で特徴語を 15 個(表 2)に絞り込んだ場合、一致率：82.3%であった。特徴語 77 個を用いた Exact Bayes では一致率 100%であり、stepwise 法で 15 語に絞り込んだ特徴語を用いた場合は一致率 94.4%となった。

表 1：主訴めまいの特徴語

40歳未満	40歳以上50歳未満	50歳以上60歳未満	60歳以上70歳未満	70歳以上80歳未満	80歳以上	内服_pos	内服_neg	飲酒_pos	飲酒_neg	喫煙_pos	喫煙_neg
アレルギー_pos	アレルギー_neg	BP_pos	BP_neg	HR_pos	HR_neg	入院_pos	入院_neg	手術_pos	手術_neg	高血圧_pos	高血圧_neg
WBC_pos	WBC_neg	心電図_pos	心電図_neg	嘔気_pos	嘔気_neg	動悸_pos	動悸_neg	SpO_pos	SpO_neg	脳梗塞_pos	脳梗塞_neg
圧痛_pos	圧痛_neg	最近_pos	最近_neg	糖尿病_pos	糖尿病_neg	眼振_pos	眼振_neg	耳鼻_pos	耳鼻_neg	MRI_pos	MRI_neg
内服薬_pos	内服薬_neg	貧血_pos	貧血_neg	頭痛_pos	頭痛_neg	胸部_pos	胸部_neg	腹部_pos	腹部_neg	血圧_pos	血圧_neg
気分不良_pos	気分不良_neg	喘息_pos	喘息_neg	ふらつき_pos	ふらつき_neg	嘔吐_pos	嘔吐_neg	痛み_pos	痛み_neg	心音_pos	心音_neg
甲状腺_pos	甲状腺_neg	吐き気_pos	吐き気_neg	性別	x_posは単語xが肯定的、x_negは否定的に含まれていることを表す						

表 2 : 主訴めまいの特徴語 (stepwise で絞り込み)

性別	飲酒_pos	喫煙_pos	アレルギー_neg	HR_pos	HR_neg	頭痛_pos	頭痛_neg	WBC_pos	動悸_pos	眼振_pos	眼振_neg
耳鼻_pos	気分不良_pos	喘息_neg									

主訴「頭痛」の場合、

診断が「脳外科領域疾患」である事前確率は 24.6%であった。
 なお、特徴語には診断の結果と思われる語、動脈瘤、S A H、脳梗塞が含まれていたのをこれらを含む語を除いた 27 語 (表 3) と、logistic 回帰の stepwise 法で特徴語を絞り込んだ 17 語 (表 4) とを対象とした。

表 3 : 主訴頭痛特徴語

性別	40歳未満	40歳以上50歳未満	50歳以上40歳未満	60歳以上70歳未満	70歳以上80歳未満	80歳以上	紹介_pos	飲酒_pos	喫煙_pos	喫煙_neg	救急_pos
jcs_pos	MRI_pos	MRI_neg	アレルギー_pos	アレルギー_neg	ecg_pos	M R A_pos	M R A_neg	肩こり_pos	肩こり_neg	経過観察_pos	構音障害_pos
構音障害_neg	感覚障害_pos	感覚障害_neg									

表 4 : 主訴頭痛の特徴語 (stepwise で絞り込み)

性別	40歳未満	40歳以上50歳未満	50歳以上60歳未満	60歳以上70歳未満	70歳以上80歳未満	80歳以上	紹介_pos	喫煙_pos	喫煙_neg	救急_pos	jcs_pos
アレルギー_neg	ecg_pos	M R A_pos	肩こり_pos	経過観察_pos							

Naïve Bayes の場合、一致率 84.1%、stepwise 法で特徴語を絞り込んだ結果、一致率 : 83.3%であった。特徴語 27 個を用いた Exact Bayes では一致率 92.5%であり、stepwise 法で特徴語を絞り込んだ場合は一致率 88.6%となった。

(2) Logistic 回帰 :

主訴「めまい」の場合

診断がめまい疾患病名であるか否かを目的変数、抽出した特徴語 77 語を説明変数として logistic 回帰モデルを構築、VIF=10 以上の説明変数を除外して最終的なモデル (表 5) とし、

表 5

内服_pos	飲酒_pos	喫煙_pos	アレルギー_pos	アレルギー_neg	BP_pos	BP_neg	HR_pos	入院_pos	入院_neg	手術_pos	高血圧_pos
内服薬_pos	貧血_neg	頭痛_pos	頭痛_neg	胸部_pos	胸部_neg	腹部_pos	腹部_neg	血圧_pos	血圧_neg	WBC_pos	心電図_pos
心電図_neg	嘔気_pos	嘔気_neg	動悸_pos	動悸_neg	SpO2_pos	脳梗塞_pos	脳梗塞_neg	圧痛_pos	圧痛_neg	最近_pos	最近_neg
糖尿病_pos	糖尿病_neg	耳鼻_pos	耳鼻_neg	眼振_pos	MRI_pos	MRI_neg	気分不良_pos	喘息_pos	喘息_neg	ふらつき_pos	ふらつき_neg
嘔吐_pos	嘔吐_neg	痛み_pos	甲状腺_pos	甲状腺_neg	吐き気_pos	吐き気_neg	PT_SEX				

個々のデータに対して得られるめまい疾患病名である確率が 0.5 以上をめまい疾患病名であるとして評価した。その結果、一致率は 88.0%、その他の指標は AUC: 0.937、95% CI: 0.903-0.970 sensitivity: 0.893, specificity: 0.861 であった。

主訴「頭痛」の場合

診断が「脳外領域の疾患」であるか否かを目的変数、抽出した特徴語から診断の結果と思われる項目である動脈瘤_pos, 動脈瘤_neg, S A H_pos, S A H_neg, 脳梗塞_pos, 脳梗塞_neg を削除した 27 語 (表 3) を説明変数として logistic 回帰モデルを構築、個々のデータに対して得られる「脳外科領域の疾患」である確率が 0.5 以上を脳外領域の疾患」として評価した。その結果、一致率は 84.4%、その他の指標は、AUC=0.858(95% CI: 0.823-0.894(DeLong)) sensitivity= 77.6%, specificity=76.1%であった。

(3) Decision Tree:

主訴「めまい」の場合

診断が「めまい疾患病名」であるか否かをクラス分類、抽出した特徴語 77 語を分岐の属性とした場合、樹構造の出発点 root の段階、すなわち分岐の条件を何も課さない状態で「めまい」と診断された確率は 71.1%であった。R の関数 rpart で、樹構造の複雑さを指定するパラメータ cp(complexity parameter)=0.0007 とした場合、末端の分岐 (terminal mode) の数は 6 個で、最も一致率が良かった分岐の一致率は 90.9%、最も悪かった分岐は 62.5%、全分岐の期待値は 82.0%であった。

主訴「頭痛」の場合

診断が「脳外領域の疾患」であるか否かをクラス分類、抽出した特徴語から診断の結果と思われる語、動脈瘤、S A H、脳梗塞を除いた 27 語 (表 3) を分岐の属性とした場合、Root の段階で「脳外領域疾患」と診断された確率は 24.7%であった。cp=0.0007 とした場合、末端の分岐 (terminal mode) の数は 9 個で、最も一致率が良かった分岐の一致率は 93.8%、最も悪かった分岐は 61.1%、全分岐の期待値は 84.4%であった。

(4) 各モデルの比較

各モデルの結果をまとめると表 6 のようになる

表 6

方法	主訴	特徴語数	一致率	備考
Naive Bayes	めまい			事前確率71.1%
		77	76.70%	
	15	82.30%	Logistic回帰stepwise 法で絞り込み	
	頭痛			事前確率24.6%
27		84.10%	診断の結果と思われる語を除外	
17	83.30%	Logistic回帰stepwise 法で絞り込み		
Exact Bayes	めまい			事前確率71.1%
		77	100%	
	15	94.40%	Logistic回帰stepwise 法で絞り込み	
	頭痛			事前確率24.6%
27		92.50%	診断の結果と思われる語を除外	
17	88.60%	Logistic回帰stepwise 法で絞り込み		
Logistic回帰	めまい	56	88.00%	77語から VIF=10以上の説明変数を除外
	頭痛	27	84.40%	診断の結果と思われる語を除外
決定木分析	めまい	77		Rootでの「めまい疾患病名」: 71.1%, (全分岐数6)
			最大90.9%	
			最小62.5%	
	期待値 ^{注1)} 82.0%			
	頭痛	27		Rootでの「めまい疾患病名」: 24.6%, (全分岐数9)
			最大93.8%	
最小61.1%				
期待値 ^{注1)} 84.4%				

注 1) (各分岐の一致率×各分岐のデータ数)の総和/全データ数

この表にみられるように、一致率が最も良かったのは Exact Bayes であったが、主訴「めまい」の場合、データに見られた特徴語の組み合わせは、ほぼ 1 症例 1 種類ずつであり、明かに over fitting であると考えられる。そのほかの方法については、Naive Bayes の主訴「頭痛」、logistic 回帰、決定木分析は、ほぼ同様の結果となった。Naive Bayes の場合、特徴語が 77 と多い主訴「めまい」の一致率は事前確率よりわずかに良い程度であったが、logistic 回帰の step wise 法で 15 語に絞り込んだ結果、一致率 82.3%まで向上した。そもそも「めまい」の場合、特徴語 77 に対してデータ数は 249 と少なく、Exact Bayes の一致率 100%のように over fitting が起きる可能性が高いため、step wise 法など適切な方法により説明変数を絞り込む必要があると考えられる。Exact Bayes の「めまい」では 15 語に絞り込んだ結果も一致率 94.4%という高い値を示し、適切な方法で特徴語を絞り込めば最も精度の高い方法と言える。

このように、一致率という点では Exact Bayes が最も良い方法と言えるが、推論システム構築のための方法として考えた場合、多数の疾患とその特徴量の組み合わせは膨大になり、必要とされる計算資源や効率の点および高速での検索が可能な知識データベース化は簡単ではないと考えられる。また、各特徴量が推論の結果にどのように影響しているかを評価することは困難であり、推論過程の「見える化」は実現できない。Logistic 回帰分析については、各特徴量が推論の

結果にどのように影響しているかを評価することは容易であるが、知識データベース化については Exact Bayes と同様の問題点がある。どちらも、教師データを作成するためには有用と考えられる。

Naive Bayes については、特徴量が多い場合の一致率は他の方法より悪かったが、特徴量を適切に絞り込んだ結果は logistic や決定木と同様の一致率となった。また、計算に必要なのは、各疾患の各特徴量に対する感度と特異度であり、データベース化およびそこから検索、計算も容易である。各特徴量が推論の結果にどのように影響しているかの評価も容易である。また、2 クラス分類だけではなく、鑑別診断の場合など候補病名が複数である他クラス分類への拡張と、複数病名の組み合わせについてすべての症状を説明できる最小限の特徴量の組み合わせを残すアブダクション推論への発展も容易である。

決定木については、導かれるルールは分かりやすく推論過程の「見える化」が容易であるという利点があったが、疾患ごとに樹構造を最適化するパラメータを求めなくてはならず、また結果の構造化も簡単ではなく、大量のデータから知識データベースを構築するのに向いているとは言いがたい。限られた数の疾患のルールベース型知識データベース構築には有用と考えられる。また、教師データの作成にも有用と考えられる。

なお、当初は Association 分析も候補の一つとして考えたが、Association 分析の「条件部」で示される「特徴語」の組み合わせには、実際に各患者に見られる「特徴語」の組み合わせの部分集合も含まれるため、不必要な大量のルールが生成される。このため、必要とされる計算資源や効率の点から他の方法より優れているとは言えない。推論システムとしてデータベース化する際にも推測すべき患者の「特徴語」の組み合わせに完全一致する「条件部」に該当するルールを冗長で膨大なルールから探し出す必要がある。また、「条件部」に含まれる「特徴語」の個数も一定ではなく、1 語から可能性のある「特徴語」全ての個数までであり、「特徴語」全てからそれらの個数ごとの「特徴語」を選ぶ組み合わせ全てを考慮する必要がある。また、絞り込みの条件とする「確信度」の下限值はクラス分類の閾値（例えば 0.5）にする必要があり、通常の下限值 0.8 より低いと絞込みの結果も膨大になる。加えて、実際に各患者に見られる「特徴語」の組み合わせを評価するためには全トランザクション中に現れる対象となる「特徴語」の組み合わせを持つトランザクションの割合「支持度」も小さく設定する必要があり、さらなる結果の組み合わせ増大につながる。このため、Association 分析の結果を効率的な検索に適したデータベースにすることは簡単ではないと考えられる。

また、教師用データの候補を抽出する方法としても、必要とされる計算資源や効率の点から優れた方法とは考えられない。

結論として、電子カルテの初診時記録から得られる大量のデータから知識データベースを作成し、臨床推論システムを構築するための方法として、分類のクラスを特徴付ける名詞を、出現率の違いによって抽出し、さらに頻出単語が否定であるのか、肯定であるのかを評価して特徴語とし、それらと疾患の感度、特異度を求めてリレーショナルデータベースなどで知識データベース化し Naive Bayes を用いた推論方式によって推論システムとする方法が最も適していると考えられる。Exact Bayes, Logistic 回帰、決定木については教師データ作成には有用であると考えられる。

< 参考文献 >

- [1] Symptom to Diagnosis. Scott D.C. Stem, A.S. Cifu, D. Altkorn, McGraw-Hill (2015)
- [2] 松田 寛 GiNZA - Universal Dependencies による実用的日本語解析 - 自然言語処理 Volume 27 Number 3 (2020.09)
- [3] Kaoru Ito, Hiroyuki Nagai, Taro Okahisa, Shoko Wakamiya, Tomohide Iwao, Eiji Aramaki: J-MeDic: A Japanese Disease Name Dictionary based on Real Clinical Usage, International Conference on Language Resources and Evaluation (LREC) (Japan, Miyazaki), pp. 2365-2369, 2018 (2018/05/11)

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計0件

〔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
研究分担者	畠山 豊 (Hatakeyama Yutaka) (00376956)	高知大学・教育研究部医療学系連携医学部門・教授 (16401)	
研究分担者	宮野 伊知郎 (Miyano Ichiro) (00437740)	高知大学・教育研究部医療学系連携医学部門・准教授 (16401)	
研究分担者	兵頭 勇己 (Hyohdo Yuki) (50821964)	高知大学・教育研究部医療学系連携医学部門・助教 (16401)	
研究分担者	瀬尾 宏美 (Seo Hiromi) (80179316)	高知大学・教育研究部医療学系医学教育部門・教授 (16401)	
研究分担者	安井 繁宏 (Yasui Sigehiro) (00535346)	高知大学・教育研究部医療学系連携医学部門・助教 (16401)	他大学異動のため2023年度より研究分担者から離脱

	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
研究協力者	竹内 慎哉 (Takeuchi Shinya)	高知大学	

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
---------	---------