

IV. 研究成果の刊行物・別刷

副作用評価におけるシグナル検出

藤田 利治*

Signal Detection of Adverse Drug Reactions

Toshiharu FUJITA *

*The Institute of Statistical Mathematics
4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan

〈Abstract〉

During the post-approval period, hypotheses about potentially new adverse drug reactions (ADR) have traditionally emerged from passive surveillance systems that collect large volumes of spontaneous case reports of suspected adverse drug reactions. With signal detection by traditional (or conventional, or manual) methods, quantitative (or statistical, or automated) methods for spontaneous reporting system (SRS) databases were introduced in the late 1990's in order to detect serious ADR as early as possible. Most quantitative methods rely on comparisons of relative reporting frequencies, also known as disproportionality analyses. In FY 2009, the Pharmaceuticals and Medical Device Agency (PMDA) plans to introduce the quantitative methods (data mining method) used on Japanese SRS database. This paper introduces the recent situation on signal detection and signal management of adverse drug reactions.

Key words : signal detection, data mining method, spontaneous reporting system, post marketing surveillance, signal evaluation

はじめに

副作用の受動的サーベイランス(自発報告制度)は、未知の副作用を検出するシステムとして、薬剤の安全確保に大きな役割を果たしてきた。近年、世界的に薬剤の安全確保システムの強化が加速している。国際医学団体協議会(CIOMS)では2006年に副作用シグナルの検出に関するWorking Group VIII(CIOMS VIII)を設置し、「ファーマコビジランスにおけるシグナル検出とシグナルマネジメント」の報告を取りまとめつつあり、近く公表される予定である。日本でも独立行政法人医薬

品医療機器総合機構(PMDA)が企業等から収集した副作用等情報に対してデータマイニング手法を用いたシグナル検出を2009年度から安全対策業務に導入すべく検討を進めており、そのために「データマイニングに関する検討委員会」が設置されている。筆者はこの検討会の委員を務めており、また、外部専門家としてCIOMS VIII報告書(案)のレビューを行う機会を得た。この講演では、CIOMS VIII報告書(案)を参考にし、またPMDAでのデータマイニング手法導入についても触れながら、副作用評価におけるシグナル検出について紹介する。

* 情報・システム研究機構 統計数理研究所 〒106-8569 東京都港区南麻布4-6-7

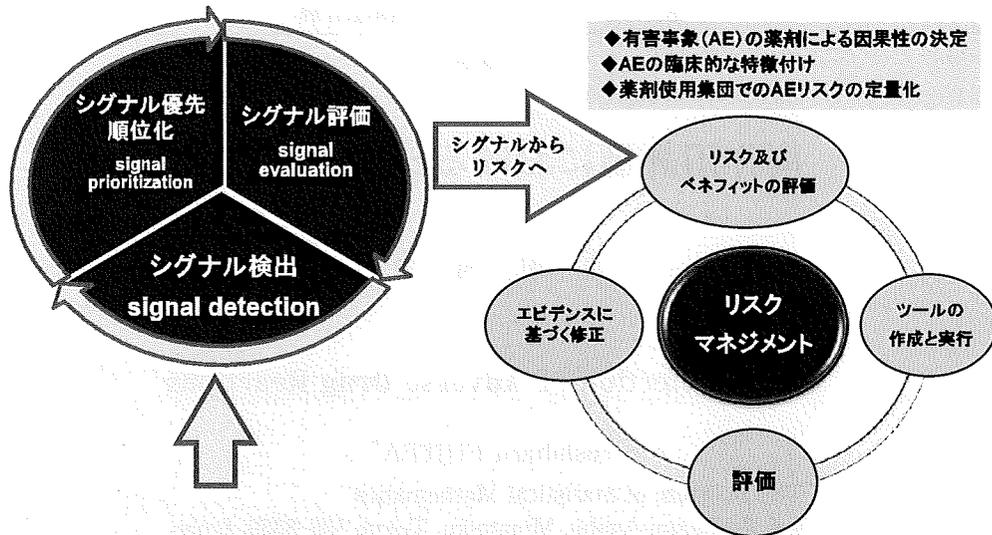


図1 シグナルマネジメントとリスクマネジメント

受動的サーベイランスに基づく シグナル検出

世界保健機関 (WHO) の定義では、シグナルとは「それまで知られなかったか、もしくは不完全にしか立証されていなかった薬剤と有害事象との因果関係の可能性に関する、さらに検討が必要な情報」である¹⁾。CIOMS VIIIでは「薬剤安全性シグナルのライフサイクル全体での、マネジメントのための系統的かつ全体的な戦略を確立する」ことを目指している。「シグナル検出」はシグナルマネジメントの入り口であり、「シグナルの優先順位化」「シグナル評価」というステップがその後に続く (図1)。そして、薬剤との因果性が確認されたリスクとなった場合には、リスクマネジメントのサイクルへと繋がっていく。

シグナル検出の目標は、

- ・新しい安全性シグナル
- ・特定の安全性シグナル (前臨床で懸念される事項など)
- ・特定集団での安全性シグナル
- ・既知のリスクについての頻度、重症度や他の特性 (性、年齢の分布) の変化のシグナルなどの検出である。使用する情報源として、世界的には、
- ・受動的サーベイランス (自発報告制度)

- ・集団についての大規模なデータベース (medical claims databases, electronic health record databases)
 - ・特別なモニタリングプログラム (全例調査など)
 - ・積極的サーベイランス (拠点医療機関、処方イベントモニタリングなど)
- などがある。CIOMS VIIIにおいては、受動的サーベイランスを情報源とするシグナル検出等の検討が中心となっている。

受動的サーベイランスの長所としては、

- ・薬剤と患者の広い範囲をカバー
- ・実際の臨床での使用実態を反映
- ・実施が安価で単純
- ・未知・重篤な副作用を検出するための最も効果的な手段

などを挙げることができる。受動的サーベイランスは未知の重篤な副作用を検出するための最も強力な手段であり、ファーマコビジランスでの必須のシステムである。安全性問題で撤退した薬剤の36~50%は、症例報告のみが根拠との報告がなされている²⁻⁴⁾。しかしながら、

- ・過少報告
- ・不安定な報告割合
- ・詳細な臨床情報の不足
- ・発生頻度の分母となる直接的な情報の欠如のため、発生頻度の算出が困難

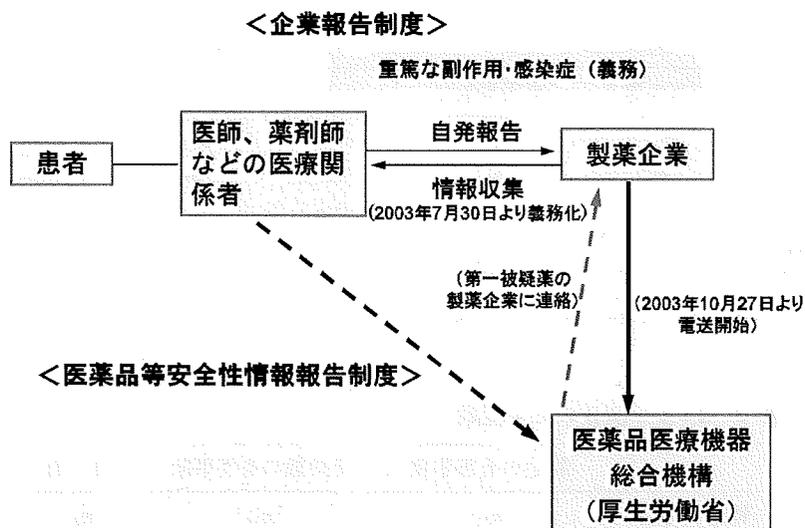


図 2 日本の受動的サーベイランス制度

- ・ 遅発性の有害作用の検出を見逃す可能性
- ・ 背景での発生率が低いとはいえない疾患を見逃す可能性（例：高齢者での虚血性心疾患，悪性新生物など）

などの問題があることはよく知られている。とくに集団での発生率が稀とは言えない有害事象を受動的サーベイランスに基づいて発見することは困難である。有害事象の臨床のおよび定量的多様性から，多数の方法やデータの使用が最適なシグナル検出には必要であり，ひとつのアプローチ・情報源への過度の過信は避けるべきである。大規模な集団についてのデータベースや積極的サーベイランスといった情報源を実際的に欠いている日本の状況の改善は，薬剤の安全確保のための喫緊の課題となっている。

日本における受動的サーベイランスは，医薬品等安全性情報報告制度と企業報告制度からなっている（図2）。これらの制度では，副作用の合理的な可能性のある有害事象の報告が想定されている。医薬品等安全性情報報告制度からの重篤な事象については第一被疑薬の製薬企業への連絡がなされ，企業による詳細な情報収集が行われる。重篤な副作用については企業報告制度が医薬品等安全性情報報告制度を含んでいる形になっている。

従来，自発報告の評価は，医学・薬学の専門の知識を持った評価者による専門的な評価という伝

統的方法で実施されてきた。一方，自発報告のデータに対する統計的アプローチは過去においては低調であった。その理由は，上述の受動的サーベイランスの短所と関連している。副作用の発生頻度を算出することができず，過少な報告しかなされず，しかも系統誤差（バイアス）が含まれる可能性の高いノイズの大きなデータに対して，統計的推測を行ったとしても妥当な推論など行えない，との自然な発想によるところが大きかった。しかしながら，マーケティング，金融などのビジネスの世界でノイズの大きな大規模データベースから意味のある情報を抽出するデータマイニングが注目され，さまざまな分野への応用がなされるようになった。データマイニングとは，期待されるよりも高い頻度で発生する首尾一貫した強い関連を検出するために，大規模なデータベースを探索する統計的技術の総称である。

健康にかかわる情報を蓄積して生きている人のために活用することについての社会的合意が伝統的になされている欧米諸国においては，副作用の自発報告についても過去からの膨大な情報の蓄積がなされてきたことから，データマイニングによる副作用のシグナル検出の情報基盤が存在していた。また，膨大な数の自発報告について個々に専門的評価を行うことは極めて大きな労力を要し，また重大な副作用の見逃しの危険の増加も懸念さ

表1 シグナル検出の元となるデータ

	有害事象1	有害事象2	……	有害事象 p	合計
薬剤1	n_{11}	n_{12}	……	n_{1p}	n_{1+}
薬剤2	n_{21}	n_{22}	……	n_{2p}	n_{2+}
⋮	⋮	⋮	……	⋮	⋮
薬剤 m	n_{m1}	n_{m2}	……	n_{mp}	n_{m+}
合計	n_{+1}	n_{+2}	……	n_{+p}	n_{++}

表2 2×2分割表のセル度数

	特定の有害事象	その他の有害事象	合計
特定の薬剤	n_{11}	n_{12}	n_{1+}
その他の薬剤	n_{21}	n_{22}	n_{2+}
合計	n_{+1}	n_{+2}	n_{++}

れる。こうした状況において、自発報告の大規模データベースを活用して、統計的手法を用いて副作用シグナルを検出するデータマイニング手法の検討が進められ、専門的評価を支援するシステムが実用化された。

しかし、自発報告のデータベースは分母情報(薬剤の使用人数や人年)を欠いたものである。したがって、信頼性のある因果推論は困難なことが多く、他の方法により確認されるべき仮説提供というのが本来の位置づけである。

シグナル検出の方法

自発報告データベースを用いたシグナル検出は、「伝統的方法」(traditional, conventional, manual)と「データマイニング手法による方法」(statistical, automated)に分けることができる。伝統的方法では、

- ・リチャレンジによる確認
- ・特定医学事象 (Designated Medical Events) (アナフィラキシー, Stevens-Johnson 症候群, 横紋筋融解症, 無顆粒球症など)
- ・標的医学事象 (Targeted Medical Events) (特定の薬剤や患者集団に関係する事象)
- ・薬剤曝露の直後の発現
- ・ゼロに近い背景での3例以上の症例

などが暗黙のうちに重視されてきた。一方、データマイニング手法では、非比例的な報告パターンの自動的・網羅的スクリーニングがなされる。伝統的な方法とデータマイニング手法による方法では異なる側面・方向から副作用シグナルの検出を目指しているものであり、いずれが他の方法に替わりえるものではない。

データマイニング手法によるシグナル検出のもととなるデータは、行に薬剤を列に有害事象を取り、その報告件数を度数とする表1のような度数表である。表1が通常の統計分析で扱われる度数表と異なる点は、薬剤の種類および有害事象の種類が極めて多く、表のセル数が膨大であることである。たとえば、日本で市販されている12,000超の薬剤と17,000超の有害事象 (MedDRAのPT) では2億を超える組合せ数(セル数)であり、個々のセルの度数は小さく、1件や2件といったセルがかなりの割合を占めることになる。表1における特定の薬剤と特定の有害事象に注目すると、表2のような2×2分割表ができる。こうしたデータを用いるデータマイニング手法にはいくつかの仮定がある。

- ① ある薬剤が特定の有害反応を起こすならば、他の薬剤と比べて、この有害反応が多く報告される

② 同じ有害反応について、(過少) 報告の程度は異なる薬剤間で同じである

③ 種々の有害反応の報告割合や報告パターンは、異なる薬剤間で類似している

しかし、これら(①~③)の仮定は弱いものであり、多くの反する例がある。たとえば、Weberカーブ、自発報告の強化、メディアの注目、緊急安全性情報などにより仮定が成り立たない状況が多々起きる。シグナル検出のデータマイニングには、ロバストな手法が求められる。検定統計量や区間推定などがデータマイニング手法でも用いられるが、それらは検証を目的とするものではなく、単なるスクリーニングの目安として用いているに過ぎない。

シグナル検出のデータマイニング手法として代表的なアルゴリズムである Proportional Reporting Ratios (PRR)^{5,6)}の指標は、報告割合の比である。

$$PRR = \frac{n_{11}/n_{1+}}{n_{21}/n_{2+}} = \frac{P_{11}}{P_{21}}$$

通常、以下の3つの条件を満たす場合(閾値)に、「シグナルあり」と判断される。

① $PRR \geq 2$

$$\textcircled{2} \chi^2 = \frac{n_{++} + (|n_{11}n_{22} - n_{12}n_{21}| - n_{++}/2)^2}{n_{1+}n_{2+}n_{+1}n_{+2}} \geq 4$$

③ $n_{11} \geq 3$

また、オッズ比である ROR (Reporting Odds Ratio) はオランダで使用されており、理論的には PRR よりもバイアスの少ない方法であることが指摘されている⁷⁾。

$$ROR = \frac{(n_{11}/n_{21})}{(n_{12}/n_{22})} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

ベイズ流のアプローチを用いる手法として、WHO Uppsala Monitoring Centre で用いられている Bayesian Confidence Propagation Neural Network Method (BCPNN)⁸⁾と、米国 FDA で使用されていた Gamma-Poisson Shrinker (GPS) Program⁹⁾があり、薬物相互作用を考慮した MGPS (Multi-Item Gamma Poisson Shrinker) が現在は使用されている。これらの4つの手法での指標は、表2の n_{11} という観察値 0 に対して、わ

ずかに異なる期待値 E を用いていることに相当し、次の関係にある。

• Proportional reporting ratio (PRR)

$$\frac{n_{11} \times n_{2+}}{n_{21} \times n_{1+}}$$

• Reporting odds ratio (ROR)

$$\frac{n_{11} \times n_{22}}{n_{12} \times n_{21}}$$

• Information Component (IC)

$$\log_2 \frac{n_{11} \times n_{++}}{n_{1+} \times n_{+1}} \quad \text{BCPNN は、これについてのベイズ流のアプローチ}$$

• Relative reporting (RR)

$$\frac{n_{11} \times n_{++}}{n_{1+} \times n_{+1}} \quad \text{GPS は、これについてのベイズ流のアプローチ}$$

ベイズ流アプローチでは、期待値が小さい場合、O/E は「1」の方向に縮小する。図3は、米国の AERS の 2001 年のデータを用いて 26,359 の薬剤・事象の組合せに対して各データマイニング手法を適用して検出されたシグナルの結果を、期待値(横軸)と各指標値(縦軸)で図示したものである。PRR と ROR では期待値の小さなシグナルの指標値に極めて大きなものがあり、偶然変動の大きな影響を受けていることが分かる。これに対して、ベイズ流アプローチの BCPNN (指標: IC) と GPS (指標: RR) では期待値が小さい場合に指標値が縮小して変動が抑制されることが示されている。なお、ROR で 64% という極めて多数のシグナル量になっているのは、PRR と異なり、 $n_{11} \geq 3$ という条件を使用していないためである。

データマイニング手法の実用化の状況

データマイニング手法の実用化とそのデータベースの状況は、次のとおりである。

- WHO-UMC (Uppsala Monitoring Centre) : メンバーの 83 カ国からの自発報告を集積した Vigibase (2007 年末において約 390 万件) において、BCPNN を使用
- 米国 FDA : AERS (Adverse Event Reporting System) に蓄積された約 400 万件 (2007 年末) の自発報告に対して、MGPS を使用。

指標値 1,000以下、期待値 8以下について表示

2001年AERSの15日報告データを使用。26,359の組合せ種類

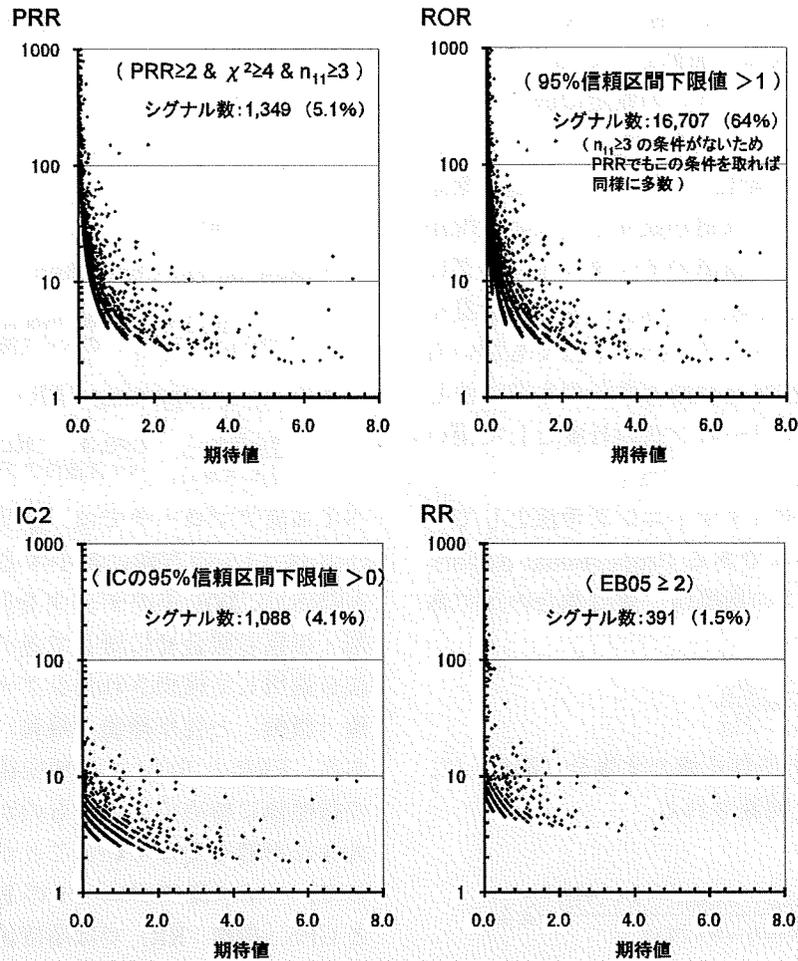


図3 検出されたシグナルの n_{11} の期待値と指標値

- (以前は, GPS)
- ・ EU: EU 諸国からの自発報告を集積した Eudra Vigilance において, PRR を使用
 - ・ 英国 MHRA: Yellow Card データベースである SENTINEL system (2007 年末において約 57 万件の自発報告) において, MGPS を使用.
(以前は, PRR)
 - ・ オランダ: Lareb データベース (2007 年末において約 6 万件の自発報告) において, ROR を使用

日本における データマイニング手法の導入

企業報告制度での電送による報告が 2003 年 10

月 27 日から開始され, 日本においても自発報告データベースの整備が進められている. 報告を受けている医薬品医療機器総合機構 (PMDA) では, 2009 年度からデータマイニング手法を安全対策業務に導入すべく, さまざまな検討を進めており, その報告書が Web サイト上で公表されている.

http://www.info.pmda.go.jp/kyoten_iryaku/dm.html

検討結果として, たとえば, PMDA のデータベース (2003 年 10 月 27 日~2005 年 9 月 30 日の期間の約 5 万件の報告) を使用したシグナル検出量を, 表 3 に示した. 第一被疑薬については, ROR では 21,384 のシグナルが検出され, これはすべての薬剤・事象の組合せの 60.9% に当たる量

表 3 PMDA におけるシグナル検出量の検討結果

基準	シグナル検出指標	2005/9/30 時点のシグナル 検出数		2005/9/30 時点のシグナル 検出割合 (総組合せ比, %)	
		第一被疑薬	全ての被疑薬	第一被疑薬	全ての被疑薬
PRR (MHRA)	PRR, χ^2 , n (報告数)	3,240	10,931	9.2	6.7
ROR (Lareb)	ROR	21,384	80,371	60.9	49.5
GPS (FDA)	EBGM (by GPS)	2,239	3,177	6.4	2.0
BCPNN (WHO-UMC)	IC	2,889	9,446	8.2	5.8

データマイニング手法の検討を行うための支援業務報告書 (平成 18 年 3 月) から転載

表 4 シグナル検出の各手法の感度と特異度

	検出数	感度	誤検出数	特異度
PRR	149	35.5%	553	97.5%
ROR	205	48.8%	1040	95.3%
GPS	105	25.0%	265	98.8%
BCPNN	146	34.8%	590	97.7%

データマイニング手法の検討を行うための支援業務報告書 (平成 19 年 3 月) から転載

であった。これは類似の手法である PRR の 9.2%と大きく異なるものであるが、3 件以上の報告数 ($n_{II} \geq 3$) という条件を ROR で用いなかったことに起因するものである。60%を拾い上げてしまうのでは、シグナルを絞り込むためのスクリーニングの基準としては適当とはいえない。一方、シグナル検出時期についての検討では、ROR が他の手法よりも報告数が少ない段階で早期にシグナルを検出すると報告されている。

2006 年度の報告書では暫定的な基準を設けて、感度と特異度についての検討結果が報告されている (表 4)。誤検出数 (false positive signals) がかなりの数ではあるが、特異度はいずれの手法も 95%を超えている。しかしながら、感度は最も高い ROR でも 48.8%に過ぎず、真の副作用 (true positive signals) を見逃す可能性が示されている。この結果は、松下らのシミュレーション研究¹⁰⁾とよく類似している。効率よく true positive signals を同定し、シグナル評価の対象となる false positive signals を如何に少なくするかは大きな課題である。このように感度・特異度やシグナル

検出量などの検討結果から、データマイニング手法への過度の期待は避けなければならないことが分かる。

さらに、副作用シグナルのデータマイニングにかかわる多様な選択の問題がある¹¹⁾ (表 5)。データマイニング手法のアルゴリズムとしてどの手法を使用するのか、その閾値をどのようにするか、といった手法にかかわるもののみではなく、有害事象の用語集とその使用法、検討する医薬品の取扱い、部分集団での検討、他剤や疾患との相互作用の検討、経時的な検討などについての多様な選択の余地があり、実際の適用の仕方を容易には決め難い。PMDA では、手法や閾値についての検討に加えて、副作用のグルーピング、薬剤の組合せ、性別や年齢階級別のサブグループ、部分データベース (特定の薬剤群や有害事象群) などについての検討を行い、異なるシグナルが検出され得ることを確認している。

一般に安全対策業務へのデータマイニング手法の導入については、次のような点が期待されるであろう。

表5 副作用シグナルのデータマイニングにかかわる多様な選択

<ul style="list-style-type: none"> ◆アルゴリズム <ul style="list-style-type: none"> ・PRR ・ROR ・BCPNN ・MGPS ◆計量的な基準 <ul style="list-style-type: none"> ・閾値 ・有害事象数の閾値 ◆有害事象の辞書 <ul style="list-style-type: none"> ・WHOART (JART), MedDRA ・用語のレベル: 基本語 (PT), 高位語 (HLT), 下層語 (LLT) ・利用者による PT の組合せ ・MedDRA 標準検索式 (SMQ) 	<ul style="list-style-type: none"> ◆医薬品 <ul style="list-style-type: none"> ・第一被疑薬 ・全ての被疑薬 ・特定の薬剤群 (基質, 薬効群...) ◆非層別, 層別 (サブグループ) <ul style="list-style-type: none"> ・年齢 ・性別 ・報告年 ・国内, 海外 ◆次元 <ul style="list-style-type: none"> ・2D (薬剤と有害事象のペア) ・3D (相互作用) ・3D (薬剤・疾患と有害事象) ◆時間的広がり <ul style="list-style-type: none"> ・横断的 (個々の時点ごと) ・縦時的動向分析
--	---

- ① 伝統的方法では検出できない有害事象の検出
- ② 伝統的方法でも同定される有害事象であるが, 早い時期に検出可能なこと
- ③ 伝統的方法で同定され, かつ同様な時期であるが, 少ない人的資源で検出できること
- ④ 伝統的方法に対するセーフティネット

PMDA での安全対策業務へのデータマイニング手法の導入は, おもに各報告の専門的レビューという定型業務において, データマイニング手法によるシグナル検出結果を併用する方向で検討されている (図4)。薬剤・有害事象の未知の組合せについては, まず伝統的方法による個別評価を行い, その後にデータマイニング手法によるスクリーニングを行うという併用が予定されている。ROR を中心的な指標とし, BCPNN と GPS を参考としての使用が考えられている。伝統的な方法では検出できないシグナルをも拾い上げようという, 伝統的方法に対するセーフティネットの役割でのデータマイニング手法の導入が考えられている。

シグナルの優先順位付けと評価

伝統的方法およびデータマイニング手法によっ

て検出されるシグナル数はかなりの量にのぼる。詳細なシグナル評価を行う前に検出されたシグナルに優先順位を付けることが実際的である。シグナルの優先順位付けにおいて,

- ・新規, 未知, 未報告
- ・新規の薬剤
- ・医学的重要性 (重篤, 非可逆など)
- ・薬剤の関与の可能性の高い特定医学的事象
- ・リチャレンジされた症例報告
- ・急激な非比例性に増大
- ・特定の集団での増加 (小児, 妊婦, 高齢者など)
- ・報告のクラスタリング
- ・社会的影響の大きさ (広範な使用, 多数の報告件数, 適用外使用など)

などの諸点が勘案されよう。

しかし, こうした検討によって有害事象の薬剤との因果性が確認されることにはならない。受動的サーベイランスの他にも大規模なデータベースが存在する欧米諸国では, 受動的サーベイランスとは異なる情報源で確認することがリスク評価の基本と考えられている。

FDA 再生法 (FDA Amendment Act) に基づ

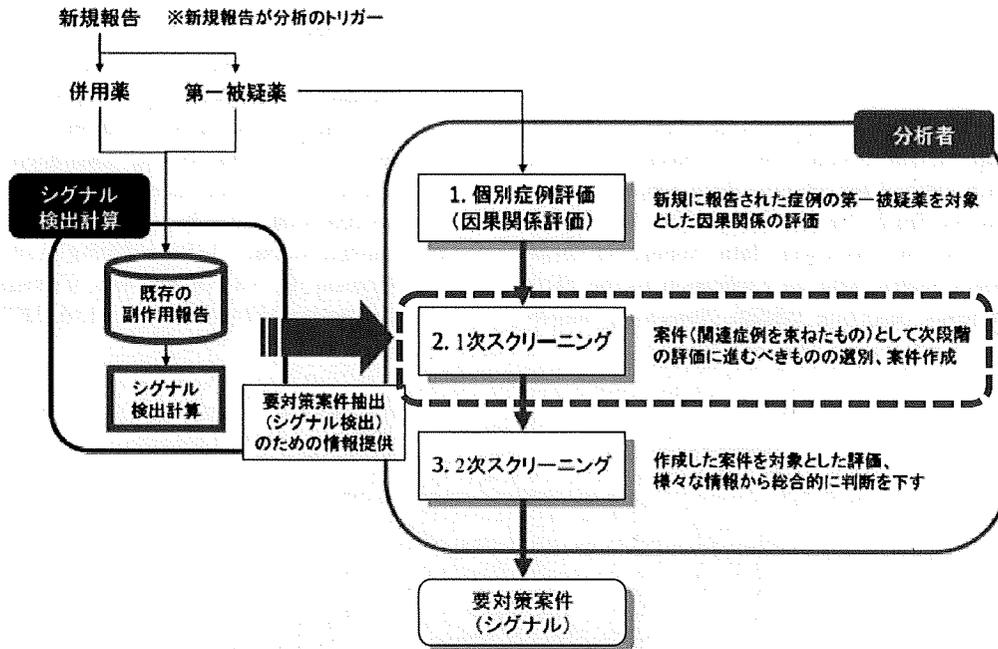


図 4 未知症例に関する新業務フロー
データマイニング手法の検討を行うための支援業務報告書（平成 20 年 3 月）から転載

いて、米国 FDA では潜在的シグナルを Web サイト上 (http://www.fda.gov/cder/aers/potential_signals/default.htm) で公表している。シグナル評価により薬剤と有害作用の因果性が確認された段階ではないが、シグナル優先順位化でかなり絞り込まれたものと考えられる。

リスク評価で有害事象の薬剤との因果性が確認された場合には、既に図 1 に示したとおり、リスクマネジメントのサイクルに入ることになる。

おわりに

日本においてもデータマイニング手法を併用したシグナル検出が 2009 年度から安全業務に導入される。実際の運用の中で gold standard となる評価が蓄積され、データマイニング手法の適切・有用な使用が今後本格的に模索していくものと考えられる。また、それに続くシグナル評価やリスクマネジメントに必須であるレセプトデータなどに基づく大規模データベースの実現に向けての努力が求められている現状である。

文 献

- 1) Edwards IR, Lindquist M, Wiholm B-E, Napke E. Quality criteria for early signals of possible adverse drug reactions. *Lancet* 1990 ; 336 : 156-8.
- 2) Clarke A, Deeks JJ, Shakir SA. An assessment of the publicly disseminated evidence of safety used in decisions to withdraw medicinal products from the UK and US markets. *Drug Saf* 2006 ; 29 (2) : 175-81.
- 3) Olivier P, Montastruc JL. The nature of scientific evidence leading to drug withdrawals for pharmacovigilance reasons in France. *Pharmacoepidemiol Drug Saf* 2006 ; 15 (11) : 808-12.
- 4) Wysowski DK, Swartz L. Adverse drug event surveillance and drug withdrawals in the United States, 1969-2002 : The importance of reporting suspected adverse reactions. *Arch Int Med* 2005 ; 165 : 1363-9.
- 5) Evans SW, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol Drug Saf* 2001 ; 10 : 483-6.
- 6) van Puijenbroek EP, Bate A, Leufkens HG, Lindquist M, Orre R, Egberts AC. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. *Pharmacoepidemiol Drug Saf* 2002 ; 11 : 3-10.
- 7) Rothman KJ, Lanes S, Sacks ST. The reporting odds

- ratio and its advantages over the proportional reporting ration. *Pharmacoepidemiol Drug Saf* 2004 ; 13 : 519-23.
- 8) Bate A, Lindquist M, Edwards IR, Olsson S, Orre R, Lansner A, De Freitas RM. A Bayesian neural network method for adverse drug reaction signal generation. *Eur J Clin Pharmacol* 1998 ; 54 : 315-21.
- 9) DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *American Statistician* 1999 ; 53 : 177-90.
- 10) Matsushita Y, Kuroda Y, Niwa S, Sonehara S, Hamada C, Yoshimura I. Criteria revision and performance comparison of three methods of signal detection applied to the spontaneous reporting database of a pharmaceutical manufacturer. *Drug Saf* 2007 ; 30 : 715-26.
- 11) Hauben M, Bate A. Data mining in drug safety. In Aronson JK (ed). *Side Effects of Drugs Annual 29*. Amsterdam : Elsevier Science Ltd. 2007 ; xxxiii-xlvi.

薬の安全性と医療データベース

久保田 潔 | Kubota Kiyoshi

東京大学大学院医学系研究科
薬剤疫学講座特任教授



■1978年北海道大学医学部卒。卒業研修を経て83年国立医療センター（現国立国際医療センター）臨床研究部臨床薬理学研究室研究員。89年カリフォルニア大学サンフランシスコ校訪問研究員。91年英国サザンプトン医薬品安全性研究ユニット（DSRU）卒業訪問研究員。96年東京大学大学院医学系研究科薬剤疫学講座助教授、2008年より現職。同年より日本薬剤疫学会理事長。特定非営利活動法人日本医薬品安全性研究ユニット（DSRU Japan）理事長。

1. 薬剤疫学と薬の安全性

薬剤疫学（pharmacoepidemiology）は、近年、薬の安全性の確保に重要な役割を果たすようになった学問領域の一つである。薬剤疫学は、「薬の」を意味する“pharmaco”と「疫学」を意味する“epidemiology”を結びつけた1980年代の造語で、「ヒトの集団における薬物の使用とその効果や影響を研究する学問」と定義されている。

薬剤疫学の起源の一つは、1950年代後半に発生したサリドマイド事件などの薬害事件である。第二次世界大戦後、抗生物質、副腎皮質ステロイドホルモン、降圧薬、経口糖尿病治療薬などが次々に登場し、薬物治療への夢と期待が広がったが、サリドマイド禍に象徴される薬害事件は「薬は有用性と同時に有害作用をもち、その有害作用は時に重篤である」という古くから知られている事実の再認識を人々に迫った。

まれで重篤な薬の副作用は、限られた人数の被験者によって実施される市販前の臨床試験では発見できないことが多く、その多くは市販後に明らかになる。未知の副作用の発見を主たる目的として、副作用症例を医師などが直接、または企業を通じて行政に報告する「自発報告制度」が1960年代後半から世界各国でスタートした。報告される症例は実際に発生した症例のごく一部であること、マスコミ報道などの様々な要因で報告率が大きく変動することも知られているが、重大な副作用が社会問題化した時点においては、自発報告が利用可能な唯一の情報源であることも多い。

2001年8月に高脂血症治療薬の一つ、スタチン系薬剤に分類されるセリバスタチン（商品名：バイコールなど）が市場から撤退した。撤退の理由は、他のスタチン系薬剤に比べて横紋筋融解症の発生率が高かったことである。横紋筋融解症は、横紋筋（骨格筋）が大量に

壊死を起こし、その結果、急性腎不全や死亡を引き起こす重篤な副作用であり、スタチン系薬剤でまれに発生することが知られている。セリバスタチンの撤退の決定は自発報告数と薬の売上から推測された十分正確とはいえない発生割合の比較を根拠にしており、撤退以後も横紋筋融解症の発生率をスタチン系薬剤間で比較する薬剤疫学研究がいくつか実施された。

薬剤疫学では、「人-年」(person-years)あたりの発生件数で発生率を示すことが多い。「人-年」は観察された期間の総和である。たとえば、50人が2年間ずつ観察されれば、100人-年であるが、25人が4年間ずつ、10人が10年ずつ観察されても観察期間の総和は100人-年と計算される。2004年に発表されたFDA(米国食品医薬品局)のGrahamらのスタチン系薬剤の横紋筋融解症に関する薬剤疫学研究^[1]では、セリバスタチン使用に伴う横紋筋融解症の発生率は10,000人-年あたり5.3と推定された。アトルバスタチン(商品名:リピトールなど)、シンバスタチン(商品名:リポバスなど)は10,000人-年あたり0.5なので、1桁高い発生率である。またフィブラート系薬剤と併用されると横紋筋融解症のリスクは100倍前後高まることも明らかになり、自発報告に基づく撤退の決定が正しかったことが示された。

2. 薬剤疫学研究とデータベース

前記セリバスタチンの横紋筋融解症の発生が10,000人-年あたり5.3件であるとのGrahamらの研究^[1]は、米国の11の“claims database”(医療費償還のための請求書をデータベース化

したもの)を用いて行われた。この11のデータベースから、過去にさかのぼって1998年1月から2001年6月までに新規に高脂血症治療薬(スタチン系またはフィブラート系薬剤)の投与が開始されたと判断された患者252,460人が特定された。「新規」とは、「使用開始日に先立つ180日間、その薬の処方がないこと」と定義されている。観察された高脂血症治療薬の新規使用者252,460人の観察期間の総和は232,940人-年であり、いずれかの高脂血症治療薬の単独使用期間225,640人-年と複数の高脂血症治療薬を併用した期間7,300人-年が含まれていた。このうち、セリバスタチン単独使用は12,695人について合計7,486人-年観察され、この観察期間内に4例の横紋筋融解症が確認され、発生率($4 \div 7486 = 0.00053$ /人-年)は、アトルバスタチン129,367人-年に7件、シンバスタチン40,940人-年に2件から得られる値(いずれも約0.00005/人-年)の10倍であった。

わが国では長く、新薬を使用した患者数千人を医師が登録し、登録された患者を数ヶ月追跡した結果を調査票に記入する「使用成績調査」(製薬会社が法の規定に基づき、自社製品使用に伴い、日常診療下での副作用の発生割合などを調べるために複数の医療機関に依頼して行う調査)が実施されてきた。実施には数年を要し、直接費用だけでも数億円を要する。観察総期間は数千人-年程度の調査がほとんどで、10,000人-年を超える調査はわずかである。Grahamらのセリバスタチンの研究に要した費用は不明であるが、日本円で数千万円を超えることはないと推測される。この推測が正しければ、上記データベース研究に

要した1人-年あたりの費用は日本の「使用成績調査」の100分の1以下であったと考えられる。また、「使用成績調査」では、通常、自社製品を使用した患者だけを調査対象とするが、Grahamら^[1]の研究では、複数のスタチン剤が一つの研究で比較されており、データベース研究で得られる情報の価値は高い。

3. 個人IDのデータベースへの保持と個人情報保護

わが国でも、平成23年度までにレセプト（診療報酬明細書）の完全オンライン化と電子化されたレセプト情報の有効活用を目指すナショナルデータベース構築が進んでいる。筆者が所属する日本薬剤疫学会は、このナショナルデータベースが薬剤疫学研究にとって有用な資源であるとの認識に基づき、平成19～20年にかけて要望書を行政に提出した（表1）。以前からレセプトのデータベースとしての利用を望む声はあったが、ここに来て、にわかには現実性を帯び、平成21年度において、厚生労働科学研究班及び行政での利用にかかわる本格的検討が開始されようとしている。

レセプトの利用に関しては、「『レセプト診断』は不正確であり、レセプトデータの有用性は限られる」との意見を耳にすることがある。日本に限らず、医療費償還のための診断名が不正確であることは周知の事実であるが、正確な情報にたどりつくための貴重な手掛かりにはなりうる。Grahamらの研究では、対象者を入院患者に絞り、まず、退院時の主たる診断名（primary diagnosis）が、「ミオグロビン尿」「筋炎」「筋の脱力」などの患者と、副次的診断名（secondary diagnosis）がこれら

の病名で、かつ入院7日前から退院後7日までの間に、筋障害で値が上昇する血清creatin kinase（CK）値の検査が実施されている患者が「候補」としてデータベースから選ばれた。それら「候補」の患者について病院のカルテに戻り、カルテから抽出された詳細な情報を専門家がレビューした。結果、上記「候補」に該当したのは194例で、その90%の174例からカルテ情報が取得され、最終的に31例が横紋筋融解症と診断された。つまり、レセプト診断を使って広めに投げた網にかかった「候補」から、真のケースを絞り込む手法がとられた。この方法を可能とする条件として、医療機関の協力姿勢、研究者の経験と熟練とともに、データベースが個人の識別を可能にする情報を保っていたことを挙げることができる。

個人の識別を可能とする情報をデータベースに残すことは、個人情報の保護の観点から問題であるという声もある。しかし、個人情報の保護と個人のID情報をデータベースに残すことは技術的に両立可能であり、薬剤疫学研究に利用されている欧米のデータベースでは、例外なく実現されている。上記レセプト病名から横紋筋融解症の症例を絞り込む過程にも見られるように、個人ID情報の保持がデータベースの価値を高める。さらに、日本薬剤疫学会からの要望書（表1）にもある通り、フィブリノゲン投与によるC型肝炎ウイルス感染における有効な対策にもつながりえたと考えられる。

また、個人IDが保持されていれば、データベースの外部にある他のデータ源と連結（リンク）させることも可能となる。たとえば、

表1 日本薬剤疫学会からの要望書の要約

日本薬剤疫学会からの要望書	
<p>I. 厚生労働大臣への医薬品等の安全確保のためのレセプト情報活用に関する要望書</p> <p style="text-align: right;">平成19年12月17日</p> <p>前文 (要約)</p> <p>高度情報通信ネットワーク社会推進戦略本部が平成18年1月19日に公表した「IT新改革戦略」に基づき、2011年度までにレセプト完全オンライン化と電子化されたレセプト情報の有効活用を目指すナショナルデータベース構築が進んでいる。医薬品の安全確保への利用はその有効活用の一環として重要。たとえばフィブリノゲン投与によるC型肝炎ウイルス感染では、レセプト情報が適切にデータベース化されていれば当該患者を迅速に特定でき、必要な対応が取りえたはず。また、レセプト情報を死亡情報やがん登録情報と連結すれば、死亡やがんと医薬品との関係について迅速かつ効率的に調査することが可能。</p> <p>多数の国で固有の医療システム下でデータベースが構築され、医薬品副作用に関する薬剤疫学研究が高い精度で迅速に実施され、市販後医薬品の安全確保が推進されているが、我国には大規模データベースが存在せず、世界から遅れを取っている。</p> <p style="text-align: center;">要望事項 記</p> <ol style="list-style-type: none"> 1. 医薬品等の安全確保を含むレセプト情報の包括的活用を図るために、全省的体制の下でレセプトデータベース構築を推進すること。 2. 重大な副作用への緊急対応時のため、レセプトデータベースに基づく当該患者の特定が図れる要件を盛り込むこと。 3. 医薬品の安全確保に関する公益性の高い疫学研究のために、構築されたレセプトデータベースを第三者が利用できるようにすること。 	<p>II. 内閣官房IT対策室への「重点計画-2008 (案) に関する意見」</p> <p style="text-align: right;">平成20年7月8日</p> <p>(要約) 日本薬剤疫学会は、「重点計画-2008 (案)」の「レセプトデータ等の利活用」(5~6ページ)に、特に重大な関心を持っている。質の高い大規模データベースを構築し、薬剤疫学的活用を図ることが、医療、特に医薬品の安心安全の確保につながる。日本には大規模データベースが欠如しており、医薬品の安全対策の大きな障害となっている。医薬品の安全確保のためにレセプトデータ等の活用を図ることは、薬害を繰り返さない恒久的な対策の一環として不可欠。「重点計画-2008」に「医薬品等の安全確保のために、レセプトデータ等の利活用を図る。」を明記すべき。</p>

注) 全文はホームページ (<http://www.jspe.jp/>) 参照。

スタチン系薬剤が発がんを予防するとの仮説がいくつかの研究から提示され、スタチンとがん発生に関する薬剤疫学研究がこれまでに複数実施されている[2][3][4]。これらの研究の多くで、医療費償還のための記録とがん登録または退院時診断のデータをリンクする「レコードリンケージ」の手法が使われている(表2)。このレコードリンケージの

鍵も、個人を識別可能な個人ID情報が、それぞれのデータベースに保存されていることである。

4. 台湾と韓国のデータベース

アジアには、日本と同様の国民皆保険制度をもつ国が存在し、そのうち、台湾と韓国では、医療費償還のためのデータのデータベー

表2 レコードリンケージによるスタチンとがん発生との関係に関する研究の例

研究者 発表年	地域	研究デザイン	アウトカム	薬の処方情報の取得に利用したデータベース	アウトカムの特定に利用したデータベース
Setoguchiら 2007[2]	アメリカのペンシルベニア州	コホート研究	肺、乳、大腸がん	Pennsylvania州のmedicareデータベース	Pennsylvania state cancer registry data
Karpら 2008[3]	カナダのケベック州	コホート研究	すべてのがん	Quebec州のdrug claims database	The Quebec hospital discharge summary database
Murtolaら 2007[4]	フィンランド	症例対照研究	前立腺がん	Nationwide prescription database of the social insurance institution of Finland	Finnish cancer registry

ス化が実現されている。台湾では、1995年に2200万人の全国民をカバーするデータベースが実現し、疾患疫学、医薬品・医療機器の使用実態、医療管理、医療経済、医療政策への利用が開始されている。韓国では、1999年のNational Health Insurance Act (NHIA) による保険者の一元化に引き続いて2000年7月にはHealth Insurance Review Agency (HIRA) が設立された。全国民4900万人をカバーするHIRAデータベースは、医薬品使用実態のレビューや新たな副作用の検出を含む薬剤疫学研究など、様々な調査研究に利用されている。

台湾と韓国のデータベースでも、個人ID情報がデータに含まれ、カルテ情報に戻って行う情報の検証 (validation) や外部データとのリンクの可能性が保たれている。

5. わが国のレセプトデータベース設計の問題

わが国のレセプトデータベースにおいては、倫理的配慮からハッシュ関数により個人識別情報が置き換えられる。平成20年2月に発表された「医療サービスの質の向上等のためのレセプト情報等の活用に関する検討会にお

ける報告書」では、特定健診のデータとのリンクは考えられているが、「医療費適正計画の作成等に必要な分析上、特定の患者等…を識別する必要はないことから、患者等については特定の個人が識別できないよう、国がデータを収集する際には、患者等の氏名など個人情報を削除する必要がある」[5]とし、個人情報を削除することが明記されている。ハッシュ関数は、氏名、生年月日、被保険者番号の組み合わせからユニークなコード (ハッシュ値) を生み出すが、ハッシュ値自体から逆に氏名や生年月日を知ること (復号化) は不可能である。同一の個人識別情報から生成されるハッシュ値は同一なので、データベース内でハッシュ値を照合し、新たなデータを追加することなどは可能だが、外部データとの連結は困難である。

レセプトデータは外部データとリンクさせなければ、十分価値のあるデータにはならない。現在、わが国で構築が進んでいるデータベースでは「医療費適正計画の作成」以外にデータベースが多様な資源となりうる事が無視されている。日本薬剤疫学会からの平成19年12月17日付けの要望書 (表1) では、

個人を識別する情報をデータベースに含めるべきであるとの認識から「重大な副作用への緊急対応時のため、レセプトデータベースに基づく当該患者の特定が図れる要件を盛り込むこと」を要望事項の二点目に含めた。また、担当の局または課での用途以外が無視される縦割り行政が、「医療費適正計画の作成」以外のデータベースの利用可能性の軽視につながっているとの認識から、「全省的体制の下でレセプトデータベース構築を推進すること」を要望事項の一点目に据えた。さらに、レセプトデータベースが、もっぱら官によって利用されるべきものと位置づけられる可能性への懸念が、「公益性の高い疫学研究のために、構築されたレセプトデータベースを第三者が利用できるようにすること」という三点目の要望につながっている。

6. その他推進すべき事柄

スタチンとがん発生との関係に関する研究(表2)では、がん登録や退院時サマリーのデータベースなどが薬使用のデータとリンクされた。日本におけるがん登録は、都道府県単位で実施されてきたが、最近、乳がんなど一部のがんについては、全国レベルの包括的な登録体制確立へ向けた学会主導の動きが見られるようになった。また、わが国でも地域の医療データ、複数の病院のデータのデータベース化と研究への利用に向けた取組みも始まっている。さらに、旧来の画一的な「使用成績調査」とは異なった多様な薬剤疫学研究の手法を用いた市販後医薬品の調査のあり方を定着させることも重要である。データベースの利用が常に最適というわけではない。日本

薬剤疫学会では、現在、日本病院薬剤師会学術委員会と共同で、比較群をもち効率性の高い「ケース・コホート研究」のデザインを使った複数のスタチンの副作用発生を比較する「JSS研究」を試行している(<http://www.jspe.jp/>の「What'sNew 2009/03/02」参照)。データベースをもつ欧米諸国でも、当該調査・研究のために収集したデータ(一次データ)を用いる研究は、依然として有用な選択肢の一つとして認識されている。今後、国家レベルの戦略を立て、全国的な疾患登録や医療データベース構築の動きを加速させ、また課題に応じて最適な調査方法が選択される市販後調査の体制を整備すべきである。

*参考文献

- [1] Graham DJ et al. (2004) : Incidence of hospitalized rhabdomyolysis in patients treated with lipid-lowering drugs: JAMA, 292, pp.2585-2590.
- [2] Setoguchi S et al. (2007) : Statins and the risk of lung, breast, and colorectal cancer in the elderly: Circulation, 115, pp.27-33.
- [3] Karp I et al. (2008) : Statins and cancer risk: Amer J Med, 121, pp.302-309.
- [4] Murtola TJ et al. (2007) : Cholesterol-lowering drugs and prostate cancer risk: a population-based case-control study: Cancer Epidemiol Biomarkers Prev, 16, pp.2226-2232.
- [5] 「医療サービスの質の向上等のためのレセプト情報等の活用に関する検討会」報告書(<http://www.mhlw.go.jp/shingi/2008/01/s0130-16.html> 2009. 5. 10アクセス).

副作用自発報告データベースからのシグナル検出

藤田 利治 | Fujita Toshiharu

情報・システム研究機構
統計数理研究所データ科学研究系教授



■1976年東京大学医学部保健学科卒。83年東京大学大学院医学系研究科保健学専攻博士課程修了。保健学博士。国立公衆衛生院疫学部、国立保健医療科学院疫学部を経て、2006年4月より現職。08年より日本薬剤疫学会副理事長。専門は疫学、薬剤疫学。

1. 受動的サーベイランスに基づくシグナル検出

医療用の薬剤は、市販前の安全性と有効性を確認する臨床試験や動物実験などによる厳密な開発プロセスと厳格な承認審査プロセスを経て社会に導入される。しかし、薬剤に未知のリスクがないことを完全に保証することはできない。すべての薬剤はリスクを伴うものであり、新しい重篤な副作用はしばしば市販後に明らかになる。

副作用の受動的サーベイランス（自発報告制度）は、未知の副作用を検出するシステムとして、市販後の薬剤の安全確保に大きな役割を果たしてきた。日本における自発報告制度は、医薬品等安全性情報報告制度と企業報告制度からなっている。従来、自発報告の評価は、医学・薬学の専門の知識を持った評価者による専門的な評価という伝統的方法で実施されてきた。しかし、膨大な数の自発報告

について個々に専門的評価を行うことは極めて大きな労力を要し、また重大な副作用の見逃しの危険の増加も懸念されるようになった。こうした状況において、自発報告の大規模データベースを活用して「副作用のシグナル」を検出するデータマイニング手法の検討が進められ、専門的評価を支援するシステムとして実用化されてきている。ここで、副作用のシグナルとは、「それまで知られなかったか、もしくは不完全にしか立証されていなかった薬剤と有害事象との因果関係の可能性に関する、さらに検討が必要な情報」である[1]。

自発報告制度の短所として、過少報告、不安定な報告割合、詳細な臨床情報の不足、発生頻度の分母となる直接的な情報の欠如のために、発生頻度の算出が困難、遅発性の有害作用の検出を見逃す可能性、背景での発生率が低いとはいえない疾患を見逃す可能性（例：高齢者での虚血性心疾患、悪性新生物な

ど)、などの問題がある。一方、長所としては、薬剤と患者の広い範囲をカバーしていること、実際の臨床での使用実態を反映していること、実施が安価で単純であること、そして、何よりも未知・重篤な副作用を検出するための最も効率的な手段であること、などを挙げることができる。安全性問題で撤退した薬剤の36～50%は、症例報告のみが根拠との報告がなされている[2][3][4]。未知で重篤な副作用の早期発見のために、自発報告制度からの大規模データベースを用いたデータマイニング手法によるシグナル検出の取組みが世界中で進行している。

2. シグナル検出の方法

自発報告データベースを用いたシグナル検出は、「伝統的方法」と「データマイニング手法による方法」に分けることができる。伝統的方法では、リチャレンジによる確認、特定医学事象（アナフィラキシー、Stevens-Johnson症候群、横紋筋融解症、無顆粒球症など）、標的医学事象（特定の薬剤や患者集団に関係する事象）、薬剤曝露の直後の発現、ゼロに近い背景での3例以上の症例報告、などが暗黙のうちに重視されてきた。

一方、統計的なデータマイニング手法では、非比例的な報告パターンについての自動的・網羅的スクリーニングがなされる。データマイニング手法によるシグナル検出の元となるデータは、行に薬剤を、列に有害事象を取り、その報告件数を度数とする表1のような度数表である。表1が通常の統計分析で扱われる度数表と異なる点は、薬剤の種類および有害事象の種類が極めて多く、表のセル数が膨大

であることである。例えば、日本で市販されている12,000超の薬剤と17,000超の有害事象では2億を超える組合せ数（セル数）であり、個々のセルの度数は小さく、1件や2件といったセルがかなりの割合を占めることになる。表1における特定の薬剤と特定の有害事象に注目すると、表2のような2×2分割表ができる。こうしたデータを用いるデータマイニング手法には、いくつかの仮定がある。

- ①ある薬剤が特定の有害反応を起こすならば、他の薬剤と比べて、この有害反応が多く報告される
- ②同じ有害反応について、（過少）報告の程度は異なる薬剤間で同じである
- ③種々の有害反応の報告割合や報告パターンは、異なる薬剤間で類似している

しかし、これらの仮定は弱いものであり、多くの反する状況がある。例えば、市販後間もない期間には報告が多いが、次第に少なくなるというWeberカーブ、メディアの注目、緊急安全性情報などにより、仮定が成り立た

表1 シグナル検出の元となるデータ

	有害事象 1	有害事象 2	...	有害事象 p	合計
薬剤 1	n_{11}	n_{12}	...	n_{1p}	n_{1+}
薬剤 2	n_{21}	n_{22}	...	n_{2p}	n_{2+}
⋮	⋮	⋮	...	⋮	⋮
薬剤 m	n_{m1}	n_{m2}	...	n_{mp}	n_{m+}
合計	n_{+1}	n_{+2}	...	n_{+p}	n_{++}

表2 2×2分割表のセル度数

	特定の有害事象	その他の有害事象	合計
特定の薬剤	n_{11}	n_{12}	n_{1+}
その他の薬剤	n_{21}	n_{22}	n_{2+}
合計	n_{+1}	n_{+2}	n_{++}

ない状況がしばしば起きる。このため、シグナル検出のデータマイニングには、ロバストな手法が求められる。

シグナル検出のデータマイニング手法として代表的なアルゴリズムであるProportional Reporting Ratio (PRR) [5][6]の指標は、報告割合の比である。

$$PRR = \frac{n_{11}/n_{1+}}{n_{21}/n_{2+}} = \frac{P_{11}}{P_{21}}$$

通常、以下の3つの条件を満たす場合(閾値)に、「シグナルあり」と判断される。

① $PRR \geq 2$

② $\chi^2 = \frac{n_{++}(|n_{11}n_{22} - n_{12}n_{21}| - n_{++}/2)^2}{n_{1+}n_{2+}n_{+1}n_{+2}} \geq 4$

③ $n_{11} \geq 3$

また、オッズ比であるReporting Odds Ratio (ROR) は、理論的にはPRRよりもバイアスの少ない方法であることが指摘されている[7]。

$$ROR = \frac{(n_{11}/n_{21})}{(n_{12}/n_{22})} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

ベイズ流のアプローチを用いる手法として、世界保健機関 Uppsala Monitoring Centre (WHO-UMC) で用いられているBayesian Confidence Propagation Neural Network (BCPNN) Method[8]と、米国FDA(食品医薬品局)で使用されていたGamma-Poisson Shrinker (GPS) Program[9]があり、現在は薬物相互作用を考慮したMGPS (Multi-Item Gamma Poisson Shrinker) が使用されている。これらの4つの手法での指標は、いずれも表2の n_{11} の非比例的な報告の増大を検出しようとするものである。その観察値 O に

対して、わずかに異なる期待値 E を用いていることに相当し、次の関係にある。

• Proportional Reporting Ratio (PRR) $\frac{n_{11} \times n_{2+}}{n_{21} \times n_{1+}}$

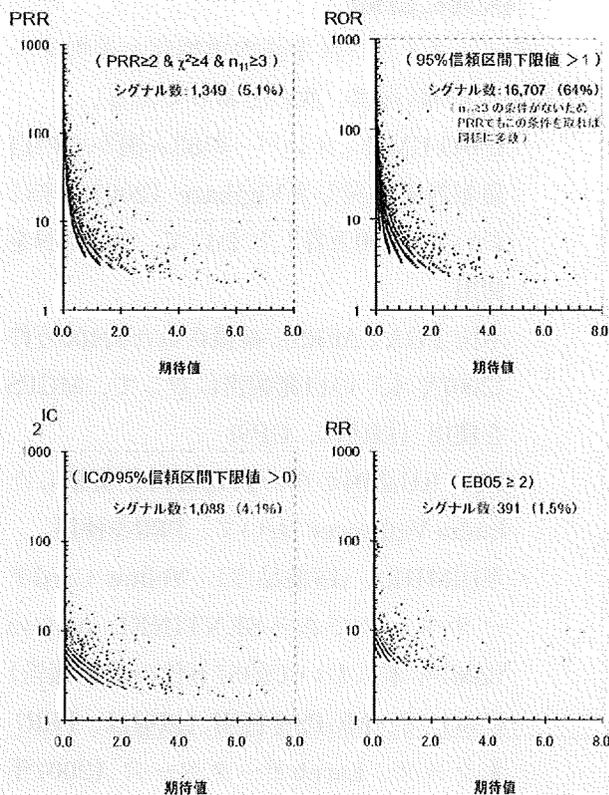
• Reporting Odds Ratio (ROR) $\frac{n_{11} \times n_{22}}{n_{21} \times n_{12}}$

• Information Component (IC) $\log_2 \frac{n_{11} \times n_{++}}{n_{1+} \times n_{+1}}$ BCPNNは、これについてのベイズ流のアプローチ

• Relative Reporting (RR) $\frac{n_{11} \times n_{++}}{n_{1+} \times n_{+1}}$ GPSは、これについてのベイズ流のアプローチ

ベイズ流アプローチでは、期待値が小さい場合、 O/E は「1」の方向に縮小する。図1は、米国のAERS (Adverse Event Reporting System) の2001年のデータを用いて26,359の薬剤・事象の組合せに対して各データマイニ

図1 検出されたシグナルの n_{11} の期待値と指標値



注) 指標値1,000以下、期待値8以下について表示。2001年-AERSの15日報告データを使用。26,359の組合せ種類。