

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 22 日現在

機関番号：17701

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25460845

研究課題名(和文) 医薬品相互作用検知システム用の有害事象シグナル検知アルゴリズムの評価に関する研究

研究課題名(英文) Study on evaluation of the algorithms for Adverse event signal detection with drug interaction detective system

研究代表者

村永 文学 (Muranaga, Fuminori)

鹿児島大学・医歯学域医学部・歯学部附属病院・講師

研究者番号：00325812

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：【背景】新薬の市販後調査時の有害事象のシグナルを早期に検知することは被害者を最小限にするために必須である。【目的】医薬品相互作用の有害事象のシグナル検知に最適なデータマイニング手法について調査する。

【方法】実際の処方記録を元に、薬歴シミュレータを開発し、薬歴データを10万件作成。アソシエーション分析、ベイジアンネットワーク分析、ニューラルネットワーク分析の有用性について検討した。

【結果及び考察】アソシエーション分析では、有害事象が発生しているレコードに絞り分析し発見できた。ベイジアンネットワーク分析では発見できなかった。ニューラルネットワーク分析では、想定併用薬の候補を発見することができた。

研究成果の概要(英文)：[Background] In the postmarketing surveillance (phase 4) of the new medicine, detecting the signal of the adverse event at postmarketing surveillance early, is essential to minimize the victim by the harmful effect. [Objectives] We investigate data mining technique that is optimal for the signal detection of the adverse event of the drug interaction. [Methods] We analyze a real prescription record, and develop a medication history simulator, and made 100,000 medication history data. We examined the utility of the association analysis, the Bayesian network analysis, the neural network analysis.

[Result and discussion] By the association analysis, we narrowed it down to the record which an adverse event developed and analyzed it and succeeded. By the Bayesian network analysis, we failed. By the neural network analysis, we were able to discover a candidate of the assumption combined medicine.

研究分野：医療情報学

キーワード：データマイニング 医薬品相互作用 医薬品有害事象 有害事象シグナル検知 データウェアハウス

1. 研究開始当初の背景

薬が市販される前の臨床試験(フェーズ3)は、モデル的な患者集団を対象とし、いわば「理想の世界」で行われるが、市販後の薬は「現実の世界、医療の現場」で使用される。また、小児、妊婦、高齢者に対する影響、および既に市販されている薬との相互作用は、市販後調査 PMS(Post Marketing Surveillance)によってのみ明らかになる事象といえ、市販後に早期に未知の副作用を検知することの重要性は、キノホルムによるスモンやゲフィチニブ(イレッサ)等の過去の経験からも明らかである。さらに近年は医療費削減策の一環として、ジェネリック薬の処方奨励され、実際にジェネリック薬を処方される症例が増えている。医薬品の商標が多様化することにより、医師の経験則に基づく相互作用の予見が困難になる恐れがある。

PMS の手法として、自発報告の中から優先的に検討すべき重要な adverse drug reaction(医薬品副作用。以下 ADR)の疑いを、統計的に選び出すシグナル検出法の研究が進んでいる。医療機関等は組織的に安全性の監視を行うことが不可欠である。これからの時代は、薬剤師や医師の知識と経験のみに頼るのではなく、処方データおよび検査結果値等から、データマイニングと呼ばれる統計的なシグナル検出法を用いてシステムティックに薬剤監査支援を行うのが主流となるだろう。

一方、知識発見の分野においては、形式化されていないテキスト情報(自由に記述されたという意味でフリーテキストとも呼ばれる)から一定の知識や発想を得る「テキストマイニング」と呼ばれる分析手法が注目されてきている。カルテの記載情報はまさにフリーテキスト形式で記録されており、その解析にはテキストマイニング技術の応用が期待される。

鹿児島大学医学部・歯学部附属病院では、過去に科学研究費補助金により医薬品相互作用による有害事象のシグナルを検知する為の薬剤疫学データウェアハウスシステムを開発した。本システムはアソシエーション分析のアプリオリアルゴリズムを採用し、有害事象を発生させた可能性のある相互作用薬の発見を可能とした。しかし、本システムの臨床応用段階において、コントロール症例群の薬歴と、調査症例群の薬歴に大きな乖離があった場合に、有意な相互作用薬が大量にリストアップされてしまう問題が残った。また、単剤で有害事象を発生させる可能性がある薬剤が薬歴に存在していた場合に事前にその薬剤を除外する方法論等については再検討の余地があった。

2. 研究の目的

本研究では、アソシエーション分析手法による相互作用薬発見機能の検知精度の向上を目指し、アプリオリアルゴリズム以外のデ

ータマイニングアルゴリズムとして、ベイジアンネットワークとニューラルネットワークについて比較検討を行った。

本研究の最終目標は、臨床的に自然発生するデータから、未知の医薬品相互作用による有害事象を自動的に発見し警告するシステムを開発することである。

3. 研究の方法

方法1) 処方薬歴の分析

2015年1月から2015年12月までに鹿児島大学病院の入院及び外来にて処方された薬剤について、その種類と分布、併用薬の数等について調査した。入院患者の持ち込み薬も含めた。すべての薬剤データには厚生労働省コードが付番されていたので、厚生労働省コードの頭8桁を用いて規格違いの同一成分薬は1種となるように工夫した。(例: ベイスン錠(0.2mg)のYJコード=3969004F1020, ベイスン錠(0.3mg)のYJコード=3969004F2027であるので、頭8桁の「3969004F」で統合した)

方法2) 方法1で調査した薬剤種と分布を近似して、ある特定の併用薬で相互作用が発生するような薬歴をシミュレーションするシステムを開発した。

方法3) 方法2で作成した薬歴データについて、アソシエーション分析(arulesライブラリ apriori)、ベイジアンネットワーク(dealライブラリ)、ニューラルネットワーク(nnetライブラリ)にて、相互作用を起こしている薬剤のペアを発見できるか、調査した。データマイニング手法のチューニング、データの前処理方法、後処理方法について検討した。加えて、検査結果に併用薬とは別の要因で有害事象が発生した場合(ノイズの混入)についても研究した。

本研究で使用するソフトウェアは、統計ソフト R ver3.3.0 の 64bit 版を利用した。データ処理手順の自動化に active perl ver.5.22.1 を利用した。

4. 研究成果

結果1)

鹿児島大学病院の処方薬歴の傾向分析について調査した結果について述べる。2015年1月から2015年12月までの1年間で、24,485名が処方を受けていた。1人の患者が複数の診療科で処方を受ける可能性があるが、1日で複数の診療科で処方を受けた場合も1件とカウントすると、1年間で206,043件の処方がなされていた。処方薬剤数は(YJコードによる規格違いを統合した結果)205,958件であった。1患者当たりの平均併用薬数は3.67剤(SD 3.60)であった。2015年の併用薬数の分布を図1に示す。

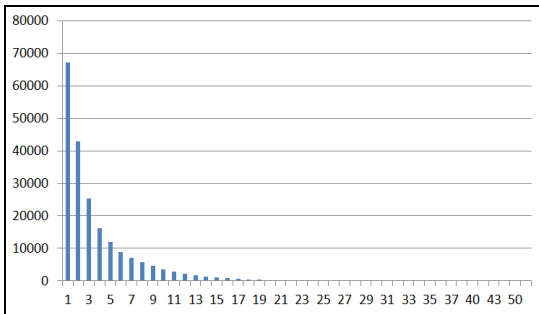


図 1 2015 年の併用薬数の分布

単剤処方が 33% , 2 剤併用が 21% , 3 剤併用が 12% , 4 剤併用が 8% , 5 剤併用が 6% , 6 剤併用が 4% , 7 剤併用が 3% , 8 剤併用が 3% , 9 剤併用が 2% , 10 剤併用が 2% , 11 剤併用が 1% であった。11 剤併用まで、全処方 of 95% を占めていた。

2015 年に 3 回以上処方されていた薬剤は 1554 種であった。薬剤種による処方量の偏りについても分析を行った。図 2 に薬剤種毎の処方量を示す。

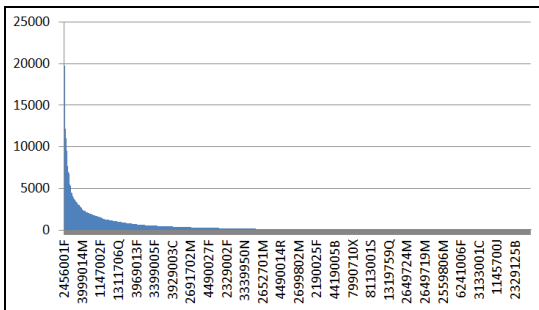


図 2 2015 年の処方薬剤種の分布

最も処方量の多い薬剤でも全体の 3% であり、2 位も 3%、3 位と 4 位が 2%、5 位以下は 1% 未満であった。全処方の 95% を占める処方種は 662 種であった。全診療科を有する大学病院では特定の薬剤にあまり偏ることなく処方されていたことが判明した。

結果 2)

結果 1 の結果を元に、薬歴シミュレータを作成した。10 万件の薬歴をシミュレートした結果について、分析を行った。併用薬の分布を図 3 に示す。2015 年の併用薬の分布 (図 1) と比較し、分布の形状が一致していることを確認できた。ただし、薬剤種に関しては、R システムで分析可能なデータ量の範囲内に収める為に、100 種以下の薬剤種が発生するようにした。

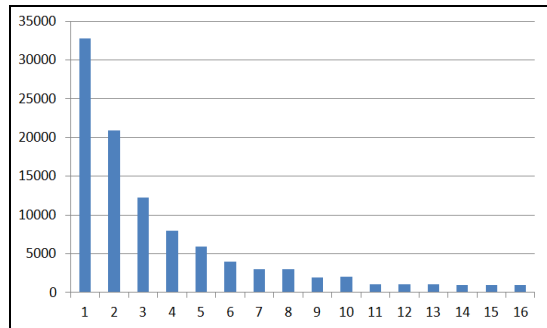


図 3 シミュレートした薬歴の併用薬の分布

薬剤種毎の分布を図 4 に示す。分布の形は、2015 年の薬剤種分布 (図 2) に類似する割合になっていることが確認できた。

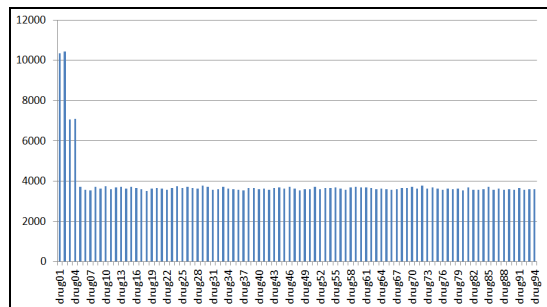


図 4 シミュレートした薬剤種の分布

なお、作成した薬歴シミュレーションでは、drug09 と drug10 を併用した場合に限り、白血球数低下を示唆する検査結果が発生すると想定した。10 万件の薬歴シミュレーションでは、drug09 と drug10 の処方出現率は 1% であった。シミュレートした薬歴における併用薬は平均 3.62 剤 (SD 3.40) であった。作成した薬歴データのうち、有害事象の発生率は $185/100,000=0.19\%$ であった。薬歴シミュレータでは、各データマイニング手法で分析しやすい構造で出力可能とした。アソシエーション分析用薬歴の部分サンプルを図 5 の 1 に、ベイジアンネットワーク分析用薬歴の部分サンプルを図 5 の 2 に、ニューラルネットワーク分析用薬歴の部分サンプルを図 5 の 3 に示す。10 万件の薬歴シミュレートに必要な時間は平均 12 秒 (SD 0.52s) であった。

No.	WBC	drug																		
1	WBC_N	drug19	drug81																	
2	WBC_N	drug86	drug59	drug03	drug01	drug18	drug62													
3	WBC_N	drug42	drug75	drug38	drug12															
4	WBC_N	drug40																		
5	WBC_N	drug73																		
6	WBC_N	drug87																		
7	WBC_N	drug78																		
8	WBC_N	drug57	drug74	drug54	drug53															
9	WBC_N	drug04	drug30	drug83	drug89	drug15	drug38													
10	WBC_N	drug81	drug46																	
11	WBC_N	drug29	drug84																	
12	WBC_N	drug36	drug47	drug05	drug92	drug34	drug71	drug39	drug22	drug01										
13	WBC_N	drug04	drug76	drug03	drug09	drug57	drug82	drug89	drug12	drug05										
14	WBC_N	drug93	drug55	drug03	drug19	drug81	drug07	drug64	drug32	drug28										
15	WBC_N	drug01	drug39	drug02																
16	WBC_N	drug65	drug29	drug50	drug92	drug90	drug08													
17	WBC_N	drug81																		
18	WBC_N	drug26																		
19	WBC_N	drug42	drug41	drug69	drug12	drug49	drug60	drug46												
20	WBC_N	drug93																		

図 5-1 アソシエーション分析用データセット例

3) ニューラルネットワーク分析の結果を示す。

3-1) 全件から探索

本研究では、ニューラルネットをデータでトレーニングし、白血球減少に対する各ニューロンの重みづけのパラメータを調査して、相互作用薬の発見を行った。10 万件、薬剤種 94 種のデータで探索を実行できた。図 7 にその結果を示す。

図 7 に示すように、中間ニューロン数 4 の場合に限り、相互作用想定薬剤である drug09 と drug10 の重みが他のニューロンと明らかに異なる結果となった。但し、同じ条件で別のデータセットを作成した場合に、相互作用想定薬剤を発見できなくなることも発生した。

処理時間は、中間ニューロン数が 4 の場合に平均 11 秒(SD 0.3s)であった。中間ニューロン数を 10 に増やすと処理時間が平均 56 秒(SD 2.1s)に延長した。10 万件の薬歴と 1 万件の薬歴の処理時間は、データ読み込みに必要な時間がデータ量に比例していたが、機械学習時間には大きな影響は無かった。

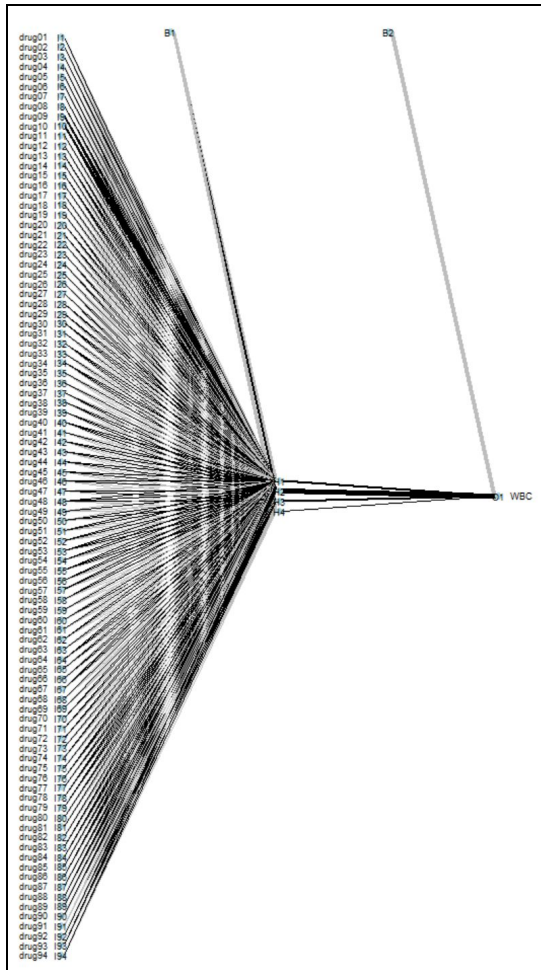


図 7 ニューラルネットワーク分析結果
(10 万件全件, 中間ニューロン=4)

中間ニューロン数の変化による各ニューロ

ンの重みづけの変化を、10 万件、薬剤種 10 種のデータで検証した。薬剤種を 10 種に限定したのは、視覚的に認知しやすくなるからである。

中間ニューロン数を 2~10 と変化させて機械学習させた結果を示す。(図 8 は中間ニューロン 3 の場合) 最初に準備したデータセットでは、中間ニューロン数 3~9 で相互作用想定薬剤である drug09 と drug10 を判別できた。別のデータセットを作成し直して分析した結果、中間ニューロン数が 4 の時のみ相互作用想定薬剤を判別できた。データセットによって適正なニューロン数にバラつきが見られた。

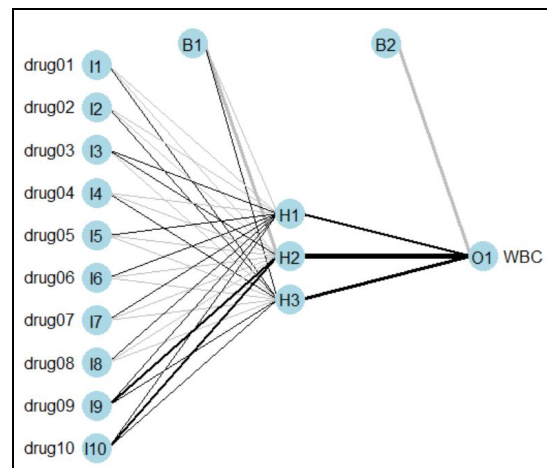


図 8 ニューラルネットワーク分析結果
(10 万件 10 薬剤種, 中間ニューロン 3)

考察

本研究によって、2つのアプローチが有用であることが判明した。

手法 1)

有害事象が発生しているレコードを絞り込み、アソシエーション分析(アプリアルゴリズム)を実施する。ただし、統計解析システム R の apriori では 100 薬剤種以下が限界であった。アソシエーション分析の特質上、薬剤種が 1 つ増える度に、計算量が倍増することになる。ある程度、有害事象や薬剤にある程度狙いをつけて絞り込んだ後に探索する必要がある。

手法 2)

ニューラルネットワークにおいて、中間ニューロン数を調整して探索する。薬剤種が極端に増えた場合においても、探索時間が極端に増加しないのは大きな魅力であると思われる。中間ニューロン数については、データセットの偏り等によって、適切な値が変化することが判明した。

近年、ディープラーニングと呼ばれる、多層構造のニューラルネットワークによって、

画像認識や問題解決の分野の研究が加速している。本研究の結果から、単層ニューラルネットワークが、医療ビッグデータに対する大規模データのマイニングに有用であることが示唆された。

今後は多層ニューラルネットワークによる、医薬品相互作用のシグナル検知システムについて、更に研究を進めていきたい。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計4件)

【DWH 構築が果たす効果検証】 総論 DWH の活用における今日的課題と将来展望 鹿児島大学病院の事例を踏まえて; 熊本 一郎, 村永 文学; 新医療, 43 巻 2 号, Page24-27(2016.02)

The significance of a Bayesian Network in Type 2 Diabetes Mellitus: Nurjannah, Fuminori Muranaga, Takashi Iwaanakuchi, Yumiko Uto, Ichiro Kumamoto; Journal of Joint conference on Medical Informatics, Vol.34 Suppl. Page 352-355 (2014.11)

診療記録の記載事項より誤嚥性肺炎の発症を予見する取り組みとその評価 データマイニングを用いて: 小牧 祥太郎, 村永 文学, 川原 浩, 山下 美穂, 岩穴口 孝, 宇都 由美子, 熊本 一郎, 福田 ゆかり; 医療情報学連合大会論文集第 34 巻 Page 750-753 (2014.11)

医薬品相互作用検知システム用の有害事象シグナル検知アルゴリズムの評価: 村永 文学, 岩穴口 孝, 宇都 由美子, 熊本 一郎; 医療情報学連合大会論文集 第 36 巻 Suppl. (掲載予定)

[学会発表](計3件)

The significance of a Bayesian Network in Type 2 Diabetes Mellitus: Nurjannah, Fuminori Muranaga, Takashi Iwaanakuchi, Yumiko Uto, Ichiro Kumamoto; The 34th Joint conference on Medical Informatics (千葉県千葉市幕張メッセ 2014.11)

診療記録の記載事項より誤嚥性肺炎の発症を予見する取り組みとその評価 データマイニングを用いて: 小牧 祥太郎, 村永 文学, 川原 浩, 山下 美穂, 岩穴口 孝, 宇都 由美子, 熊本 一郎, 福田 ゆかり; 第 34 回医療情報学連合大会 (千葉県千葉市幕張メッセ 2014.11)

医薬品相互作用検知システム用の有害事象シグナル検知アルゴリズムの評価: 村永 文学, 岩穴口 孝, 宇都 由美子, 熊本 一郎; 第 36 回 医療情報学連合大会 (神奈川県横浜市パシフィコ横浜にて発表予定)

[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]

特記事項無し

6. 研究組織

(1)研究代表者

村永 文学 (MURANAGA Fuminori)

鹿児島大学・医歯学域医学部・歯学部附属病院・講師

研究者番号:00325812

(2)研究分担者

熊本 一郎 (KUMAMORO Ichiro)

鹿児島大学・医歯学域医系・教授

研究者番号:40225230

宇都 由美子 (UTO Yumiko)

鹿児島大学・医歯学域医系・准教授

研究者番号:50223582

岩穴口 孝 (IWAANAKUCHI Takashi)

鹿児島大学・医歯学域医学部・歯学部附属病院・助教

研究者番号:80619198