

## 学位論文審査の結果の要旨

令和 5 年 1 月 12 日

審査委員	主査	辻 見 仁 (計)		
	副主査	平 尾 智 光 (産)		
	副主査	宮 武 伸 行 (産)		
願出者	専攻	医学	部門	(平成27年度以前入学者のみ記入)
	学籍番号	19D738	氏名	間島 行則
論文題目	Using Natural Language Processing Techniques to Detect Adverse Events From Progress Notes Due to Chemotherapy			
学位論文の審査結果	<input checked="" type="radio"/> 合格 <input type="radio"/> 不合格      (該当するものを○で囲むこと。)			

## [ 要 旨 ]

本研究に関する学位論文審査委員会は令和5年1月12日に行われた。

本研究は、がん薬物療法を受けた患者に関して医師が電子カルテへ叙述的に記載した経過記録から、悪心嘔吐と下痢の症候の発現を検出するための自然言語処理技術を用いたコンピュータプログラムを新規開発し、その検出精度について医師の目視による検出結果との一致率を指標に評価した。その結果、悪心嘔吐について83.5%、下痢について97.7%と、両者とも高い一致率であったことを示し、有害事象を同プログラムによって効率的に検出する事が可能であることを示した。さらに、両症候間の一一致率に有意な差が検出されたことに着目し、それらの一致率に影響を及ぼす要因を探索するため、エラー分析を行った結果、悪心嘔吐の陰性所見に対する誤検出が多く発生したことを指摘した。さらに、今回使用したデータセットの特性を詳細に分析した結果、悪心嘔吐に関連した記載量が下痢に関するものより多く、特に悪心嘔吐の陰性所見の記載の多さが、悪心嘔吐の検出精度が下痢の検出精度に対して有意に下回った要因であることを明らかにした。そして、電子カルテに蓄積された叙述的記録から自然言語処理技術を用いたコンピュータプログラムにより症候を的確に検出するためには、その分析対象となる医療文書の特性について十分な理解が必要であるとの見解が示された。本研究で得られた成果は、医療における自然言語処理技術の正しい利活用を進めていく上で意義があり、学術的価値が高い。委員会の合議により、本論文は博士(医学)の学位論文に十分値するものと判定した。

審査においては、次に記載する項目をはじめ多数の質問が行われた。

1. 試行用データセットに配分する症例数について

(回答) 本研究では可能な限り精度の高い正解ラベルを用意するため、医師3名が手作業でカルテレビューを行った。個々の症例を目視で確認する必要があり、その作業量から合計200例程度が実現可能な症例数の上限と考えた。さらに、エラー分析を含めシステムの評価を詳細に行うため、そこから評価用データセットを多く確保した。一方、システムの精度を高めるためには、一般に試行用データセットを多く確保する方が好ましく、本システムの精度を改善するための今後の検討課題にしたいと考えている。

2. 効率的な正解ラベル付与の方法について

(回答) 手作業での正解ラベル付与は、担当者にとって大きな負担が発生する作業であり、その効率化が期待される。ご指摘頂いたように、既に実施されている臨床研究が構造化されたデータを持っているなら、またそのプロトコルが許すようであれば、それらの情報を活用してより大きな症例数でシステムを設計、検証することは有効であると考えている。

3. Limitationのひとつに挙げた、カルテの個性について

(回答) 医師3名が手作業でカルテレビューを行った。担当者は作業の後半になるにつれ、記載者を確認する前に、記載パターンからその記載者がある程度推測できるようになった。今回具体的な数値は求めているが、これらからはその記載パターンがある程度限定的であったと認識している。ただし、各症候を検出するための2つのデータセットは、無作為に分割されており、含まれる病名や投与薬剤に偏りはなく、記載パターンも同様と考えている。

4. 今後の展望について

(回答) 本研究を通じて、電子カルテの情報に不均一性があることを認識した。近年では電子カルテの情報を人工知能(AI)に学習させることで、その出力を医学的判断に用いる試みも多い。偏ったデータセットから学習したシステムは、その出力も偏ったものになってしまうことが危惧されるため、それらのシステム開発を担う研究者に対して情報の発信を続けることで、将来的には医療におけるAIの正しい利活用に貢献したいと考えている。

5. 症候記載が看護師の記録に多いことが話題に上がった後の質問：がん薬物療法は看護師の記録が少ない外来での投与に移行しつつあり、今後有害事象検出が困難となる可能性について

(回答) 現在、本研究とは別に、医師と看護師の経過記録を比較する研究を行っており、先日その要旨を学会報告した。それによると、医師の経過記録には診断名の記載が、看護師の経過記録には症候名の記載が、それぞれ多いことが明らかとなった。これに基づくと、症候名ではなく、診断名として検出することができれば対応できるのではないかと考えている。申請者はいずれにも明確に回答し、博士(医学)の学位授与に値する十分な見識と能力を有することが認められた。

掲載誌名	Cancer Informatics 第 21 巻		
(公表予定) 掲載年月	2022 年 3 月	出版社(等)名	SAGE Publishing

(備考) 要旨は、1, 500字以内にまとめてください。