

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 4 月 5 日現在

機関番号：13101

研究種目：基盤研究（B）

研究期間：2009～2011

課題番号：21390159

研究課題名（和文） 電子カルテからのダイレクトリスクマイニングの試み

研究課題名（英文） Detection of adverse events in electronic medical records using natural language processing

研究代表者

鳥谷部 真一（TOYABE SHIN-ICHI）

新潟大学・危機管理本部危機管理室・教授

研究者番号：20227648

研究成果の概要（和文）： インシデントレポートには、報告されない事例があることや、事例発生から報告までにタイムラグがあるという問題点がある。そのため、診療情報の多くを占めながら活用が進んでいない、テキストデータから事例を検出することを試みた。自然言語処理技術を用いて、転倒・転落事例を検出する判定規則を作成した。インシデントレポートで評価したところ、この判定規則の感度は 86.5%、特異度が 97.5%であった。実際の診療情報データでは、感度は診療経過記録、インシデントレポート、画像オーダエントリをデータ源とした場合に、特異度はインシデントレポート、画像オーダエントリで優れていた。画像オーダエントリを用いた場合、患者影響度が高い事例を、より短いタイムラグで検出でき、重大転倒事例の検出が 23%増加した。画像オーダエントリ情報のテキストマイニング処理は、迅速に重大事例を検出するのに適し、インシデントレポートを補完する方法として有用であった。今後は他のインシデント・有害事象についても今回と同様の方法論が有用であるかどうか検討する必要がある。

研究成果の概要（英文）： Incident reporting system is widely used in hospital to detect patient safety incidents and adverse events. However, there are several problems in the system such as under-reporting and lag-time between events and submission of the reports. To resolve these issues, we tried to detect incidents by using natural language processing of electronic medical records. We constructed syntactic-semantic category decision rules and applied these rules to narrative text data in medical records. As for inpatient falls, these rules showed satisfactory performance (sensitivity 86.5% and specificity 97.5%). Progress notes and discharge summaries were not appropriate for data sources to be used in this method because of their low specificity or low sensitivity. In contrast, natural language processing of text data on image order entries enables rapid detection of injurious falls. This method is useful to compensate the shortcomings of incident reporting system such as the under reporting and the lag-time.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009 年度	3,300,000	990,000	4,290,000
2010 年度	2,000,000	600,000	2,600,000
2011 年度	1,700,000	510,000	2,210,000
年度			
年度			
総計	7,000,000	2,100,000	9,100,000

研究分野：医療安全学

科研費の分科・細目：境界医学・医療社会学

キーワード：インシデントレポート、電子カルテ、テキストマイニング、自然言語処理、チャートレビュー、医療事故医療事故

1. 研究開始当初の背景

- (1) インシデントレポートの限界： 病院内でインシデントや有害事象が起こると、当事者や目撃者はインシデントレポートを医療安全管理部門に提出する。インシデントレポートはインシデントや有害事象の把握には最も有用な手段ではあるが、①個々の報告自体は義務ではなく自発的な行為であるために報告されない事例がある、②多忙な医療関係者に報告書作成という負担をかけることになる、③インシデントや有害事象の当事者にとっては、事例そのものの影響に加えて、報告することも心理的負担になると報告されている、③事例発生から報告までに無視できないタイムラグがあり、医療安全管理者が事例を速やかに検知できないことがある、などの問題点がある。
- (2) 活用されていない診療情報： インシデントや有害事象は診療録・診療記録に記載されることが多い。診療録・診療記録に残された記録は、検査結果に代表されるような数値など量的データについては活用されている。しかし、診療録・診療記録に記載されるデータのほとんどはテキストデータであり、活用が進んでいない。
- (3) 自然言語処理 natural language processing (NLP)・テキストマイニングの進歩： 近年、NLPあるいはテキストマイニング手法の開発が進み、質的データであるテキストデータを処理して、定量的・定性的に解析することができるようになりつつある。しかし、NLPの医療分野への応用は、本邦では報告が非常に少ない。
- (4) 診療録・診療記録からのNLPによるインシデント・有害事象検出： 診療録・診療記録をNLPで処理して、インシデントや有害事象を検出することによって、(1)や(2)で掲げた課題を解決にする糸口になる可能性がある。

2. 研究の目的

- (1) インシデント・有害事象検出の判定規則 decision rule の作成： 診療録・診療記録のテキストデータからNLPを用いて、インシデント・有害事象を検出するための、検出システムを構築する。あるいは、インシデント・有害事象を検出するた

めの判定規則を作成する。

- (2) 最適なデータ源の同定： 海外では過去に、入院総括や画像診断報告書を有害事象抽出のデータ源とした報告がある。診療録・診療記録には様々な種類の記録が含まれることから、どの種類のデータを用いれば最も効率よくインシデント・有害事象を検出できるかを調べる。
- (3) インシデントレポートとの性能比較： NLPによる判定規則を用いた場合と、従来のインシデントレポートを用いた場合とで、インシデント・有害事象の感度、陽性的中率、事例発生から事例検出までのタイムラグの長短を比較する。

3. 研究の方法

- (1) 判定規則の作成： 特定のインシデント・有害事象（例として、転倒・転落事故）に焦点を絞って、そのインシデント・有害事象に関連があるインシデントレポートを陽性コントロール、関連がないインシデントレポートを陰性コントロールとする。インシデントレポートのテキスト部分に形態素解析、係り受け解析を行い、陽性コントロールに検出され、陰性コントロールには検出されない単語、係り受け、それらの集合を選択する。選択には、 χ 二乗検定またはFisher正確検定を用いる。2群で有意に頻度に差があるもののうち、p値の小さいものから優先的に判定規則に組み入れる。
- (2) 判定規則の性能検証： 判定規則作成に用いたものとは異なるデータセット（インシデントレポートのテキスト部分）に、作成した判定規則を適用する。インシデント・有害事象に関連したデータのうちの程度の割合を正しく検出できたかを感度とし、不関連インシデントレポートのうち、どの程度の割合を不関連と判断したかを特異度とする。
- (3) インシデント・有害事象検出のためのチャートレビュー chart review： 有害事象を検出するための標準的な方法は、チャートレビュー manual chart review であるとされている。インシデント・有害事象の真値とするため、一定期間に入院したすべての患者の診療録・診療記録を目視して、その中に含まれるインシデントや有害事象を調査する。
- (4) 実際の診療録・診療記録を用いた判定規

則の性能評価： 判定規則を実際の診療録・診療記録に適用し、チャートレビューの結果と照合することで、どの程度の検出率があるかを調べる。また、

4. 研究成果

- (1) 判定規則の作成： 本院で半年間に報告されたインシデントレポートを、転倒に関連があるものと、そうでないものの2群に分けた。それぞれを転倒関連情報を検出する判定規則を作成する上での陽性コントロール、陰性コントロールとした。2群の比率の差の検定においてp値の小さいものから順に170の規則を採用した。判定規則をカテゴリに分けると、1)転倒が発生したという情報に関する規則(81)、2)転落が発生したという情報に関する規則(40)、3)転倒・転落によって患者が外傷を負ったという情報に関する規則(23)、4)身体バランスを失ったという情報に関する規則(14)、5)離床センサーを使用したという情報に関する規則(12)、という内訳だった。
- (2) 判定規則の性能検証： 判定規則を作成するのに用いたインシデントレポートとは別の、本院で半年間に報告されたインシデントレポートを性能検証用のデータとした。転倒に関連したインシデントレポートに判定規則を適用したところ、86.5%を正しく転倒に関連があると判定した。一方、転倒に関連しないインシデントレポートに判定規則を適用したところ、2.5%のみを誤って転倒に関連していると判定した。以上から、この判定規則は十分実用的であると判断した。
- (3) チャートレビュー： 本院に1ヶ月間の間に入院した全ての患者について、診療経過記録を逐一検閲し、記録に含まれたインシデントおよび有害事象を抽出した。チャートレビューは、検閲者によって検閲結果が異なることが知られていることから、判定基準を一定に保つために研究代表者が全ての判定を行った。
- (4) データ源の違いによるインシデント・有害事象検出能力の差： チャートレビューを行った期間と同じ期間に報告されたインシデントレポート、同じ期間に出された画像オーダエントリ、入院総括、診療経過記録について、上記で作成した判定規則を適用した。転倒の疑いとして検出された情報に関して、実際に転倒発生に関する情報が含まれているかどうか(感度)を検討した。診療記録はチャートレビューで発見できた全ての転倒事例を検出できた(図1)。ついで、インシデントレポート、画像オーダエントリの順に感度が優れていた。

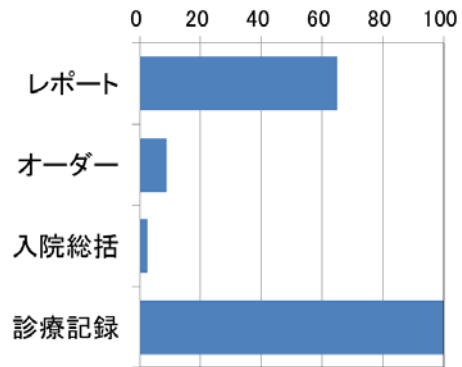


図1 各データ源を用いた場合に検出できた転倒の比率

一方、転倒の疑いと判定規則が判断したデータのうち、実際には転倒発生とは関係がないデータの割合は図2の通りだった。

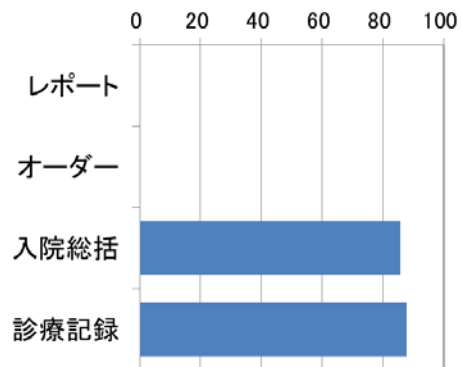


図2 転倒と判断されたが実際には転倒発生ではなかったデータの比率

インシデントレポートと画像オーダエントリの特異度が優れていた。実際に臨床で活用することを念頭に置いた場合、特異度が低いことは大きな欠点である。診療記録や入院総括は転倒事例検出のデータ源として不向きであると考えた。

- (5) データ源としてのインシデントレポートと画像オーダエントリの比較： 上記の結果からデータ源としては、特異度ついで感度に優れるインシデントレポートと画像オーダエントリ情報が有用であると考え、本院に入院した患者の1年間の診療情報を用いて、この二つの比較を行った。インシデントレポートの弱点の一つとして、事例発生から報告まで無視できないタイムラグがあることが報告されている。タイムラグをこの二つのデータ源で比較すると画像オーダエントリ情報の方が有意に短時間で病院情報システム上で検出できた(図3、

p<0.001、Wilcoxon 順位和検定)。

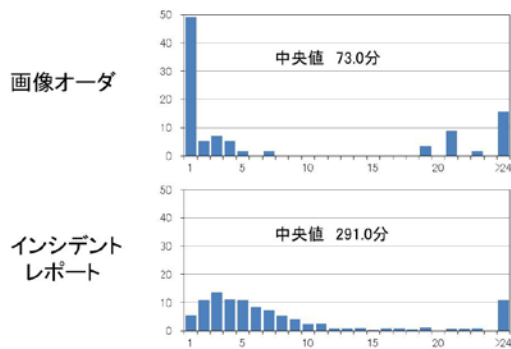


図 3 事例発生から検出までのタイムラグ

WHO の基準による事例の重大性・患者影響度は、有意に画像オーダエントリーの方がレベルが高かった (p<0.001、Fisher 正確検定)。

インシデントレポート単独の場合と比較して、画像オーダエントリー情報の NLP 処理を併用すると、検出できた転倒事例は 3%増えたに過ぎなかった。しかし、患者影響度が高い事例 (WHO の基準で moderate 以上) に限ると、画像オーダエントリーを併用することで、転倒事例の検出率は 23%増加した。したがって、画像オーダエントリー情報を NLP 処理して転倒事例を検出する方法は、重大事例を迅速に検出するという利点がある。

- (6) 応用： 以上に述べた方法で転倒事例を検出し、転倒後に重大な外傷 (骨折、頭蓋内出血) を起こすリスク因子を検討した。また、チャートレビューの結果と対比することで、インシデントレポートで報告されていない転倒事例を検出することができた。インシデントレポートで報告されない転倒事例には、どのような背景要因があるかを解析した。

5. 主な発表論文等

[雑誌論文] (計 2 件)

- (1) 鳥谷部真一、Use of risk assessment tool for inpatient traumatic intracranial hemorrhage after falls in acute care hospital setting、Global Journal of Health Science、査読あり、4、2012
- (2) 鳥谷部真一、World Health Organization fracture risk assessment tool in the assessment of fracture after falls in hospital、BMC Health Services Research、査読あり、10、2010、106

[学会発表] (計 4 件)

- (1) 鳥谷部真一、テキストマイニングを用い

た電子カルテからのインシデント情報検出、第 31 回医療情報学連合大会、2011 年 11 月 21 日、鹿児島

- (2) 鳥谷部真一、インシデントレポートで報告されない転倒事例の特徴、第 6 回医療の質・安全学会、2011 年 11 月 20 日、東京
- (3) 鳥谷部真一、頭蓋内出血を伴う転倒イベントのリスク評価、第 5 回医療の質・安全学会、2010 年 11 月 27 日、千葉
- (4) 鳥谷部真一、骨折を伴う転倒イベントのリスク評価、第 4 回医療の質・安全学会、2009 年 11 月 21 日、東京

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

6. 研究組織

(1) 研究代表者

鳥谷部 真一 (TOYABE SHIN-ICHI)
新潟大学・危機管理本部危機管理室・教授
研究者番号： 20227648

(2) 研究分担者

なし

研究者番号：

(3) 連携研究者

なし