

平成29年度厚生労働科学研究費補助金
(政策科学総合研究事業(臨床研究等 ICT 基盤構築・人工知能実装研究事業))
分担研究報告書

自然言語処理による MACE イベント捕捉の検討に関する研究

研究分担者 西村 邦宏 国立循環器病研究センター・予防医学・疫学情報部・部長
中村 文明 国立循環器病研究センター・循環器病統合センター・室長
竹上 未紗 国立循環器病研究センター・予防医学・疫学情報部・室長
安田 聡 国立循環器病研究センター・副院長・心臓血管内科部門長
宮本 恵宏 国立循環器病研究センター・循環器病統合センター・センター長

研究要旨

これまでの研究において、電子カルテシステム内に蓄積された所見・報告書・サマリなどのテキスト情報から、自然言語処理および機械学習を用いて、カルテ記載内における「症状記載」について、判別・予測する方法論の検討を行った。臨床研究業務担当者が実際に必要とする症状記載データについて、カルテ記載情報から手動で抽出を行った。これらを用いて教師データを作成し、カルテ記載における「症状記載」と「その他の記載」についての自動判別器を作成した。こちらは感度・特異度はともに70~80%の性能を有していることがわかった。

この結果をもとにさらに自然言語処理技術に関して先進的な IBM ワトソンにより Major Cardiac event をとらえることを目的に辞書チューニングを行った。心筋梗塞レジストリ MIDAS 研究を中心とした約2000人の国立循環器病センター入院患者に関して、最も記述が的確と考えられる退院時サマリの記述をもとに虚血性心疾患、心不全、脳卒中、心臓死、全死亡に関して IBM ワトソンエクスプローラーにより抽出を行った。死亡イベントに関しては、電子カルテ上の死亡退院により100%の把握が可能であった。初回の入院に関しては、入院契機が虚血性心疾患、心不全、脳卒中である場合もほぼ捕捉可能であった。死亡と入院契機の虚血性心疾患、心不全、不整脈項目により心臓死の確認が可能であった。辞書チューニング前は accuracy として65%前後であるが、チューニング後は95%以上の精度達成が可能であった。

入院時における血性心疾患、心不全、脳卒中、心臓死、全死亡に関して IBM ワトソンエクスプローラーによる自然言語処理でほぼ正確に MACE の捕捉が可能であった。

A．研究目的

前年度までの研究において、電子カルテシステム内に蓄積された所見・報告書・サマリなどのテキスト情報から、自然言語処理および機械学習を用いて、カルテ記載内における「症状記載」について、判別・予測する方法論の検討を行った。臨床研究業務担当者が実際に必要とする症状記載データについて、カルテ記載情報から手動で抽出を行った。これらを用いて教師データを作成し、カルテ記載における「症状記載」と「その他の記載」についての自動判別器を作成した。こちらは感度・特異度はともに70~80%の性能を有していることがわかった。

この結果をもとにさらに自然言語処理技術に関して先進的な IBM ワトソンにより Major Cardiac event をとらえることを目的に辞書チューニングを行った。

B．研究方法

心筋梗塞レジストリ MIDAS 研究を中心とした約2000人の国立循環器病センター入院患者に関して、最も記述が的確と考えられる退院時サマリの記述をもとに虚血性心疾患、心不全、脳卒中、心臓死、全死亡に関して IBM ワトソンエクスプローラーにより抽出を行った

(倫理面への配慮)

- 6 人を対象とする医学系研究に関する倫理指針を遵守し研究を遂行する。

C. 研究結果

死亡イベントに関しては、電子カルテ上の死亡退院により 100%の把握が可能であった。初回の入院に関しては、入院契機が虚血性心疾患、心不全、脳卒中である場合もほぼ捕捉可能であった。死亡と入院契機の虚血性心疾患、心不全、不整脈項目により心臓死の確認が可能であった。辞書チューニング前は accuracy として 65%前後であるが、チューニング後は 95%以上の精度達成が可能であった。

D. 考察

入院契機ではなく例えば予定 PCI として紹介され、既往歴としてイベントとなる病名が記述されている場合は、日付けが何年何月など明記されず入院時から去年など曖昧さを残す記述であることがあり、これら相対日付けに関しては今後の精度向上が必要と考えられた

E. 結論

入院時における血性心疾患、心不全、脳卒中、心臓死、全死亡に関して IBM ワトソンエクスプロラーによる自然言語処理でほぼ正確に MACE の捕捉が可能であった。

G. 研究発表

1. 論文発表 なし
2. 学会発表
 - 1) 医療情報連合大会 2017年11月22日
日本循環器学会合同シンポジウム
人工知能応用による自然言語処理の活用
電子カルテ情報のセマンティック登録と全国登録事業への将来展望

H. 知的財産権の出願・登録状況

1. 特許取得 なし
2. 実用新案登録 なし
3. その他 なし