

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場合にシグナルの指標値が極めて大きくなることもあり、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は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。

3. データマイニング手法の実用化の状況と課題

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：メンバーの83ヵ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（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（医薬品庁）：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

日本においては、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、自発報告データベースの整備が進められている。このデータベースを用いて、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したが、その導入に先立ってさまざまな検討を行い、報告書をWeb上で公表している。

例えば、PMDAのデータベース（2003年10月27日～2005年9月30日までの期間の約5万件の報告）を使用した第一被疑薬についてのシグナル検出数は、RORでは21,384も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%に当たる量であった。これは類似の手法であるPRRの9.2%（3,240）と大きく異なるものであるが、3件以上の報告数（ $n_{11} \geq 3$ ）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。また、BCPNNとGPSでは、それぞれ8.2%（2,889）と6.4%（2,239）であった。

また、暫定的な基準を設けて、PRR、ROR、BCPNNおよびGPSの感度と特異度についての検討結果が報告されている。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、PRRが35.5%、BCPNNが34.8%、GPSが25.0%であり、真の副作用（true positive signals）を見逃す可能性が示されている。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となっている。

さらに、副作用シグナルのデータマイニン

表3 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

◆アルゴリズム	◆医薬品
<ul style="list-style-type: none"> PRR ROR BCPNN MGPS 	<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…)
◆計量的な基準	◆非層別、層別(サブグループ)
<ul style="list-style-type: none"> 閾値 有害事象数の閾値 	<ul style="list-style-type: none"> 年齢 性別 報告年 国内、海外
◆有害事象の辞書	◆次元
<ul style="list-style-type: none"> WHOART(IART), MedDRA 用語のレベル: 基本語(PT), 高位語(HLT), 下位語(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"> 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

グにかかわる多様な選択の問題がある[10] (表3)。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が可能であり、実際の適用の仕方を容易には決め難い。PMDAでも、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。

以上のように、データマイニング手法によるシグナル検出には、今後とも検討を要する多くの課題が山積している。

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。

- ① 伝統的方法では検出できない有害事象の検出が可能なこと
- 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、主に各報告の専門的評価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ものである。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手順で進められている。この場合、中心的な指標としてRORを用い、BCPNNとGPSを参考指標としている。このようにPMDAでは、伝統的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なされたといえる。今後、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・有用な使用方法が本格的に模索されるものと考えられる。

4. シグナルの評価と分母情報のある大規模データベース

シグナルは「薬剤と有害事象との因果関係の可能性に関する情報」であり、「さらに検討が必要な情報」である。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、実際は、かなりの量にのぼる。シグナルの評価に当たっては、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、自発報告のみによる検討によって、有害事象と薬剤との因果関係を科学的に確認することは、多くの場合、困難である。世界的には、自発報告データベースとは異なる情報源を用いて、シグナルを評価することが基本とされている。

欧米諸国や韓国、台湾には、薬剤の安全確保に利用できる数十万人～数千万人規模の集団についての大規模データベース（medical claims databases, electronic health record databases）が多数存在している。これらの大規模データベースには、有害事象を発生した個人を含む、薬剤を使用したすべての個人のデータが保管されており、有害事象の発生頻度を算出することができる。自発報告データベースから検出されたシグナルを、こうした分母情報のある大規模データベースにおいて、迅速に検証することが基本となっている。さらに、遅発性の有害作用や背景での発生頻度が低いとはいえない疾患に対する薬剤によるリスクの増大について、これ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も試みられている。

日本では、大規模な集団についてのデータベースを欠いた状態が続いており、そ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。折しも、政府の高度情報通信ネットワーク社会推進戦略本部が推進する「IT新改革戦略」（2006年1月19日）に基づき、2011

年度までにレセプト完全オンライン化ならびに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の構築が、現在進められている。国民が繰り返し被っている薬害、すなわち被害拡大を未然に予防し得る薬剤による被害を軽減するため、国民の財産であるレセプト情報をどのように役立てていくかの具体的検討が期待されている昨今である。

* 参考文献

- [1] Edwards IR, Lindquist M, Wiholm B-E, et al. (1990) : Quality criteria for early signals of possible adverse drug reactions : *Lancet*, vol.336, pp.156-158.
- [2] Clarke A, Deeks JJ, Shakir SA (2006) : An assessment of the publicly disseminated evidence of safety used in decisions to withdraw medicinal products from the UK and US markets : *Drug Saf*, vol.29(2), pp.175-181.
- [3] Olivier P, Montastruc JL (2006) : The nature of scientific evidence leading to drug withdrawals for pharmacovigilance reasons in France : *Pharmacoepidemiol Drug Saf*, vol.15(11), pp.808-812.
- [4] Wysowski DK, Swartz L (2005) : Adverse drug event surveillance and drug withdrawals in the United States, 1969-2002; The importance of reporting suspected adverse reactions : *Arch Int Med*, vol.165, pp.1363-1369.
- [5] Evans SW, Waller PC, Davis S (2001) : Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports : *Pharmacoepidemiol Drug Saf*, vol.10, pp.483-486.
- [6] van Puijnenbroek EP, Bate A, Leufkens HG, et al. (2002) : A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions : *Pharmacoepidemiol Drug Saf*, vol.11, pp.3-10.
- [7] Rothman KJ, Lanes S, Sacks ST (2004) : The reporting odds ratio and its advantages over the proportional reporting ratio : *Pharmacoepidemiol Drug Saf*, vol.13, pp.519-523.
- [8] Bate A, Lindquist M, Edwards IR, et al. (1998) : A Bayesian neural network method for adverse drug reaction signal generation : *Eur J Clin Pharmacol*, vol.54, pp.315-321.
- [9] DuMouchel W (1999) : Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system : *American Statistician*, vol.53, pp.177-190.
- [10] Hauben M, Bate A (2007) : Data mining in drug safety : In Aronson JK (ed) : *Side Effects of Drugs Annual 29*. Amsterdam; Elsevier Science Ltd, xxxiii-xlvi.

