

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25年 6月 6日現在

機関番号：14401

研究種目：基盤研究（C）

研究期間：2009～2012

課題番号：21590157

研究課題名（和文） データマイニングを用いた、医薬品併用による有害事象の発見

研究課題名（英文） Study on Adverse Events of Drugs using Data Mining

研究代表者

高木 達也（TAKAGI TATSUYA）

大阪大学・薬学研究科(研究院)・教授

研究者番号：80144517

研究成果の概要（和文）：

FDAの有害事象報告データベース(AERS)2004年版データを利用し、まず相関ルールを用い、報告された薬剤を患者単位で二剤ずつ組み合わせることによって、被疑薬の中でも重要な薬剤をピックアップできるような新たなルールを構築した。次にLogistic回帰(LRA)、Partial Least Squares(PLS)を用いて、薬剤の化学的部分構造とSJSとTENとの関連をモデル化した。

研究成果の概要（英文）：

The database constructed by the FDA Adverse Event Reporting System (AERS) was used in order to find some drugs which are more likely to cause Stevens - Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN) using data mining methods. Especially, we utilized association rule applying to the combinations of drugs in each patient. As a result of the mining, we constructed a new process to find significant suspect drugs. In addition we also revealed some chemical features which frequently exist in the suspect drugs.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	1,400,000	420,000	1,820,000
2010年度	1,200,000	360,000	1,560,000
2011年度	1,000,000	300,000	1,300,000
総計	3,600,000	1,080,000	4,680,000

研究分野：医歯薬学

科研費の分科・細目：薬学・医療系薬学

キーワード：

AERS、Stevens-Johnson 症候群、シグナル検出、データマイニング、ノンパラメトリック回帰、相関ルール

1. 研究開始当初の背景

Stevens-Johnson 症候群、Lyell 症候群（以下、併せて SJS と略す）は、そのかなりの部分が医薬品の副作用と思われ、死亡率が低くない（10-40%との報告例あり）、皮膚、粘膜の炎症、壊死を伴う副作用で、回復後も後遺症率が高いものの、多数の医薬品で惹起されることが報告されているため、報告のあった医薬品全てを販売中止することは現実的

ではない。このため、医薬品の適正使用という観点から、SJS の発生をでき得る限り低下させることが期待されている。

既に多数の疫学的研究により、SJS を惹起しやすい医薬品は幾つか明らかにされている。また、遺伝子的には、HLA-B75、HLA-B58などが、重要な serotype として挙げられている。しかしながら、これらの報告例に当てはまらない～serotype も服用医薬品も～少

なくなく、さらなる疫学研究が期待されている。

2. 研究の目的

多くの場合、患者は単剤を投与されるのではなく、複数の医薬品を併用している。単剤ではSJSを惹起する率は低いものの、幾つかの医薬品の併用がSJSを惹起する可能性があり、このようなケースは、従来の分子生物学的、及び疫学的研究では発見できなかったもので、高度な統計的、機械学習的手法による研究が不可欠である。

本研究計画では、相関ルール、新規非加法型ノンパラメトリック判別分析法を用い、主としてSJSを取り上げ（その他、抗悪性腫瘍剤、抗真菌剤等の副作用も取り上げる）、医薬品の併用により出現する重篤な副作用を未然に防ぐため、単剤を目標とした疫学研究では発見の困難な事例について、高度な解析手法により発見するとともに、やはりデータマイニング手法を用いて、多形滲出性紅斑（EM）とSJS、TENの報告例の比を考え、よりSJS、TENを惹起しやすい化学構造について解明することを目的とする。

3. 研究の方法

本研究計画では、相関ルール、新規非加法型ノンパラメトリック判別分析法を用い、主としてSJSを取り上げ（その他、抗悪性腫瘍剤、抗真菌剤等の副作用も取り上げる）、医薬品の併用により出現する重篤な副作用を未然に防ぐため、単剤を目標とした疫学研究では発見の困難な事例について、高度な解析手法により発見することを目的とする。

相関ルールは、社会科学、金融学方面で広く使用されている手法であり、2つ以上の事象（例えば服用医薬品）の間に何らかの相互作用があり、結果として（例えば有害事象として）現れる可能性を見つけ出す手段である。このデータマイニング手法を用いることにより、単剤ではSJSを発症しないが、併用により発症する危険性のある薬剤の組み合わせを見いだすことができる。

4. 研究成果

(1) 序論

本研究ではFDAの公開データベースの中のAdverse Event Reporting System (AERS)[1]の2004年版データを利用し、association rule、Logistic Regression Analysis (LRA)を適用することにより、医薬品の重篤な副作用の一つであるStevens Johnson 症候群(SJS)、Toxic epidermal necrolysis(TEN)、Erythema multiform (多型紅斑: EM)に関する危険性を医薬品サイドから検討し、報告された薬剤を患者単位で二剤ずつ組み合わせることによって、被疑薬の中でも重要な薬剤をピックアップできるような新たなルールを

構築した。SJSまたはTENの研究で予後に焦点を当てた研究は筆者の知る限り報告されていない。E2Eにおいては“Data mining techniques facilitate the evaluation of spontaneous reports by using statistical methods to detect potential signals for further evaluation. This tool does not quantify the magnitude of risk, and caution should be exercised when comparing drugs.”[2]と述べられているが、本研究計画では、添付文書内でSJS、TENの記載がある薬剤について予後を解析することによって、同種同効薬の中でリスクの比較を行い、シグナル検出から一歩踏み出した解析と言える。

また、同じ皮膚病変を伴う重篤な副作用でも、EMにとどまる傾向のある薬剤と、SJS/TEN傾向の薬剤との違いを、医薬品の化学構造の面から研究し、一定の成果を得ることができた。

(2) 方法

データセットは、AERS ファイルから Microsoft Access 2003を使用することによって抽出し、そして、Microsoft Excel 2003を使用してデータの統計解析を行った。すべてのグラフ描画にはSpotfireを使用した。

① 二剤結合型

下式で表わされる相関ルールを適用する。

$$X_1 \Rightarrow Y_1$$

X_1 : 薬剤 D_1 と D_2 を併用する

Y_1 : SJS を発症する

ここではこのルールを“二剤結合型”と称する。今回は各々の薬剤の支持度と確信度については、SJSを発症した患者を含むデータセットからすべての薬剤の組み合わせについて計算される。上記ルールの支持度と確信度は以下のようになる。

$$\text{sup}(X_1) = \frac{D_1 \cap D_2}{T}$$

$$\text{sup}(X_1 \Rightarrow Y_1) = \frac{(D_1 \cap D_2) \cup Y_1}{T}$$

$$\text{conf}(X_1 \Rightarrow Y_1) = \frac{\text{sup}(X_1 \Rightarrow Y_1)}{\text{sup}(X_1)}$$

ここでTはケース数であり、確信度は具体的にはSJSの発症例の中で薬剤 D_1 と D_2 を同時に使用している割合である

② 二剤分離型

同様に、

$$X_2 \Rightarrow Y_2$$

X_2 : 薬剤 D_1 を使用した患者が SJS を発症した。

Y_2 : 薬剤 D_2 を使用していた。

とすると、

$$\text{sup}(X_1) = \frac{D_1 \cap D_2}{T}$$

$$\text{sup}(X_1 \Rightarrow Y_1) = \frac{(D_1 \cap D_2) \cup Y_1}{T}$$

$$\text{conf}(X_1 \Rightarrow Y_1) = \frac{\text{sup}(X_1 \Rightarrow Y_1)}{\text{sup}(X_1)}$$

(3)結果

まず、DRAGONにより算出された3224個の記述子のうち、分散が0の記述子を除去した。次に、説明変数間に強い相関があると、得られた結果の信頼性が低くなることから、記述子間の相関係数の絶対値が0.7以上である組み合わせは目的変数との相関係数が高い方の記述子のみを残した。残った記述子は245個となる。

次に、重み付きロジスティック回帰を用い、化学記述子を用いて、EM傾向、SJS/TEN傾向の薬剤の分類を試みた。重みをつけたのは、EM傾向の薬剤の数とSJS/TEN傾向の薬剤の数が異なるからである。その結果、外部バリデーションに於いても、平均70%を超える正答率が得られた。

Table 1. 本プロジェクトで使用した医薬品データ数

SJSでのみ報告のあった医薬品 ^{*1}	59
EMでのみ報告のあった医薬品 ^{*2}	37

Table 2. ロジスティック回帰分析の結果

CV正答率(%)(平均±SD)	86.7±3.1
モデル構築に使用した記述子の数(平均±SD)	15.8±6.4

Fig. 1は、正答率が86%前後であったCross Validation法の結果を、目的変数を横軸に、予測値を縦軸にプロットした図である。Fig. 1に示されるように、目的変数が1の予測がより正確になされていた。

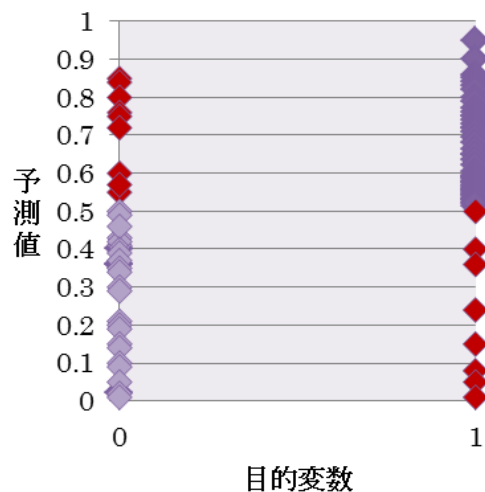


Fig.1 重み付きロジスティック回帰分析の予測値-実測値プロット

Table 3. 有意な説明変数

	標準偏回帰係数	P値
nCq	-0.8296	0.004296
JGI6	0.7650	0.01427
B08[C-C]	0.7623	0.002857

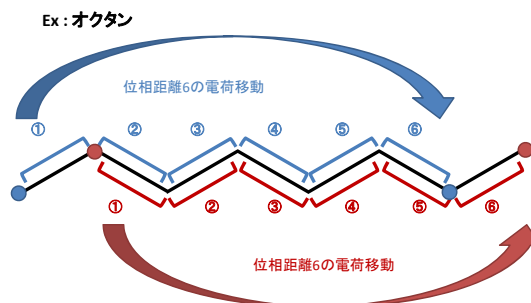


Fig. 2 JGI6 概要図

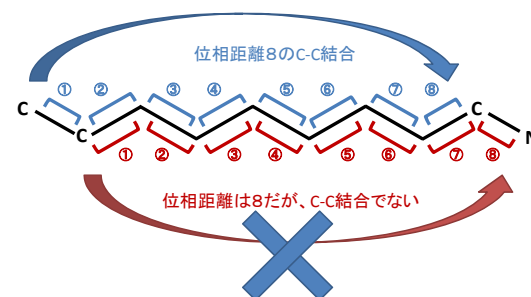


Fig. 3 B08[C-C]の概要図

Table3のうち、nCqは4級炭素の数で、4級炭素がない方がEM傾向であることを示して

いる。B08 と JGI6 は、いずれも炭素鎖の長さが関連していると考えられ、長い炭素鎖があるものほど SJS/TEN 傾向が強いと考えられる。

Table 4 PMDA で報告の多い推定原因医薬品の変数の数値

	nCq	JGI6	B08[C-C]
Acetaminophen	0	0.054	0
Allopurinol	0	0	0
Azithromycin	0	0.023	1
Carbazequine	0	0.021	0
Cefcapene pivoxil	1	0.022	1
Cefdinir	0	0.02	1
Clarithromycin	0	0.022	1
Diclofenac	0	0.014	1
Levofloxacin	0	0.028	1
Loxoprofen	0	0.022	1
Phenobarbital	1	0.028	0
Phenytoin	0	0.016	1
Salicylamide	0	0	0
Zonisamide	0	0.021	0

Table 4 に、国内で SJS/TEN の報告例の多い医薬品をあげ、その有意で残った説明変数の値を表示した。概ね、重み付きロジスティック回帰の結果から、SJS/TEN 傾向であることが予想されることが示されている。

これらの結果は、今後、医薬品開発に於いて、重要な情報となることが期待される。

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 2 件)

1. Y. Shirakuni, K. Okamoto, N. Kawashita, T. Yasunaga, T. Takagi, "Signal Detection of Drug Complications Applying Association Rule Learning for Stevens-Johnson Syndrome" *Journal of Computer Aided Chemistry*, **10**, 118-127 (2009).
DOI: 10.2751/jcac.10.118
2. Y. Shirakuni, K. Okamoto, E. Uejima, S. Inui, J. Takahara, T. Ohgaru, H. Yamasaki, Y.-S. Tian, N. Kawashita, R. Inoue, T. Yasunaga, T. Takagi, "A practical estimation method for analyzing adverse drug reactions using data mining," *Therapeutic Innovation & Regulatory Science*,

47(2), 235-241 (2013).

[学会発表] (計 5 件)

1. 米倉聡、岡本晃典、川下理日人、日高伸之介、山崎広之、川瀬雅也、高木達也、「新規ノンパラメトリック回帰手法の応用」、第 32 回情報化学討論会(宇部)、2009.
2. Y. Shirakuni, E. Uejima, J. Takahara, K. Okamoto, N. Kawashita, T. Yasunaga, S. Inui, T. Takagi, "Statistical Study on Chemical Risk Factors for SJS/TEN," 9th ISoP Annual Meeting. 2009, Reims, France
3. 高木達也、「多変量解析法を用いた疾患・副作用の予後予測支援」、日本薬学会第 130 年会(岡山)、S12-5 (招待講演)、2010
4. 井上藍、白國優子、川下理日人、岡本晃典、高木達也、「FDA データベースを用いたデータマイニング」、第 33 回情報化学討論会(徳島)2010.
5. T. Takagi, "Prediction of the Likelihood of Adverse Reactions of Medicines using Chemical Properties," ICCA-LRI and Health Canada Symposium, 2011, Quebec-city (invited lecture).
6. 研究組織
 - (1) 研究代表者
高木 達也 (TAKAGI TATSUYA)
大阪大学大学院薬学研究科・教授
研究者番号：80144517
 - (2) 研究分担者
上島 悦子 (UEJIMA ETSUKO)
大阪大学大学院薬学研究科・教授
研究者番号：80432441

川下 理日人 (KAWASHITA NORIHITO)
大阪大学大学院薬学研究科・助教
研究者番号：00423111

岡本 晃典 (OKAMOTO KOUSUKE)
大阪大学大学院薬学研究科・助教
研究者番号：70437309