

データマイニング手法の導入に関する 検討結果報告書

平成 20 年 3 月

独立行政法人 医薬品医療機器総合機構

目次

1. 調査概要.....	1
1.1 調査目的.....	1
1.2 調査内容.....	1
1.3 検討委員会の設置.....	3
2. 海外規制当局におけるデータマイニングシステムに関する調査.....	5
2.1 海外調査概要.....	5
2.2 Lareb.....	6
2.2.1 オランダの副作用報告制度の概要.....	6
2.2.2 Lareb の構成.....	8
2.2.3 Lareb における安全対策業務フロー.....	9
2.2.4 L@reb2002 システムの概要.....	12
2.2.5 ROR の利用状況.....	13
2.2.6 その他ヒアリング内容.....	15
2.3 EMEA.....	16
2.3.1 EU の副作用報告制度の概要.....	16
2.3.2 EMEA における安全対策業務フロー.....	17
2.3.3 EudraVigilance システムの概要.....	20
2.3.4 PRR の利用状況.....	21
2.4 まとめ.....	23
3. 新業務フローの検討.....	26
3.1 分析者アンケートの実施.....	26
3.1.1 アンケート概要.....	26
3.1.2 アンケート項目.....	26
3.1.3 アンケート結果.....	27
3.1.4 アンケートまとめ.....	30
3.2 新業務フローの検討.....	31
3.3 1次スクリーニングの試行実験.....	32
3.3.1 実験の目的.....	32
3.3.2 実験方法.....	32
3.3.3 実験結果.....	34
3.4 今後の課題.....	37

4. シグナル検出手法の高度化検討.....	38
4.1 非定型業務でのデータマイニング利用	38
4.2 Ad Hoc なシグナル指標値の利用	38
4.2.1 検討方法.....	39
4.2.2 検討結果.....	39
4.3 関連する取組.....	41
4.3.1 海外規制当局における関連する取組.....	41
4.3.2 関連研究.....	42
4.4 分析支援ツールのプロトタイプ	44
4.5 今後の課題	47
5. 調査のまとめ.....	49
参考文献	51

略語一覧

規制機関名称

略語	正式名称	日本語名
PMDA	Pharmaceuticals and Medical Devices Agency	独立行政法人 医薬品医療機器総合機構
Lareb	Netherland Pharmacovigilance Centre	オランダ薬剤監視センター
MEB	Medicines Evaluation Board	オランダ医薬品評価委員会
EMA	European Medicines Agency	欧州医薬品庁
WHO	World Health Organization	世界保健機関
UMC	Uppsala Monitoring Center	WHO 医薬品副作用情報収集センター
MHRA	Medicines and Healthcare products Regulatory Agency	英国医薬品庁
FDA	Food and Drug Administration	米国食品医薬品局

シグナル検出手法

略語	正式名称	日本語名
BCPNN	Bayesian Confidence Propagation Neural Network	—
PRR	Proportional Reporting Ratios	—
GPS	Gamma-Poisson Shrinker	—
MGPS	Multi-item Gamma Poisson Shrinker	—
ROR	Reporting Odds Ratio	—

MedDRA用語

略語	正式名称	日本語名
MedDRA	Medical Dictionary for Regulatory Activities	ICH 国際医薬用語集
PT	Preferred Terms	基本語
SOC	System Organ Class	器官別大分類

その他

略語	正式名称	日本語名
CAP	Centrally Authorized Products	EU 承認薬
ICSR	Individual Case Safety Report	個別症例安全性報告
MAH (DOS)	Marketing Authorisation Holders	医薬品市販承認取得者（当該医薬品の製造承認を世界で初めて取得した企業）
MPR	Medical Product Report	医薬品報告
NCA	National Competent Agency	EU 内の各国機関
PIL	Patient Information Leaflets	患者向け医薬品情報書
PSUR	Periodic Safety Report	定期的安全性最新報告
SmPC	Summary of Product Characteristics	欧州製品概要
SUSAR	Suspected Unexpected Serious ADR	重篤未知疑いのある有害事象
ETL	Extraction, Translation, Loading	—
EVDAS	EudraVigilance Data Analysis System	—
EVDBMS	EudraVigilance DataBase Management System	—
EVMPD	EudraVigilance Medical Product Dictionary	—
EVDWH	EudraVigilance Data Warehouse	—
EVPRM	EudraVigilance Product Message	—

1. 調査概要

1.1 調査目的

独立行政法人 医薬品医療機器総合機構(PMDA)の中期計画(平成16年度～平成20年度)では、安全対策業務に資するデータマイニング手法を導入すると定められている。この中期計画を実現するため、平成16年度に中期計画期間内に導入するデータマイニング手法の概要と中期計画期間内の実施スケジュールの策定を行った。また平成17年度は、業務に導入するデータマイニング手法を確定するために、主にデータマイニング手法導入後の業務フローの検討、および基本的シグナル検出手法の詳細検討を行った。さらに、平成18年度は基本的シグナル検出手法の精度の向上と高度化の検討を行い、導入するデータマイニング手法について確定した。中期計画期間中のスケジュールを図1-1に示す。

平成19年度は、データマイニング手法を導入した業務システムの開発を行うのに並行して、データマイニング手法を用いた安全対策業務の業務フローの検討、およびデータマイニング手法の高度化検討を行った。

1.2 調査内容

本年度、調査および検討を行った主な項目は次のとおりである。

(1) データマイニングを用いた業務システムの仕様作成

平成20年度末までにPMDAの安全対策業務で利用を開始するデータマイニングを用いた業務システムを導入するために、開発するデータマイニングシステムの仕様を作成した。仕様の作成に当たっては、具体的に以下の検討を行った。

① システム必要要件の検討

データマイニングシステムが擁するべき機能、およびデータマイニング実施に必要とされるデータ項目等、システムの必要要件をまとめた。

② システムインタフェースの検討

データマイニングシステムによる分析結果情報を効果的に分析者に提示するためのインタフェースについて検討を行い、必要要件をまとめた。

③ システム利用者へのヒアリング

システムの利用者に対してヒアリング調査を行い、①、②で検討した内容について改良が必要な点を抽出し、仕様に反映した。

(2) データマイニング導入後の新業務フローの検討

データマイニングを導入した後、開発する業務システムを用いて実施する安全対策業務フロー（新業務フロー）について検討した。

① 海外規制当局におけるデータマイニングシステムに関する調査

海外規制当局におけるデータマイニングを用いた業務フロー、業務フローの検討の経緯、およびデータマイニングを用いた業務システムのインタフェース等について調査を行い、参考とすべき点について整理した。

② 新業務フロー案の作成

平成 18 年度までの検討結果に基づきデータマイニングシステムを導入した際の、PMDA における安全対策業務の新業務フロー案を作成した。

(3) データマイニングの高度化に関する検討

データマイニングの高度化、および将来的に業務に導入するデータマイニング手法の検討として、非定型業務へのデータマイニングの活用に向けた検討とツールの開発を行った。

なお、(1)で作成した仕様については、本報告書とは別に仕様書案としてまとめている。

以上とは他に、データマイニング手法の導入に向けた検討として、PMDA において、新業務フローに必要となるデータ整備の一環としての未知・既知データベースの整備が、また新業務フローにおける業務量評価の試行が行われた。

1.3 検討委員会の設置

本調査検討の遂行にあたっては、以下の委員からなる「データマイニングに関する検討委員会」を設置し、調査結果の報告内容について検討を頂いた。

【データマイニングに関する検討委員会】¹

座長	：藤田 利治	統計数理研究所 データ科学研究系 教授
委員	：岩崎 学	成蹊大学 工学部 経営・情報工学科 教授
	岡田 孝	関西学院大学 理工学部 教授
	岡田 美保子	川崎医療福祉大学 医療情報学科 教授
	酒井 弘憲	日本製薬工業協会 医薬品評価委員会 統計・DM 部会 副部長
	櫻井 靖郎	財団法人 日本公定書協会 JMO 事業部 事業部長
	藤原 昭雄	日本製薬団体連合会 安全性委員会 副委員長
	望月 眞弓	共立薬科大学 医薬品情報学講座 教授
	鷲尾 隆	大阪大学 産業科学研究所 教授

本報告書は、平成 19 年度「データマイニング手法の検討を行うための支援業務」において株式会社三菱総合研究所が作成したものである。

¹ 所属は 2008 年 3 月時点

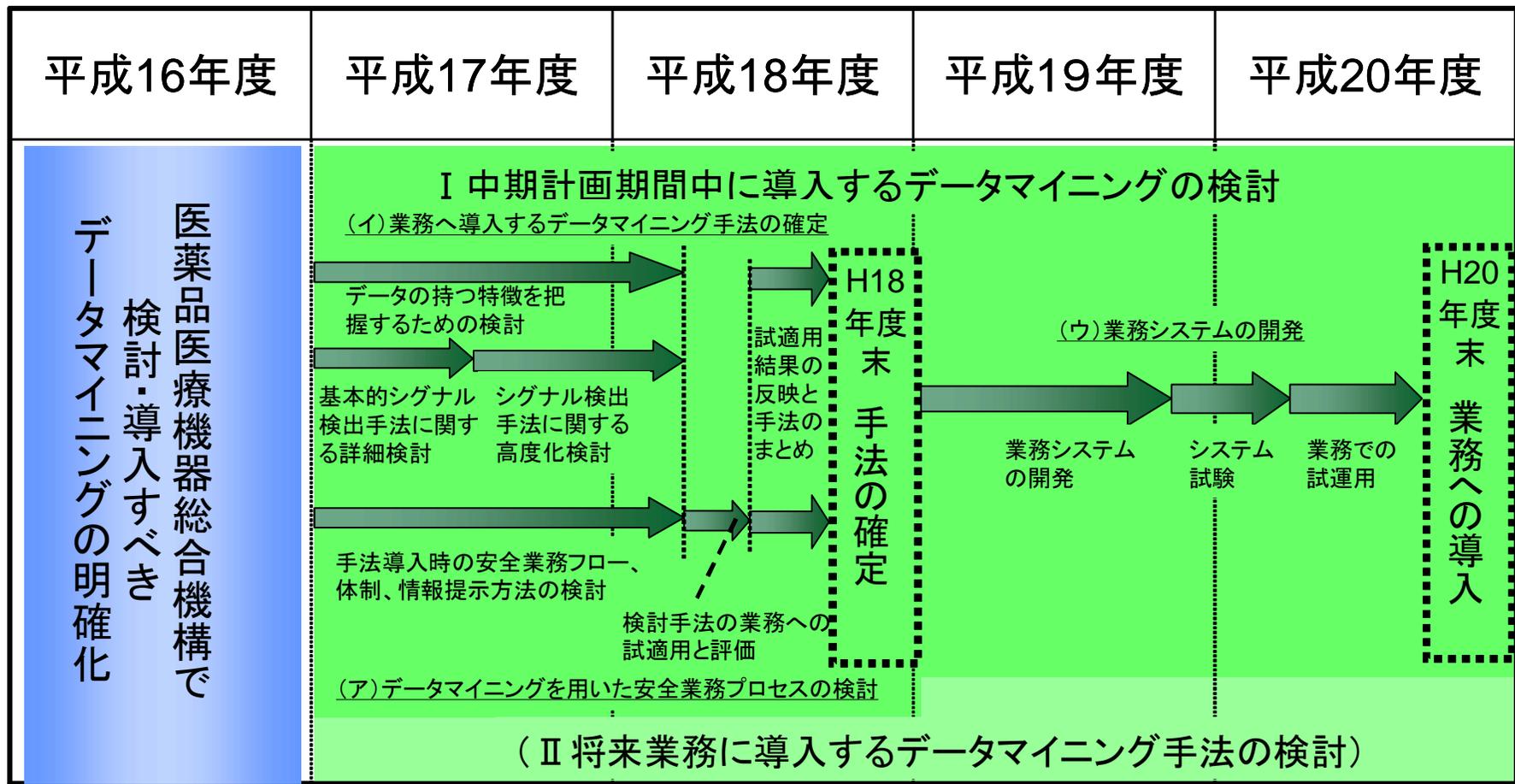


図 1-1 中期計画期間中のスケジュール

2. 海外規制当局におけるデータマイニングシステムに関する調査

2.1 海外調査概要

安全対策業務にシグナル検出手法を取り入れたデータマイニングシステムを導入している海外規制当局を訪問し、ヒアリング調査を行った。主な調査項目は以下の通りである。

- ・ シグナル検出手法を含んだ安全対策業務フロー
- ・ 業務フロー検討経緯
- ・ 実運用されているデータマイニングシステムの概要（インタフェースや機能）

訪問先および訪問日時を表 2-1 海外調査訪問先に示す。訪問先の 2 機関はいずれもシグナル検出手法を実業務に導入しており、上記項目について情報を得ることで、PMDA で開発するデータマイニングシステムおよびシグナル検出手法を用いた新業務フローの検討をする上での参考とすることを目的として、調査を実施した。

表 2-1 海外調査訪問先

訪問先	Lareb (Netherland Pharmacovigilance Centre : オランダ薬剤監視センター)	EMEA (European Medicines Agency : 欧州医薬品庁)
訪問日時	2007年11月16日 10:00~17:00	2007年11月20日 9:30~17:30
面会者	Eugene van Puijenbroek, Head scientific department Jerry Labadie, Regional officer	Noel Wathion, Head of Unit Panos Tsintis, Head of Sector Sabine Brosch, Deputy Head of Sector Francois Maignen, Principal Administrator Paolo Alcini, Scientific Administrator Jim Slattery, Scientific Administrator Xanier Kurz

2.2 Lareb

Lareb はオランダの規制当局である MEB (Medicines Evaluation Board) からの基金で設立した国立の機関であるが、他の機関からは独立した組織である。Lareb では 2000 年から安全対策業務に ROR 手法²を導入している。

2.2.1 オランダの副作用報告制度の概要

オランダの副作用報告制度の流れを図 2-1 に示す。副作用報告は規制当局である保健・福祉・スポーツ省内に設置された MEB が収集している。この MEB への報告経路には大きく製薬企業 (DSO) からのものと Lareb を経由するものの 2 つがある。2 箇所からの報告が MEB の所管する副作用報告データベース上で統合されることとなる。この副作用報告データベースの管理は Lareb に委託されており、Lareb で副作用報告データの管理と副作用報告の分析が行われる。Lareb での分析結果は MEB に報告され、MEB が措置の有無を決定する。

DSO 経由の報告は重篤な副作用のみであり、Lareb 経由の報告には軽微・重篤いずれの副作用も報告される。Lareb には医師や薬剤師といった医療従事者から報告に加え、2003 年からは患者から Lareb へ直接報告できるようになっている。

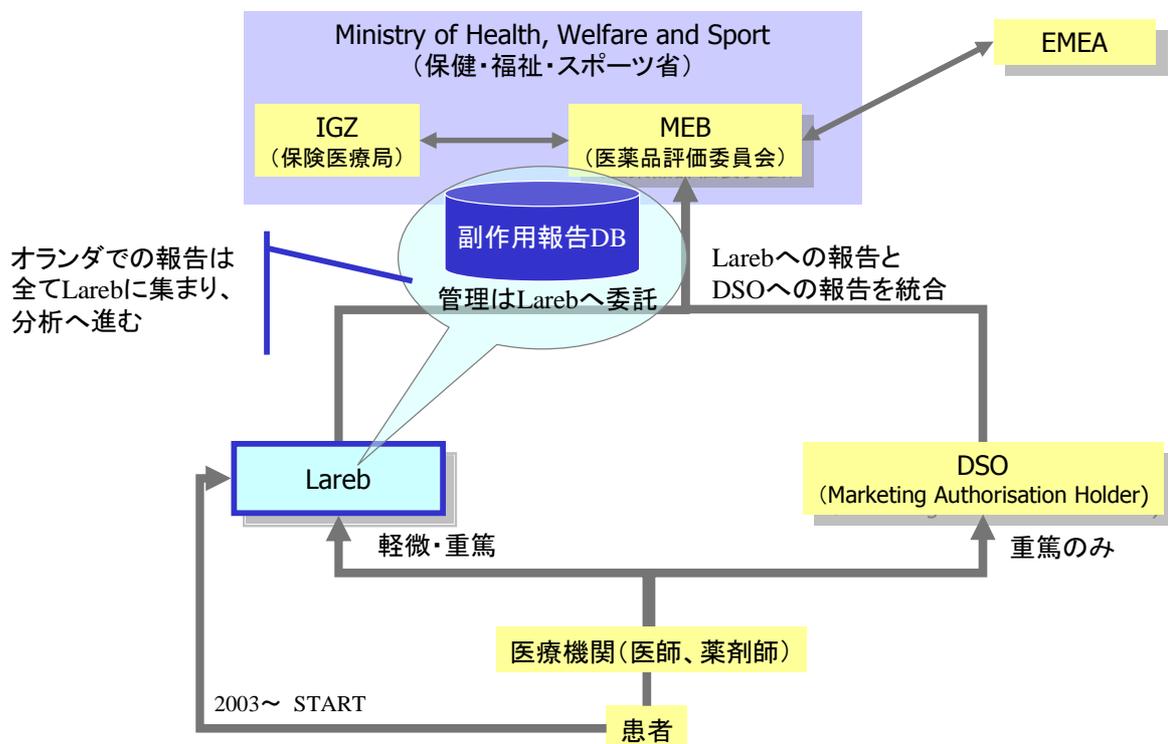


図 2-1 オランダの副作用報告制度

² 詳細は参考文献[1] p.30 を参照のこと。

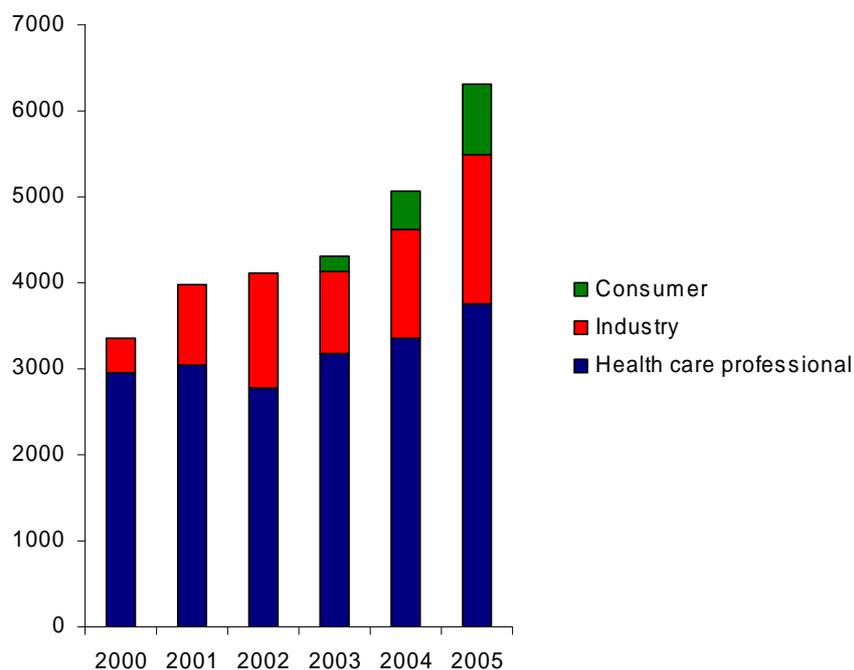


図 2-2 オランダの副作用報告件数推移

図 2-2 にオランダの副作用報告件数の推移を示す。患者からの報告、製薬企業からの報告、医療従事者からの報告の内訳も併せて図中に示されている。なお、2006 年の全報告件数は 6334 件、うち 830 件は患者からの報告である。

全報告の 7~8 割が電子報告であり、紙での報告についても PDF ファイルとしてスキャンし、電子データ化されているため、データベースには全報告が納められている。なお、患者からの報告はインターネットによる電子報告のみである。

報告様式は E2B に準じており、コーディングには副作用に MedDRA、医薬品にオランダ国内の独自コードが利用されている。また、オランダでは、被疑薬は最終的にひとつに特定した上で報告されている。

2.2.2 Lareb の構成

Lareb ではオランダ国内を 5 つの地域に分割し、地域ごとに報告を収集するという形を取っている。各地域の国立病院内にオフィスが設置されており、Regional Officer と呼ばれる担当者が 5 つの地域に 1 名ずつ（計 5 名）配置されている。図 2-3 に Lareb の構成を示す。5 つの地域は報告件数のボリュームが同程度となるように区分されている。

Regional Officer は医師または薬剤師の資格を有した Lareb 職員であり、次の役割を担っている。

1. 副作用報告の受付

担当地域からの医療従事者および患者からの報告の窓口となる。受付けた報告の記載内容を確認し、必要があれば報告者への確認を行う。

2. 報告データの分析

担当地域からの報告の分析担当者となり、各報告の分析を実施する。

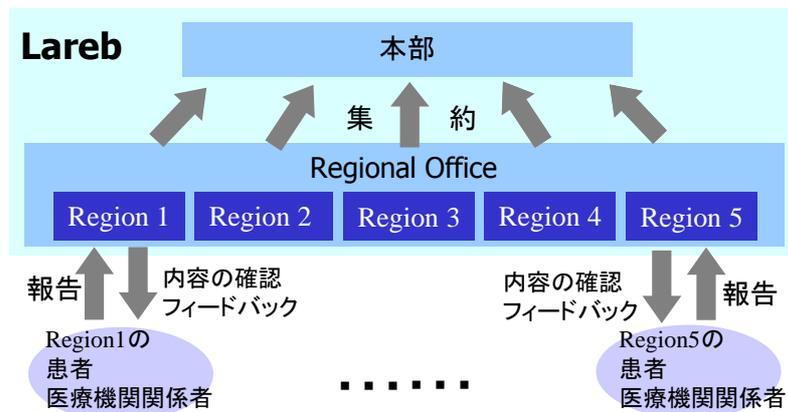


図 2-3 Lareb の構成

2.2.3 Lareb における安全対策業務フロー

Lareb における安全対策業務フローの概要を図 2-4 に示す。現在の業務フローは 1990 年代初頭に定まったものであり、2000 年からは業務フロー内に ROR 値が導入された。

Lareb では新規に報告された軽微なものも含めた全ケースについて必ずチェックを実施している。したがって、年間数千件（2006 年では 6000 件強）が分析の対象となる。Lareb では次のようなことが実施されている。

- ・ アセッサー分析 (case-by-case)
- ・ Scientific Review
- ・ スクリーニング
- ・ Analysis
- ・ フィードバックレポートの送付

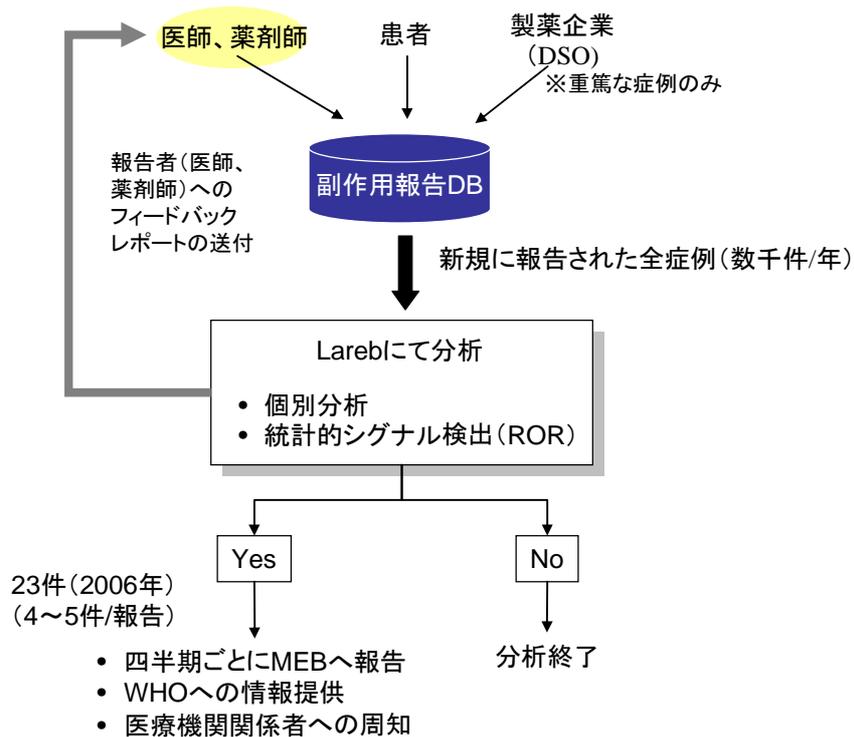


図 2-4 Lareb における安全対策業務フロー

(1) アセッサー分析 (case-by-case)

分析の第一段階であり、Regional Officer 5名によって実施される。分析対象は新規に報告された全症例であり、年間数千件である。分析には L@reb2002 システム（詳細は2.2.4に示す）を利用しており、システム上に表示される情報（報告フォーム記載内容、累積報告件数、ROR 値、等）から総合的に判断を下す。この際、標準化された基準や手順は設けられておらず、アセッサーごとに経験に基づいて判断する。アセッサー分析で要注目と判断された医薬品と副作用の組合せは (4) Analysis へ進むことになる。

(2) Scientific Review

毎週開催されるミーティング形式の分析の場であり、全新規報告を対象として実施される。ミーティングに参加できるメンバーは総勢 15 名程度であり、構成は次の通りである。

- ・ regional officer 5 名
- ・ アセッサー 複数名
- ・ アナリスト 4 名

上記メンバーのうち、1 回のミーティングには 3~5 名（アセッサーおよび医師が含まれる）が出席する。なお、当該ミーティングで評価対象となるレポートの担当アセッサーは必ず出席しなくてはならない。

ミーティングでは報告フォーム記載内容、累積報告件数、ROR 値等、得られる情報を参照しながら総合的に判断を下す。Scientific Review において要注目と判断された医薬品と副作用の組合せは (4) Analysis へ進むことになる。

(3) スクリーニング

(1) および (2) の分析を経た医薬品と副作用の組合せのうち、次段階へ進まなかったものについては、分析終了時にアセッサーが当該組合せを再分析するための条件を設定することになっている。再分析を実施する条件には次のような定量的な値が用いられる。

- ・ 累積報告件数
- ・ ROR 値の 95%下限値

これらの項目についてアセッサーは閾値を設定する。それ以降はシステム内で自動的にスクリーニングが実行されることになる。ある組合せが予め設定された条件に合致するとシステム画面上に自動的に表示されるようになっており、抽出された組合せについては、アセッサーは再度分析を行わなくてはならない。このスクリーニングという段階を設けることで、見逃しを防いでいる。

スクリーニングで抽出された組合せを再度分析した結果、当該組合せが要注目と判断された場合には (4) Analysis へ進むことになる。

(4) Analysis

前述した (1) ~ (3) のステップを経て、要注目と判断された医薬品と副作用の組合せについて、より詳細な分析を行うプロセスであり、次のような流れで実施される。

- ① アセッサー1名による再分析
- ② プレナリーミーティングの実施
- ③ Scientific Review のミーティングの場での検討

つまり、Scientific Review の場では (2) の段階での分析、および Analysis に進んだものの詳細分析、の 2 種の評価会議が実施されている。したがって、Analysis の③のミーティングは (2) と同じメンバーで行われる。

Analysis の対象となった症例は現在の業務フローがスタートしてから (5~7 年間で) 約 500 件である。

Analysis の結果、対応が必要だと判断された医薬品と副作用の組合せは、四半期ごとの MEB への報告、WHO への情報提供、医療従事者への周知がなされる。緊急に対応が必要と判断されたものについては四半期ごとの報告時期を待たずに MEB へ報告される。MEB への報告までに至った件数は、2006 年は 23 件 (1 回の報告につき 4~5 件) であった。

(5) フィードバックレポート

Lareb では医療従事者からの報告については、分析結果が報告者にフィードバックレポートとして返されている。Lareb では Regional Officer が地域に密着した活動を行っているため、このような取組みが実現できると考えられ、他の機関ではあまり見られない特徴的な取組みとなっている。

フィードバックレポートには分析結果の他に、レポート作成時点での当該組合せの累積報告件数等の定量的な値も記載される。ROR 指標値から得られる情報も記載されることがあるが、ROR 値自体を記載することはほとんどなく、アセッサーが数値を解釈した上で定性的な表現としている。

フィードバックレポートは自発報告の活性化につながると考えられており、重要な取組みとされている。

2.2.4 L@reb2002 システムの概要

Lareb で導入されているデータマイニングシステムは L@reb2002 と呼ばれるシステムである。E2B に沿った形式のインタフェースを有しており、さらにアセッサーの使いやすさを勘案し、業務フローに沿う形でシステムが設計されている。表 2-2 に L@reb2002 が有する主な機能を示す。

表 2-2 L@reb2002 の主な機能

実装機能	概要
症例票検索機能	<ul style="list-style-type: none"> ● 医薬品または副作用を指定しての検索が可能 ● 関連症例の呼出結果は症例概要一覧で表示、ここから詳細情報画面へ遷移
関連情報閲覧機能	<ul style="list-style-type: none"> ● WHO 提供データ (WHO-ART/MedDRA 対応マスター、ATC-code/オランダ医薬品コード対応マスター、UMC データ) ● CYP database : CYP による代謝機構が原因で発生する相互作用の予測に利用 ● 処方データ (保険請求における処方回数データ) 等
評価結果の入力・閲覧機能	<ul style="list-style-type: none"> ● アセッサー評価結果、フィードバックレポートの閲覧 等
ROR 値算出機能	<ul style="list-style-type: none"> ● ROR 値の計算の実行、結果の表示

各症例は safety report と呼ばれ、報告日時や患者情報、被疑薬・併用薬の情報、副作用の情報等から構成されている。

さらに、各症例に関する定量情報が提供されており、次のような情報が閲覧できる。

- **Disproportionality**
ROR 値、95%信頼区間、2×2 表の各度数 (a, b, c, d) を表示
指標値を色分けして表示することによりシグナル検出の有無を表現
赤：シグナル検出 (95%下限値が 1 以上)、緑：因果関係は薄い (95%上限値が 1 以下)、黒：その他
- **Interactions**
CYP による相互作用情報の提示

- Causality Assessment
Naranjo アルゴリズム³に基づき、算出されたスコアから因果関係を判定
- Prescriptions
処方量の推移をグラフで表示
処方量データは公的機関（public health insurance center）で収集
2003 年からは処方回数データに加え処方患者数のデータも収集

2.2.5 ROR の利用状況

Lareb では 2000 年から、安全対策業務にシグナル検出手法として ROR を導入している。

シグナル検出手法の導入に至った経緯は、報告件数の増加によりアセッサーが副作用報告データベース全体の状況を把握することが困難となったことによる。導入以前は分析対象組合せが他の報告全体に対してどの程度の割合を占めているか、等を感じ覚的に持ちながら分析が行われていたが、データベースの規模が大きくなるにつれて個人の記憶に頼ることが難しくなり、記憶に変わる指標としてシグナル検出手法が導入された。

ROR の採用理由は、「理解しやすく、シンプルで現実的、回帰分析にも利用可能」という利点からである。ROR 導入当初にはベイズアプローチ（MGPS や BCPNN）が未開発であったことも大きな理由の 1 つではあるが、ベイズアプローチでは過大評価となりアセ

³ Naranjo アルゴリズムとは因果関係の有無を判断するためのアルゴリズムであり、多数の機関でも採用されている一般的な手法である。以下のような 10 項目の質問で構成され、回答によって決定するスコア値に応じて因果関係の有無が判定できる。

質問項目	Yes	No	Don't Know
1. 当該副作用に関する以前の報告書が存在するか	+1	0	0
2. 被疑薬の投与後に AE が発現したか	+2	-1	0
3. 被疑薬中断したときに ADR は良くなったか	+1	0	0
4. 被疑薬を再投与したときに ADR が再発していないか	+2	-1	0
5. それ自体が当該副作用の原因と成り得る代わりの要因は存在するか	-1	+2	0
6. 偽薬を与えたときに当該副作用が発現しないか	-1	+1	0
7. 医薬品は中毒と知られる濃度の血液(または他の体液)から検出されたか	+1	0	0
8. 当該副作用は服用を増やすとより重篤になる、もしくは服用を減らすと軽減するか	+1	0	0
9. 同じ医薬品または類似薬による類似の副作用を以前に発症した経験があるか	+1	0	0
10. AE を支持する客観的な証拠があるか	+1	0	0

※合計スコア 9 以上→ADR であるのはほぼ確実 5～8→ADR の可能性が高い
1～4→ADR の可能性あり 0 以下 → ADR は疑わしい

ッサーの判断を誤らせる可能性も高いのではないかという懸念がある。

Lareb におけるシグナル検出手法の利用状況を表 2-3 に示す。

表 2-3 Lareb におけるシグナル検出手法利用状況

シグナル検出手法	ROR
シグナル検出閾値	$1 \leq \text{ROR}$ の 95% 下限値, and 累積報告件数(N_{11}) > 3
計算対象データ範囲	MEB 保有の副作用報告データベース内データ Lareb 収集分 (軽微+重篤) + DSO 収集分 (重篤)
シグナル検出指標値 算出頻度	ROR 値を確認する都度 Crystal Report 画面を表示する際に、そのときの 最新データを元に値を算出

計算対象データは、基本的には副作用報告データベース全体であるが、Lareb 収集分のみ、DSO 収集分のみ、いずれかでも計算できるようになっている。また、WHO Drug Dictionary と自国の独自コードとの対応マスターを所持していることから、WHO から提供されるデータを用いて ROR 値を算出することも可能である。さらに、これらの計算された ROR 値と、WHO から提供される WHO の副作用報告データベースで計算された BCPNN の指標値を並べて表示し、比較できるようになっている。

さらに、注目した医薬品群のみのデータセット (スモールセット) を作成し、これに対して指標値を算出できるような枠組みも用意されている。スモールセットによる分析はアナリストのみが実行できるものであり、アセッサーが利用しているシステムでは計算を実行することはできない。したがって、アナリストが分析を行った結果をアセッサーに提供する。

シグナル検出の閾値には ROR の 95% 下限値に加え、累積報告件数での条件を設けている。累積報告件数の閾値を 3 件とした理由は、これまで蓄積された分析経験、およびベイズアプローチの代表的な手法である BCPNN における指標値 IC025⁴との指標値の相関を研究した結果⁵による。

⁴ 詳細は参考文献[1] p.32 参照のこと。

⁵ ROR 値と IC025 値の相関は累積報告件数が 3 件以下では弱く、4 件以上で強くなるという研究結果が得られている[4]。

2.2.6 その他ヒアリング内容

(1) データクリーニング

オランダの副作用報告経路は前述したとおり、医療従事者、製薬企業、患者の3通りがあるため、同一症例が複数の経路から報告されることがある。したがってデータクリーニングにおいては、重複症例を排除することが重要となる。

重複症例の処理には次の3段階が準備されており、この検知システムによってデータクリーニングが行われている。

- ① システム上で患者の性別、誕生日（年齢）、被疑薬名の一致する他のレポートの有無を自動的にチェック、一致したものについては画面上に警告が表示され、アセッサーが最終的な判断を下す
- ② 定期的な蓄積全データを対象とした重複症例のチェック、予め指定したパラメータ（例えば、性別・誕生日・被疑薬、等）に基づきチェックをかける
- ③ アセッサーによる分析開始前の重複症例の有無のチェック

(2) アセッサーの育成方法

Lareb ではアセッサーを育成するための特定のマニュアルはなく、分析プロセスについて明文化したテキストも存在しない。Lareb は組織規模が小さいため、マニュアルを作成する必要はなく、先輩アセッサーについて経験を積むという形で教育を行う。なお、ROR に関しては、利用する上での利点とリスク（過大評価となる）が教えられる。

(3) 今後の展開

Lareb では現状の業務フローの見直しは予定されていない。現状のフローで問題はないとの認識である。さらに、ROR の利用に関しても見直しは予定されておらず、他の手法や閾値の変更も必要なく、現在の利用方法で十分実践的であると受け止められている。

2.3 EMEA

EMEAはEU承認薬に対する規制当局として1995年に設立された機関である。EU諸国内で市販されているEU承認薬(CAP⁶)の副作用報告を収集、分析し、安全対策業務を行っている。シグナル検出手法としてPRR⁷を採用しており、シグナル検出計算結果を提供するシステム(EVDAS: EudraVigilance Data Analysis System)やシステムを活用するガイドライン[6]を提供している。

2.3.1 EUの副作用報告制度の概要

欧州では各国での承認薬の他にEU承認薬も市販されている。各国での承認薬については各国の規制機関(オランダではMEB、イギリスではMHRA、等)により承認や市販後調査等が行われる。これに対しEU承認薬を規制するのがEMEAである。

EUにおける副作用報告の流れを図2-5に示す。EMEAに報告される副作用症例はEU内の各国機関(NCA)からのものと、製薬企業(MAH)からのものがあり、EU域内における重篤な副作用、およびEU域外での未知・重篤な副作用の情報を収集している。EU諸国の規制機関とは重篤な副作用報告についての情報を交換しており、製薬企業からはEU域内の重篤副作用、およびEU域外の未知・重篤な副作用の報告を受けている。ただし、EMEAが規制対象とする医薬品はEU市販薬であり、副作用症例の分析はEU承認薬に関する副作用報告のみである。

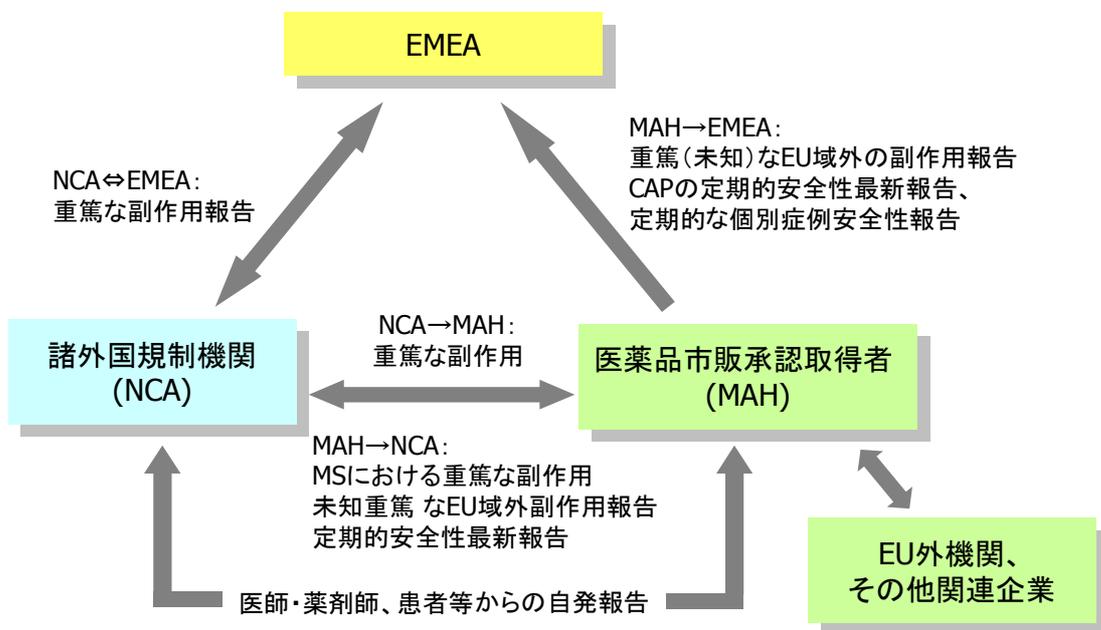


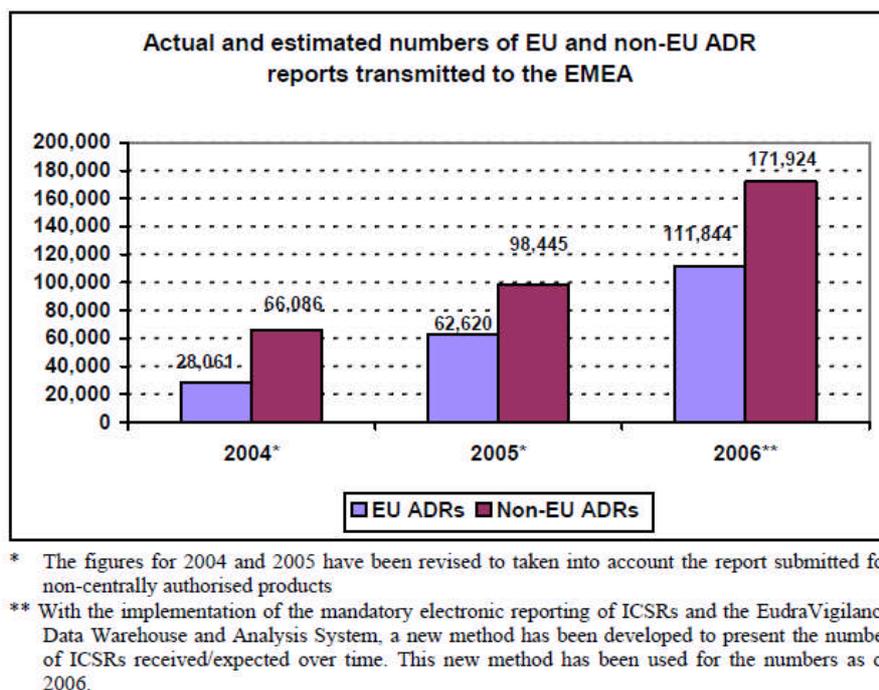
図 2-5 EU 副作用報告制度の概要

⁶ EU承認薬(CAP)は2007年11月現在、約480医薬品

⁷ 詳細は参考文献[1] p.31を参照のこと。

図 2-6 に EMEA で受け付けた副作用報告件数の推移を示す。2006 年末時点での累積報告件数は 677,976 件である。

EMEA への報告は電子報告が一般的となっており、26 カ国の規制機関、および全製薬企業の 95%強に当たる 201 企業が電子報告である。



出典：The European Medicines Agency in 2006

図 2-6 EU の副作用報告件数推移

2.3.2 EMEA における安全対策業務フロー

EMEA において分析の対象とするのは当該期間中の新規報告であり、該当する報告から SD (Signal Detection) print を作成する。この SD print を参照しながら各アセッサーは注目すべき医薬品と副作用の組合せを抽出し、抽出した組合せについては分析担当者によるミーティングで検討がなされ、対応の有無を決定する。

(1) 安全対策業務の実施体制

EMEA のアセッサーは医師、薬剤師、科学者で構成されている。人材育成は OJT で行われ、研修期間は専門や経歴ごとに異なる。つまり、標準化された教育プログラムは有していない。アセッサーは医薬品の種類に応じたグループに属しており、グループは取扱う

医薬品数が同程度となるように決められている。

分析を行うサイクルは医薬品に応じて2週間または1ヶ月ごととなる。2週間ごとに評価を行う医薬品は集中的モニタリング医薬品、1ヶ月ごとに評価を行う医薬品は定期的モニタリング医薬品と呼ばれる。集中的モニタリング医薬品は認可されてからの期間が短いもの等が含まれており、集中的モニタリング医薬品以外が定期的モニタリング医薬品に分類される。

(2) SD print の作成

当該期間における新規報告のうち、EU 市販薬で重篤な副作用のみを対象に作成される。医薬品と副作用の組合せごとに一行で構成されるラインリスト形式で副作用情報を一覧表示したものである。SD print の出力頻度は評価サイクルと同様である。

- 集中的モニタリング医薬品：2週間ごと
- 定期的モニタリング医薬品：1ヶ月ごと

SD print の出力項目を表 2-4 に示す。SD print は医薬品ごとにテーブルが作成されるようになっており、副作用は MedDRA の SOC もしくは PT タームで記載される。また、SD print 上には PRR 値に関連する計算結果も記載されている。

表 2-4 SD print 出力項目

基本情報	医薬品名、副作用名 (SOC、PT)
シグナル検出指標値	PRR 値、95%下限値、95%上限値 (閾値を超えた組合せは赤でマークして表示)
新規報告件数	新規報告総数、EU 内新規報告件数、EU 外新規報告件数、 新規死亡症例数
累積報告件数	累積報告件数、EU 内累積報告件数、EU 外累積報告件数

(3) SD print に基づく評価の実施

SDprint を元に、アセッサーは担当医薬品について分析を行う。分析は以下の流れで行われる。

1) 既知未知の分類

SD print 上の医薬品と副作用の組合せの既知未知の判断を行う。既知未知を分類する際には SmPC (欧州製品概要) や PIL (患者向け医薬品情報書) を参照する。分類の結果、既知と判断されたものについては分析の対象から除外される。

2) 注目すべき組合せ候補リストの作成

未知症例について **SD print** 上の情報から注目すべき組合せをリストアップする。

PRR 値も判断項目の 1 つとはなるが、**PRR** を用いた統計的な判断のみで組合せを抽出することはしていない。また、アセッサーが **PRR** の数値を見る際には、閾値を超えているか（赤のマークがあるか）を主にチェックしており、**PRR** 値の数値自体のチェックはしていない。

3) リストアップされた組合せについての詳細評価の実施

リストアップされた組合せについて、**SD print** 上に表示された情報以外の情報も参照しながら、最終的にミーティングへ進めるべき組合せの抽出を行う。また、この段階で重複症例の確認やデータ精度についてのチェックも行われる。

詳細評価の結果、ミーティングに進めるべきと判断された組合せについては、アセッサーがショートレポートを作成する。

この評価プロセスの手順は標準化されておらず、アセッサーの経験によって分析が行われている。現在、標準的な分析手順の検討が進められており、ドラフト段階まで進んでいる。検討結果が公開されるかは未定である。

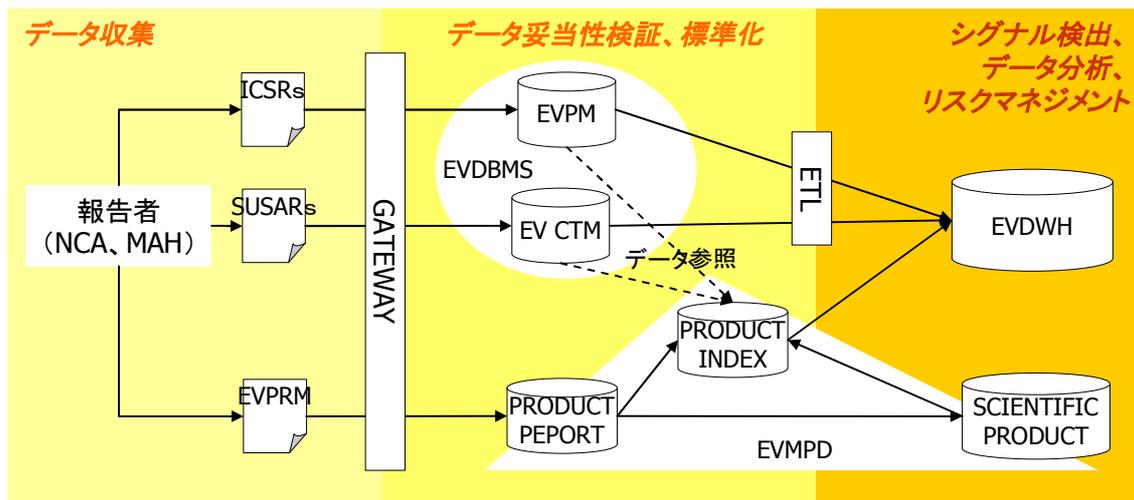
(4) アセッサーミーティング

(3) の結果、抽出された医薬品と副作用の組合せはミーティングにかけられ、最終的に対応を行うか否かを決定する。ミーティングには医薬品ごとに分けられた各グループから 1、2 名ずつが参加する。この段階で初めて他のグループの担当医薬品についての報告内容が共有されることになり、異なるグループのメンバー間でクロスチェックが行われることになる。

ミーティングの結果、対応を要する、モニタリングを継続、対応必要なし、いずれかの判断が下される。

2.3.3 EudraVigilance システムの概要

EMEA では副作用報告（承認前・後）の処理からマネジメントまでを統一的に扱う EudraVigilance システムが利用されている。EudraVigilance システムの構成を図 2-7 に示す。



出典：EMEA サイト内掲載資料を元に作成

(<http://www.emea.europa.eu/pdfs/conferenceflyers/sme/5-3maignen.pdf>)

図 2-7 EudraVigilance システムの構成

(1) EVDBMS (EudraVigilance DataBase Management System)

E2B に準拠して構築された副作用報告データベースシステムである。EVDBMS は主に承認後モジュール (EVPM)、および治験モジュール (EV CTM) の 2 つのモジュールから構成される。

EVDBMS には複数機関 (EU 諸国、製薬企業) からのデータの翻訳、統一コードへの変換、重複症例のチェック、等を自動的に行うアルゴリズムが実装されている。自動変換可能なデータは全体の 75%程度であり、自動変換できないもの (25%) については担当者 3 名が手作業で修正を行っている。EMEA で用いられている医薬品および副作用のコードは以下の通りである。

- 医薬品：EVMPD (EudraVigilance Medical Product Dictionary、EMEA 独自の医薬品データ体系、階層構造を持つ)
- 副作用：MedDRA

(2) EVMPD (EudraVigilance Medical Product Dictionary)

EVDBMS は Product Report Database、Scientific Product Database、Product Index Database の 3 つのデータベースから構成される。

- Product Report Database : データ収集をサポートするために設計されたデータベース。EVPRM (EudraVigilance Product Reporting Message) と呼ばれる EU 内の MAH や治験依頼者から提供される承認や開発段階での関連情報が含まれるデータが蓄積される。
- Scientific Product Database : データ分析をサポートするために設計されたデータベース。EVMPD に蓄積された医薬品について、グルーピングや階層化を行う。
- Product Index Database : 医薬品のコードと当該医薬品についての基本的な情報 (医薬品の正式名称、略称、一般名、企業名、薬効と薬効番号、剤型) の対応を提供するデータベース。Product Report Database および Scientific Product Database のデータを元に作成される。

EVDBMS に格納された副作用報告データは Product Index Database を参照することになる。

(3) ETL (Extraction, Translation, Loading)

副作用報告データ、EVMPD、機関情報、MedDRA 等の情報を統合し、後述する EVDWH に格納するためのデータコンバージョンシステムである。

(4) EVDWH (EudraVigilance Data Warehouse)

副作用報告のケースレビューとデータ分析をサポートするためのシステムであり、この部分がデータマイニングシステムに相当する。この中に EVDAS (EudraVigilance Data Analysis System) があり、統計的手法 (PRR) による計算結果 (表、グラフ等) の表示や CIOMS フォームの表示 (ラインリストで表示された一覧画面から遷移する) 機能を有している。また、アセッサの分析内容を入力することで履歴を蓄積することもできる。

2.3.4 PRR の利用状況

EMEA では統計的手法として PRR を採用している。他のシグナル検出手法に関しても検討を進めており研究はなされているが、アセッサが分析業務で利用しているのは PRR 手法である。PRR を採用した理由は、技術的に導入が容易でありアセッサが理解しやすい点にある。ただし、統計的手法は判断する材料の 1 つという位置付けでありシグナル検出計算によるフィルタリング等は実施されていない。PRR 値を含めて得られる情報から総合的にアセッサが判断を下している。

表 2-5 に EMEA におけるシグナル検出手法の利用状況を示す。PRR の閾値としては①および②の 2 種類が提案されており、PRR では①の閾値を採用している。これは、これまでの検討において①の方が②に比べて false-positive が少ないと判断したためである。

表 2-5 EMEA におけるシグナル検出手法利用状況

シグナル検出手法	PRR
シグナル検出閾値	① PRR の 95% 下限値 ≥ 1 and $N_{11} \geq 3$ ② $PRR \geq 2$ and $Chi-sq. \geq 4$ and $N_{11} \geq 3$ ※①の閾値をメインに利用
計算対象データ範囲	EMEA 保有の副作用報告データベース内データ (重篤事象のみ)
シグナル検出指標値 算出頻度	集中的モニタリング医薬品：2 週間ごと 定期的モニタリング医薬品：1 ヶ月ごと

なお、EMEA では 2005 年の後半から PRR を含めたシグナル検出手法の妥当性評価の検討が進められており、検討結果は 2008 年中旬に公表される予定である。

2.4 まとめ

安全対策業務にシグナル検出手法を取り入れたデータマイニングシステムを導入している海外規制当局として、オランダの **Lareb**、EU 機関の **EMA** を訪問し、ヒアリング調査を行った。本調査から次のような知見が得られた。

- 安全対策業務へのシグナル検出手法の導入

Lareb では **ROR**、**EMA** では **PRR** を安全対策業務へ導入し、アセッサーがシグナル検出手法から得られる値を分析の判断材料に用いている。したがって、いずれの機関においても、アセッサーが数値の意味を理解できる（わかりやすい）ことを重視し、比較的計算式がシンプルな手法を採用している。

- シグナル検出指標値の位置付け

Lareb、**EMA** どちらの機関においてもシグナル検出指標値のみによる評価対象とする医薬品と副作用の組合せのフィルタリングは行っていない。計算値は参考情報の 1 つであり、他に報告症例内の他の情報やアセッサーの背景知識から総合的な判断を下している。

日本におけるシグナル検出指標値の位置付けも **Lareb** や **EMA** と同様であり、これらの機関の業務フローやシステム構成は大いに参考になると思われる。

- シグナル検出指標値に対するアセッサーの理解度

Lareb ではアセッサーは **ROR** 値の意味を理解して評価を実施している。したがって、アセッサーには閾値を超えたか否かのシグナル検出の有無の情報のみではなく、指標値自体と、2×2 表の各度数が提示されている。

これに対し、**EMA** のアセッサーは閾値を超えている（シグナル検出されている）か、否かを判断指標としており、アセッサーには指標値とシグナル検出の有無の情報が提示されているが、値は分析に利用されていない。

シグナル検出手法による計算結果はアセッサーが分析の際に利用するものであり、その理解を深めることは重要である。また理解度に応じた適切な提示方法としなければならないということも忘れてはならない。

- 安全分析業務フローに沿ったシステムを整備

Lareb では **L@reb2002** と呼ばれる業務をサポートするシステムが整備されており、画面上で症例情報、関連情報の閲覧、評価の履歴の入力等が可能となっている。**L@reb2002** ではアセッサーが必要とする情報を全て統一システムにより提供されており、安全対策業務をサポートする理想的なシステムである。

EMEAでもEVDBMSを構築しており、業務をサポートする複数のツール(SDprint作成、EVDAS、分析履歴の入力)を利用しながら分析が行われている。特にEVDASはEU内の各国機関にも提供されている統計解析システムであり、シグナル検出計算結果の提示方法は参考とできる点も多い。

最後に、2005年の海外調査の結果⁸も加え、海外規制機関のシグナル検出手法を取り巻く状況を表2-6に示す。

⁸ 2005年の調査ではWHO Uppsala Monitoring Center、MHRA（イギリス）、FDA（アメリカ）を訪問し、ヒアリング調査を実施した。詳細は参考文献[2]の2章（p.5～p.23）を参照のこと。

表 2-6 海外規制機関比較表

機関名		PMDA	WHO (UMC)	MHRA	FDA	Lareb	EMA
国名		日本	--	イギリス	アメリカ	オランダ	EU
調査年		--	2005年	2005年	2005年	2007年	2007年
副作用報告制度	自発報告対象	副作用重篤(2005年以降)	各機関での報告制度による(重篤・非重篤どちらも含まれる)	全て(重篤・非重篤)	全て(重篤・非重篤)	全て(重篤・非重篤)	EU域内の重篤副作用、EU域外の未知・重篤副作用
	報告元	製薬企業	83の医薬品モニタリングプログラムメンバ国	医療従事者(医師、歯科医、薬剤師、看護師、他)／製薬会社／患者	医療機関／患者／製薬会社	医療機関／患者／製薬会社	諸外国規制機関／製薬会社
データベース	概要	2003年10月から電送開始	各国からの報告はVigibaseと呼ばれるDBに格納した後、分析プロセスに沿ったDBへデータを順に格納	Yellow Card DB	AERS、VAERS	DBはMEB(医薬品評価委員会)が所管、管理をLarebへ委託	EVDBMSと呼ばれるシステムでEUの副作用報告(承認前・後)の処理からマネジメントまでを統一的に扱う
	副作用コード	MedDRA	WHO-ART※2	MedDRA	MedDRA	MedDRA	MedDRA
	医薬品コード	医療用:再審査コード 一般用:経済課コード	WHO Drug Dictionary	MHRA開発の独自コード	FDA開発の独自コード	オランダ独自コード	EMA独自コード(EVMPD)
	規模	累積報告件数(国内・2007年末時点) 約164,000件(83,000例) 年間報告件数(国内) 約50,000件以上(約25,000例)	年間報告件数 約20万件 累積報告件数 約350万件	年間報告件数 約17,000~18,000件	累積報告件数約250万件	2006年は6,334件	2006年は111,844件(EU域内) 2006年末蓄積データ総数 ICSR: 677,976件
定量的手法	導入手法	ROR/BCPNN/GPS	BCPNN	PRR(2005年時点)※1	MGPS	ROR	PRR
	基準値	文献値	IC025(ICの95%下限値)>0	PRR>=2 and Chi-sq.>=4 and N11>=3	EB05>=2	1≤ROR値 95%下限値, and 累積報告件数(N11)>3	PRR値の95%下限値>=1 and N11>=3
	計算実行間隔	基本は1週間、任意で逐次	3ヶ月	1週間		ROR値を確認する都度	2週間/1ヶ月 (医薬品の分類により異なる)
分析プロセス	分析対象	報告された全症例	報告された全症例 (人が目を通すのはトリアージ手法を通過したもののみ)	報告された全症例		報告された全症例	2週間または1ヶ月間に新規に報告された症例
	利用システム	開発中	--	分析プロセスに沿った分析システムを利用 (リンカーン社システムの導入が決定)	WebVDME	L@reb2002	EVDBMS内のEVDWH (Eudra Vigilance Data Warehouse)
	プロセスと担当者数	新業務フロー 1. 因果関係・既知未知の判断をする個別症例評価の実施 2. 機械的な判断を含む1次スクリーニングの実施 3. 1次を通過したものについて詳細な検討を行う2次スクリーニングを実施 4. 要対策案件として、外部専門委員による検討会の開催し措置を決定	1. Vigibaseデータの医薬品と副作用の組合せへ分解し、Total Combination DBへ格納 2. Total Combination DBデータを対象にBCPNNを算出(3ヶ月ごと)し、基準に合致する組合せを抽出、Combination DBへ格納(Combination DBデータの各国規制当局への提供) 3. Total Combination DBデータをトリアージ手法によりフィルタリング 4. 3を通過した組合せについての内部アセスメントの実施(内部の専門家4名(うち1名は専任)が実施) 5. 4で詳細検討が必要とされた組合せについての外部アセスメントの実施(外部専門家も加わる)	1. インパクト分析により担当者の評点付け、カテゴリ分類の実施(担当者10~20名で実施、データ入力も含めた人員総数は50名程度) 2. 1でカテゴリA(優先度高)に分類された組合せに対するケースレビューの実施(各担当者が実施) 3. リスクアセスメント会議の開催(専門家で構成される複数名で実施)	標準的なプロセスについて議論中(調査年時点)	1. アセッサ分析(毎日実施、Regional Officer 5名) 2. Scientific Reviewの実施(毎週実施、ミーティング参加者3~5名) 3. スクリーニング(不定期実施、システム上で自動) 4. 1~3のプロセスから抽出された注目すべき症例へのAnalysisの実施(毎週実施、ミーティング参加者3~5名)	1. ラインリスト(SDprint)の作成 2. 注目すべき組合せの抽出とショートレポートの作成(1ケースを2名、総勢20名) 3. 2で抽出された組合せを対象とするSDGミーティングの実施(8~10名)
	定量的手法の利用場面	1次スクリーニング時の基準のひとつ、ならびに、各プロセスにおいて参考情報として利用する予定。	分析者によるアセスメントを行う組合せを抽出するトリアージ手法におけるフィルタの1つとしてBCPNNの基準値を採用している。	プロセス1のインパクト分析においてフィルタリング基準のひとつとしてPRRの値を算出し、この値に基づき評点付けを行っている。	非定常業務において、注目している医薬品や副作用の分析に利用(2005年訪問時の回答)	全プロセスにおいて分析者が判断する際の参考情報の1つとして利用されている	SDprint上に指標値を掲載し、2の段階において分析者が判断する際の参考情報の1つとして利用されている
特徴および参考とすべき点		--	BCPNNのみではなく他のルールも組合せてフィルタリングを行うことで、SD手法の弱点を補っている点は参考になる。ただし、規制当局ではなく扱うデータ量も膨大である点はPMDAと状況が異なる。そのため、ここで採用しているフィルタリングをそのままPMDAへ導入することは難しい。	PRR値のみではないインパクト分析という枠組みにより、実用に耐える体系的な方法を確立している先進的事例と言える。同様の形をすぐにPMDAに導入することは難しいと考えられるが、定量的手法の利用方法として今後も継続的に情報を得ることで参考とできる点は多いと思われる。	MGPSという独自の評価指標を用いた、優れたヒューマンインタフェースを有するシステム開発を行っている。非定常業務をサポートするシステムとして、参考とすべき点が多い。	利用システムL@reb2002は分析プロセス全般をサポートするシステムとして優れたIFを有しており、PMDAの新システム設計に参考とすべき点が多い。また、ROR値はアセッサの参考情報の1つであるという位置付けはPMDAにおける位置付けと合致しており、大いに参考とできる。	PRR値はアセッサの参考情報の1つであるという位置付けはPMDAにおける位置付けと合致しており、大いに参考とできる。また、EMAではPRR以外の手法も含めた手法の妥当性検証を進めており、この検討結果については注目すべきである。

※2 2008年3月からMedDRAも可能に ※1 2006年からはMGPSを利用

3. 新業務フローの検討

2009年のデータマイニングシステムの導入にともない、安全対策業務フローについてもシグナル検出手法を活用できるような検討が必要である。このシグナル検出手法を取り入れた安全対策業務フローを新業務フローと呼び、本章において検討経過をまとめた。

3.1 分析者アンケートの実施

3.1.1 アンケート概要

新業務フローの検討を行うにあたり、従来の安全対策業務フローを整理するために分析担当者8名を対象に、分析の視点等を確認するアンケートを実施した。本アンケートにより、新業務フロー内に含むべき項目、および本年度開発がスタートしたデータマイニングシステムに実装すべき機能の過不足を確認することができる。

アンケートは、特定のケースを想定して回答するパートと一般的な視点として回答するパートの2つのパートで構成した。想定する特定のケースは、担当した措置済み案件のうち重大な副作用を改訂した直近の案件とした。

アンケート作成において、分析業務におけるチェック項目候補はMHRAで導入されているインパクト分析の項目[8]を参考にした。

3.1.2 アンケート項目

アンケートでは分析担当者の方が、分析時にどのような情報をチェックしているかを尋ねる構成とした。チェック事項としてアンケートに示した主な項目は、以下の通りである。

- 報告件数に関する項目
 - ・ 集計単位（コンビネーション⁹、案件¹⁰またはクラス¹¹）ごとの累積報告件数
 - ・ 集計単位（コンビネーション、案件またはクラス）ごとの直近1年間の報告件数
 - ・ 集計単位（コンビネーション、案件またはクラス）ごとのトレンド
 - ・ 類似の医薬品や副作用内での報告件数の相対比較 等
- 副作用症例報告書記載内容に関する項目
 - ・ 投薬時期との時間的關係
 - ・ 類似副作用の報告状況

⁹ コンビネーションとは、医薬品の一般名（再審査コード上7桁）に副作用名（MedDRA PT）が一对一で構成される単位

¹⁰ 案件とは、医薬品の安全対策等の検討に用いられる単位で、複数の医薬品の複数のコンビネーションをグループ化したもの

¹¹ クラスとは、複数の案件をグループ化したもので、安全対策業務上の必要に応じて作られる単位

- ・ 相互作用の可能性
- ・ 死亡例の割合 等

以上の項目に加え、現在はシステム上で提供されていないが、容易に参照できれば参考にしたい情報（例えば、処方量データ、報告件数トレンド、層別での動向、等）についての要望を尋ねる問を設定した。

3.1.3 アンケート結果

前述した通りアンケートは、特定のケースを想定して回答するパートと一般的な視点として回答するパートの2つのパートから構成しているが、ここでは一般的な視点での回答結果を示す。

表 3-1 には報告件数に関する項目での回答結果を示す。

コンビネーション単位、および案件またはクラス単位での累積報告件数、直近1年間の報告件数、報告件数のトレンドについて確認状況を回答いただいた。累積報告件数についてはいずれの集計単位でも全員が確認しており、直近1年間についても概ね確認しているとの回答結果を得た。

報告件数のトレンドについても概ね確認しているとの回答を得た。現在はトレンド情報を自動で確認できる仕組みとはなっていないため、分析者は自身の記憶もしくは自身で集計を行っていると考えられる。したがって、トレンド情報を分析者に自動的に提供できる機能は有益と言える。

表 3-1 報告件数に関する項目での回答結果

確認内容		確認率		
		累積	直近1年間	トレンド
評価 単位	コンビネーション	100%(8/8)	88%(7/8)	75%(6/8)
	案件/クラス	100%(8/8)	75%(6/8)	75%(6/8)

表 3-2 には、副作用症例報告書記載内容に関する項目の回答結果を示す。

副作用症例報告書の記載内容については、設問とした項目について、ほぼ全員が確認しているという回答であった。

表 3-2 副作用症例報告書記載内容に関する回答結果

確認内容	確認率
被疑薬の投薬時期（使用開始、中断、再開）と副作用の発生に時間的な関係が存在するか。	100%(8/8)
同一症例で報告されている他の被疑薬・併用薬について、同様の副作用が既に知られている、または既に報告されていないか。	100%(8/8)
特定の医薬品と併用した場合に、当該副作用の発生が確認できるか。	88%(7/8)
副作用は、当該患者の原疾患、他疾患に見られるものか。	100%(8/8)
医薬品と副作用の1対1のコンビネーション（ラインリストの1行）が含まれる全症例に占める死亡例（副作用転帰欄が「死亡」）の割合が高くないか。	100%(8/8)
評価対象の案件（またはクラス）に含まれる症例中の死亡例（副作用転帰欄が「死亡」）の割合が高くないか。	100%(8/8)
評価対象の症例（または症例群）における障害・後遺症レベルは高くないか。	100%(8/8)

表 3-3 に副作用症例報告書以外の項目に関する回答結果を示す。本項目についても確認している割合がかなり高い。副作用症例報告以外にチェックしているデータとしては、研究報告が 100%であり、海外措置報告と CCDS もかなりの割合でチェックされている。ここで確認していると回答されたデータについては、新システムで容易に分析者に提供されることが望まれる。

表 3-3 副作用症例報告書以外に関する回答結果

確認内容	確認率
作用機序	88%(7/8)
類似医薬品での同種副作用	100%(8/8)
副作用報告以外のデータによる因果関係確認	88%(7/8)
確認する情報	
海外措置報告 :	88%(7/8)
面会記録 :	50%(4/8)
研究報告 :	100%(8/8)
CCDS :	88%(7/8)
	その他 (海外添付文書 ×2)

表 3-4 今後提供が期待される情報に関する回答結果

必要な情報	Yes の割合
処方量データ	75%(6/8)
当該医薬品について報告されている全副作用のリスト・件数	63%(5/8)
当該副作用について報告されている全医薬品のリスト・件数	13%(1/8)
報告件数の推移	75%(6/8)
類似薬における副作用発生傾向比較	75%(6/8)
年齢層と報告傾向	50%(4/8)
性別と報告傾向	38%(3/8)
併用薬による相互作用の有無	38%(3/8)
その他	
<ul style="list-style-type: none"> 副作用 DB の全データを日々 CSV として抽出してほしい。 副作用発生傾向の変化を定量的に示す情報 	

表 3-4 には今後提供が期待される情報に関する回答結果を示す。必要だとの回答が多かったのは、処方量データ、報告件数の推移、類似薬における副作用発生傾向であった。報告件数の推移に関しては、報告件数に関する設問においても、既にトレンドを確認しているとの回答が得られており、自動的な提供が望まれていることが確認できる。

「当該医薬品について報告されている全副作用のリスト・件数」および「当該副作用について報告されている全医薬品のリスト・件数」については、前者の要望が強いことがわかった。分析者は医薬品をキーとして分析を進めることから、このような結果となったと推測される。この回答から、前者のリストの整備を優先させるべきと判断できる。

さらに、年齢層と報告傾向、性別と報告傾向、併用薬による相互作用の有無についても半数近くが必要と回答しており、今後、データマイニングの高度化に向け引き続き検討していくべきであることが確認された。

3.1.4 アンケートまとめ

新業務フローの検討を行うにあたり、従来の安全対策業務フローを整理するために分析担当者 8 名を対象に、分析の視点等を確認するアンケートを実施した。本アンケートにより、以下の知見が得られた。

- 評価に用いられる項目に関して
 - ・ 報告件数以外の情報も含め、様々な情報に基づき総合的な評価が行われていることが確認された。
 - ・ 特に、報告件数については累積報告数以外の直近期間の報告件数、トレンド等の情報も評価に用いられており、これらの情報を業務システムで容易に確認できることが重要である。
- 評価項目の一般性に関して
 - ・ 一般的な視点としての回答からは、回答者の多くが同様の項目に基づいて評価を行っていることが確認された。
 - ・ 一方、特定のケースを想定しての回答結果からは、案件ごとに個別の評価に基づき判断が行われる場合があることが確認された。
 - ・ したがって、標準的な評価項目を準備する際には、特定ケースでの差異を考慮する必要がある。
- 他の評価用情報源に関して
 - ・ 処方量データ等へのニーズが高いことが確認された。

3.2 新業務フローの検討

本年度の検討では分析担当である医薬品安全課の方にも加わっていただき、新業務フローについても検討を重ねた。新業務フローでは新規に報告された症例について、まず始めに未知既知の判定を行い、未知か既知かによってその後の分析フローを異なる形とすることを検討している。

データマイニングシステムを活用した新業務フローのうち、未知症例についての業務フロー案を図 3-1 に示す。フローは主に次の3段階からなる。

- ① 個別症例評価（因果関係評価）
- ② 1次スクリーニング
- ③ 2次スクリーニング

各評価プロセスの概要を図 3-1 に示す。まずは症例単位で因果関係の評価を行い、評価記号を付ける。次に1次スクリーニングにより案件評価に進むべきものの抽出と、2次スクリーニングでの評価単位となる案件を作成する。1次スクリーニングにおける抽出は主に分析者による判断、および因果関係評価の結果付与される評価記号の蓄積件数、シグナル検出計算結果により行われる。

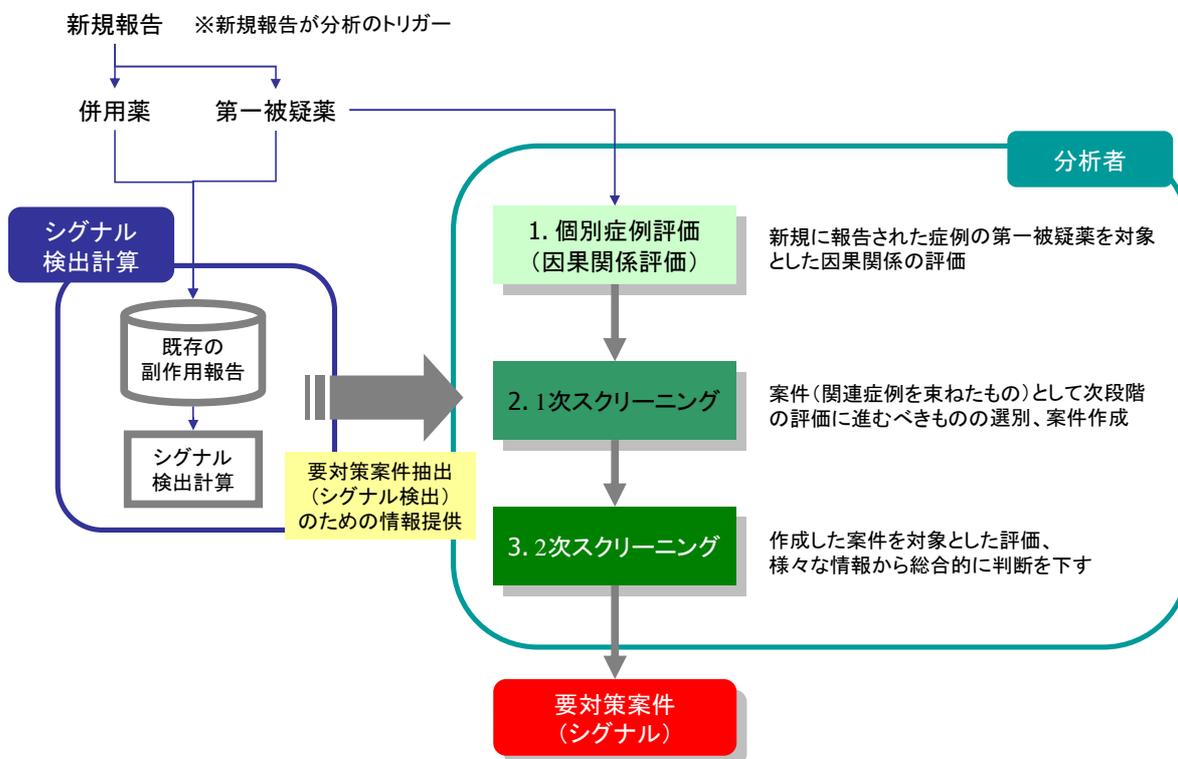


図 3-1 未知症例に関する新業務フローの概要

1 次スクリーニングにシグナル検出手法を導入することで、人の目だけでは発見されないシグナルを検出することが期待される。さらに、シグナル検出手法に用いるデータの度数は被疑薬群（第一被疑薬＋併用被疑薬）により作成することとする。人の目での分析は第一被疑薬を基本的に対象とするため、シグナル検出手法の導入により、より広い範囲をカバーすることができる。

2 次スクリーニングでは、1 次スクリーニングを経て作成された案件について、案件単位での評価が行われる。分析者は評価する上で、関連する情報（定量的、定性的いずれの情報も含む）を参照し、総合的に判断を下す。分析者に提供される情報の中にはシグナル検出指標値も含まれる。

3.3 1 次スクリーニングの試行実験

新業務フローの3つの段階うち、1 次スクリーニングは、人手によるスクリーニングと累積報告件数やシグナル検出指標値を用いて自動的なスクリーニングを組合せにより実施される。本節では、1 次スクリーニングにおける自動的なスクリーニング部分における基準を検討するための試行実験の内容について述べる。

3.3.1 実験の目的

1 次スクリーニングの自動的なスクリーニングの基準を検討するデータを得ることを目的とするが、本実験では以下の2つの視点について評価を行う。

- ・ 量の評価：1 次スクリーニングの通過件数から、1 次スクリーニングにおける機械的基準（件数、シグナル検出指標値）の知見を得る。
- ・ 内容の評価：シグナル検出計算に用いる手法を、ROR, BCPNN, GPS の3手法から絞ることの可能性に関する知見を得る。

3.3.2 実験方法

1 次スクリーニングおよび試行実験の枠組みを図 3-2 に示す。1 次スクリーニングはまずは人手により行われ、ここで抽出されなかった組合せについては定量的指標に基づき自動的なスクリーニングが行われる。

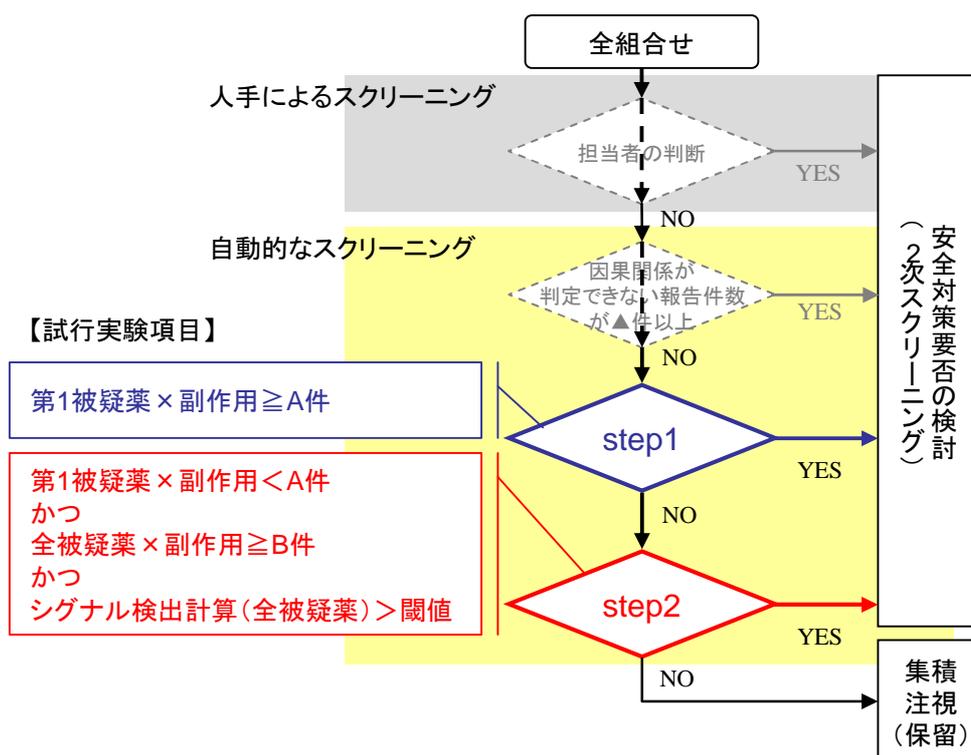


図 3-2 1次スクリーニングおよび試行実験の枠組み

自動的なスクリーニングの段階としては、次の3段階を想定している。

- ① 因果関係が判定できない報告件数が一定件数以上の組合せ
- ② ①に該当しない組合せのうち、第1被疑薬と副作用の組合せでの報告件数が一定数（A件）以上の組合せ（step1）
- ③ ②に該当しない組合せのうち、全被疑薬と副作用の組合せでの報告件数が一定数（B件）以上の組合せであり、かつ、シグナル検出指標値が閾値を超える組合せ（step2）

上記3段階のうち、本試行実験では②と③、つまり step1 と step2 について、スクリーニング基準値を変化させ、1次スクリーニングを通過する組合せの量と内容の評価を行う。表 3-5 に試行実験の条件を示す。なお、本来このフローは未知症例に適用されるものであるが、現時点では未知既知データベースが整備されておらず、対象を未知症例のみとするのは困難であるため、未知既知の区別はせず、評価対象期間中の新規の副作用報告全てを対象とする。

表 3-5 試行実験の条件

評価対象期間	2007年11月17日～12月21日の5週間に報告された副作用症例 (既知症例も含む)
評価対象組合せ	対象期間中に報告された全症例に含まれる被疑薬と副作用の組合せ
評価間隔	1週間ごと(計5回)
スクリーニング 基準	図 3-2 の step1、step2 のパラメータとして以下の値を設定 <ul style="list-style-type: none"> ● A、B (組合せ件数) の値：3件、4件、5件 ● シグナル検出指標値の閾値 ROR (95%信頼下限) > 1、 BCPNN (95%確率区間の下限) > 0、 GPS EB05 (5%点) \geq 2

3.3.3 実験結果

実験の結果のうち、一部抜粋したもの(2007/11/16-22に報告された副作用症例が対象)の結果を図 3-3 に、評価対象期間である5週分の全ての結果を表 3-6 に示す。

これらの結果から、量および内容の評価の観点について次の知見を得た。

- 量の評価：
A、Bの基準値は3～5の範囲でずらしても数十組程度変動するだけで、劇的な変化は見られない。
- 内容の評価：
RORは他2つのシグナル計算結果を包含している。このことから、使用するシグナルを絞る可能性が示唆される。

2007/11/16-22				
この週に報告された全症例数: 1333				
基準	A	3	4	5
	B	5	5	5
step1	この週来た組み合わせ数	999	999	999
	第1被疑薬A件以上	620	561	513
	第1被疑薬A件未満 かつ 全被疑薬B件以上	10	26	54
step2	// かつRORの閾値をクリア	3	14	32
	// かつBCPNNの閾値をクリア	3	14	31
	// かつGPSの閾値をクリア	1	5	15
	// かつROR/BCPNN/GPSのいずれかの閾値をクリア	3	14	32

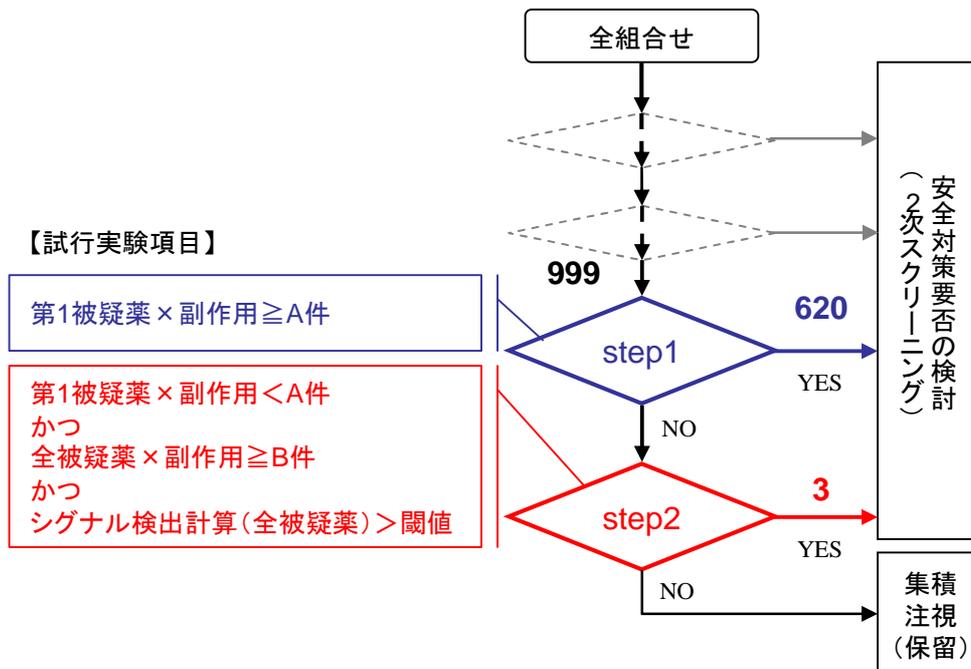


図 3-3 1次スクリーニング実験結果 (2007/11/16~22)

表 3-6 1次スクリーニング試行実験結果

2007/11/16-22 この週の症例数:1333 (うち、15日報告: 650)											
基準	A	3	3	3	4	4	4	5	5	5	
	B	3	4	5	3	4	5	3	4	5	
step1	この週末の組み合わせ数	999	999	999	999	999	999	999	999	999	
	第1被疑薬A件以上	620	620	620	561	561	561	513	513	513	
	第1被疑薬A件未満 かつ 全被疑薬B件以上	43	21	10	102	47	26	150	95	54	
step2	// かつRORの閾値をクリア	19	10	3	51	26	14	81	56	32	
	// かつBCPNNの閾値をクリア	16	10	3	45	25	14	71	51	31	
	// かつGPSの閾値をクリア	7	4	1	22	10	5	38	26	15	
	// かつROR/BCPNN/GPSのいずれかの閾値をクリア	19	10	3	51	26	14	81	56	32	

2007/11/23-29 この週の症例数:1019 (うち、15日報告: 489)											
基準	A	3	3	3	4	4	4	5	5	5	
	B	3	4	5	3	4	5	3	4	5	
step1	この週末の組み合わせ数	780	780	780	780	780	780	780	780	780	
	第1被疑薬A件以上	505	505	505	450	450	450	406	406	406	
	第1被疑薬A件未満 かつ 全被疑薬B件以上	42	21	11	97	46	27	141	90	51	
step2	// かつRORの閾値をクリア	25	9	4	61	23	14	87	49	26	
	// かつBCPNNの閾値をクリア	14	5	3	44	17	12	68	41	24	
	// かつGPSの閾値をクリア	6	2	1	25	8	7	41	24	16	
	// かつROR/BCPNN/GPSのいずれかの閾値をクリア	25	9	4	61	23	14	87	49	26	

2007/11/30-12/6 この週の症例数:1226 (うち、15日報告: 557)											
基準	A	3	3	3	4	4	4	5	5	5	
	B	3	4	5	3	4	5	3	4	5	
step1	この週末の組み合わせ数	960	960	960	960	960	960	960	960	960	
	第1被疑薬A件以上	588	588	588	527	527	527	483	483	483	
	第1被疑薬A件未満 かつ 全被疑薬B件以上	62	33	19	123	53	33	167	97	58	
step2	// かつRORの閾値をクリア	34	16	9	69	28	18	94	53	34	
	// かつBCPNNの閾値をクリア	30	14	8	56	26	17	81	51	33	
	// かつGPSの閾値をクリア	21	7	6	33	13	11	50	30	22	
	// かつROR/BCPNN/GPSのいずれかの閾値をクリア	34	16	9	69	28	18	94	53	34	

2007/12/7-13 この週の症例数:1257 (うち、15日報告: 559)											
基準	A	3	3	3	4	4	4	5	5	5	
	B	3	4	5	3	4	5	3	4	5	
step1	この週末の組み合わせ数	963	963	963	963	963	963	963	963	963	
	第1被疑薬A件以上	614	614	614	550	550	550	514	514	514	
	第1被疑薬A件未満 かつ 全被疑薬B件以上	64	35	24	128	63	42	164	99	58	
step2	// かつRORの閾値をクリア	26	15	9	59	27	17	81	49	25	
	// かつBCPNNの閾値をクリア	23	14	8	51	23	13	72	44	20	
	// かつGPSの閾値をクリア	4	2	2	18	7	5	29	18	7	
	// かつROR/BCPNN/GPSのいずれかの閾値をクリア	26	15	9	59	27	17	81	49	25	

2007/12/14-20 この週の症例数:1408 (うち、15日報告: 667)											
基準	A	3	3	3	4	4	4	5	5	5	
	B	3	4	5	3	4	5	3	4	5	
step1	この週末の組み合わせ数	1119	1119	1119	1119	1119	1119	1119	1119	1119	
	第1被疑薬A件以上	718	718	718	627	627	627	570	570	570	
	第1被疑薬A件未満 かつ 全被疑薬B件以上	64	36	26	155	80	50	212	137	84	
step2	// かつRORの閾値をクリア	37	20	15	90	42	29	121	73	49	
	// かつBCPNNの閾値をクリア	29	18	14	72	38	28	100	66	47	
	// かつGPSの閾値をクリア	13	10	6	33	19	13	49	35	24	
	// かつROR/BCPNN/GPSのいずれかの閾値をクリア	37	20	15	90	42	29	121	73	49	

3.4 今後の課題

新業務フローの策定するために、今後検討すべき課題を以下にまとめる。

(1) 1次スクリーニングにおけるシグナル検出手法の導入方法

抽出条件としてシグナル検出手法による計算結果を用いることとしているが、導入に向けて次の2点を決定しなくてはならない。

- 適用するシグナル検出手法の決定

昨年度の検討で導入対象となった ROR、BCPNN、GPS からさらに手法を絞る、もしくは複数を組合せて利用する、等を検討する必要がある。

- スクリーニング基準の決定

スクリーニングの基準となるシグナル検出指標値、累積報告件数の閾値を決定する必要がある。また、複数手法を用いる場合は、各手法での評価結果の組合せ方（AND、OR、多数決等）を検討しなくてはならない。

これらの項目を決定するためには、実データを用いた内容と量の検証をしなくてはならない。内容の評価では、手法ごとに抽出される医薬品と副作用の組合せの傾向について確認する。これにより、ROR、BCPNN、GPS のいずれを、どのように利用すべきかを検討することができる。量の評価では、閾値を複数設定した上で、実データを業務サイクルである1週間分用意し、これに対して閾値ごとに抽出される医薬品と副作用の組合せ件数調べる。この件数と現状の業務量とを比較することで、業務量という観点からの妥当な閾値を決定することができる。

(2) 既知症例に関する業務フローの検討

本年度の検討では未知の副作用報告を対象とした業務フローについて議論してきた。今後は、既知の副作用報告に対しての業務フローについても検討を進める必要がある。シグナル検出計算結果についても活用方法を検討進めなくてはならない。

(3) 新業務フローの試行の実施

新業務フローの検討がある程度進んだ段階で試行を行うべきである。試行の際にはフローが全て確定している必要はない。試行結果をフィードバックして再度検討する、このプロセスを重ねながらフローの細部を決定し、実用に耐えるものとするのが重要である。

また、一度業務フローが決定しても、業務を回す中で継続的に課題を抽出し、見直していかなくてはならない。

4. シグナル検出手法の高度化検討

4.1 非定型業務でのデータマイニング利用

これまでデータマイニングの導入を検討してきた業務は主に定型業務であった。ここで、定型業務とは、新たに到着した症例報告について、それに含まれる医薬品と副作用への対応の必要性の有無を判断するという、一定サイクルで行われる業務である。

一方、定型業務を行っていく中で、特定の症例の到着が直接的には関与しないが、医薬品（群）、あるいは副作用（群）について、報告されている内容の把握と分析を行う場合がある。このような業務を定型業務に対して、非定型業務とよぶことにする。

この非定型業務においてデータマイニングを利用する方法について、データマイニングの高度化、および将来導入するデータマイニングの検討の一部と位置づけ、検討に着手した。本章ではその経過について整理する。

4.2 Ad Hoc なシグナル検出指標値の利用

実際に報告される副作用症例報告は、特定の医薬品や特定の副作用の報告数が蓄積される傾向がある。したがって、定型業務で対象とする全医薬品と全副作用の組合せから算出されるシグナル検出指標値だけを参照していると、報告数の多い組合せ以外の組合せに関する指標値の変化が顕著に現れにくく、分析者が判断する上で有用な参照データとできないことがある。

そこで、シグナル検出指標値の変化をより明確にとらえるために、特定の薬効群に限定して副作用報告データを切り出したデータベースを作成し、このデータベースを用いてシグナル検出指標値を算出する妥当性について検討を行う。

4.2.1 検討方法

本検討では医薬品 A における血球系の副作用に注目する。検討対象とする医薬品、副作用、および検討に用いたシグナル検出手法を表 4-1 に示す。対象医薬品は、注目する医薬品 A に加え、医薬品 A の薬効群に含まれる医薬品全てであり、対象とする副作用は血球系の副作用である 6 種の PT（好中球数減少、白血球減少症、白血球数減少、汎血球減少症、無顆粒球症、顆粒球減少症）である。また、シグナル検出手法としては ROR を用いる。

表 4-1 Ad Hoc 検討の概要

検討対象	医薬品	当該薬剤（医薬品 A）の含まれる薬効群
	副作用	注目している血球系の副作用群 <ul style="list-style-type: none"> • 好中球数減少 • 白血球減少症 • 白血球数減少 • 汎血球減少症 • 無顆粒球症 • 顆粒球減少症
シグナル検出手法		ROR (Lareb での検出基準を採用し、ROR 95%下限値 ≥ 1 をシグナル検出閾値とする)
N ₁₁ とする医薬品の分類		次の 2 つの場合を設定し、それぞれについて報告件数およびシグナル検出指標値を算出 <ol style="list-style-type: none"> ① 第 1 被疑薬として報告されている症例 ② 被疑薬のいずれかとして報告されている症例
度数カウント方法		症例ベース

4.2.2 検討結果

検討結果のうち N₁₁ とする医薬品の分類が「② 被疑薬のいずれかとして報告されている症例」の場合の結果を表 4-2 に示す。表における全体 DB とは、特定の医薬品群や副作用群を設定せずに、全医薬品と全被疑薬から作成されるデータ群であり、定型業務に用いられるデータベースのことを指す。一方、部分 DB とは表 4-1 で検討対象とした医薬品が含まれる症例のみから作成されるデータ群を指す。つまり、部分 DB として「① 第 1 被疑薬として報告されている症例」により構成される部分 DB、および「② 被疑薬のいずれかとして報告されている症例」により構成される部分 DB の二種があり、表 4-2 に示す結果では後者の DB を用いている。したがって、全体 DB で作成される 2 重クロス表と部分 DB で作成される 2 重クロス表では、N₁₁ の度数は一致するが、その他の 3 つのセル (N₁₂、N₂₁、N₂₂) の値は異なる。

表 4-2 全体 DB と部分 DB での比較結果（全被疑薬）

副作用名	医薬品名	報告件数	全体 DB		部分 DB	
			ROR	ROR 95%信頼 下限	ROR	ROR 95%信頼 下限
医薬品 A	無顆粒球症	15	0.902	0.543	2.665	1.316
医薬品 B	無顆粒球症	4	0.435	0.163	0.762	0.266
医薬品 C	無顆粒球症	8	0.889	0.443	1.826	0.815
医薬品 D	無顆粒球症	2	0.189	0.047	0.308	0.074
医薬品 E	無顆粒球症	2	1.068	0.266	1.740	0.413
医薬品 A	顆粒球減少症	4	0.303	0.113	2.839	0.710
医薬品 B	顆粒球減少症	1	0.137	0.019	0.735	0.090
医薬品 C	顆粒球減少症	1	0.140	0.020	0.749	0.092
医薬品 D	顆粒球減少症	2	0.239	0.060	1.495	0.302
医薬品 A	白血球減少症	7	0.062	0.029	0.902	0.385
医薬品 B	白血球減少症	5	0.080	0.033	1.072	0.408
医薬品 C	白血球減少症	12	0.196	0.111	3.719	1.772
医薬品 D	白血球減少症	5	0.070	0.029	0.934	0.356
医薬品 A	好中球数減少	1	0.064	0.009	2.838	0.177
医薬品 E	好中球数減少	1	0.573	0.080	25.244	1.575
医薬品 A	汎血球減少症	19	0.556	0.354	1.457	0.836
医薬品 B	汎血球減少症	13	0.692	0.400	1.559	0.836
医薬品 C	汎血球減少症	18	0.980	0.615	2.497	1.421
医薬品 D	汎血球減少症	5	0.230	0.096	0.438	0.174
医薬品 E	汎血球減少症	1	0.259	0.036	0.456	0.063
医薬品 A	白血球数減少	9	0.376	0.195	25.597	3.241
医薬品 D	白血球数減少	1	0.066	0.009	0.498	0.063

表 4-2 より、全体 DB を対象としてシグナル検出計算を行った場合には、いずれの医薬品と副作用の組合せについてもシグナル検出閾値（ROR 95%下限値 ≥ 1 ）を超えてシグナル検出されるものは存在しない。これに対し、部分 DB を対象とした場合には、5 つの組合せ（医薬品 A と無顆粒球症、医薬品 C と白血球減少症、医薬品 E と好中球数減少、医薬品 C と汎血球減少症、医薬品 A と白血球数減少）でシグナル検出閾値を超える結果が得

られた。

以上の検討から、非定型業務として特定の医薬品群（副作用群）のみで作成される部分 DB を対象とする Ad Hoc な解析を行うことで、定型業務では発見できないシグナル検出指標値の変動が明らかとなり、部分 DB を用いたシグナル検出指標値が有効な参照情報となる可能性が示唆された。したがって、シグナル検出手法の高度化検討を行うに当たり、シグナル検出計算を部分的なデータで行い、粒度の違ったシグナル検出指標値が得られるシステムを構築することも有用であることが確認された。

4.3 関連する取組

4.3.1 海外規制当局における関連する取組

海外規制当局において非定型業務でデータマイニングを利用している事例として、オランダ・Lareb の事例とアメリカ・FDA の事例がある。

(1) オランダ・Lareb

オランダでは、定型業務においてアセッサーにシグナル検出指標値（ROR 値）が業務システム上に提示され、安全対策業務に利用されている。

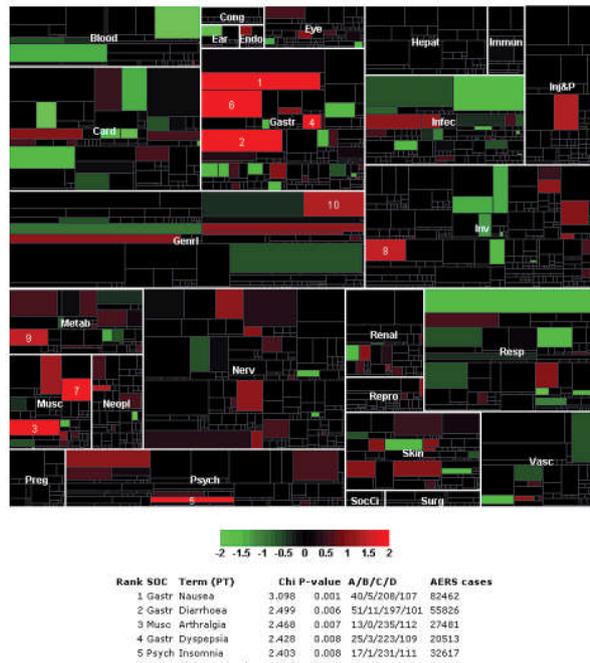
一方、これとは別にアナリストが使用しているシステムは、注目した医薬品群のみのデータセット（スモールセット）についてシグナル検出指標値を算出する機能を有している。このスモールセットによるシグナル検出計算機能は、アセッサー全員が使用しているシステムには搭載されておらず、アナリストが使用するシステムのみに搭載されている。アナリストは、非定型業務に相当する特定の医薬品群や副作用群に注目したスモールセットによる分析の依頼をアセッサーから受け、その結果を返すという業務形態をとっている。

もし、スモールセットでのシグナル検出指標値がデータベース全体での値に比べて大きい場合には、その医薬品群、副作用群は要注目であると判断される。

(2) アメリカ・FDA

アメリカ FDA では、アナリスト向けに特定の医薬品に関する分析をシステム上で支援する機能が導入されている。具体的には、図 4-1 のようなセクターマップという表示インタフェースを有している。このセクターマップは、MedDRA における SOC レベルでマップ上の各領域（セクター）を構成しており、マップ上の色はシグナル検出指標値(MGPS)のレベルに対応している。このマップを見ることにより、特定の医薬品の副作用の報告状況、あるいは医薬品間の副作用の報告状況の相違について把握することができる。

また、マップは操作インタフェースを兼ねており、マップ上の特定の領域を指定することにより、該当する副作用に関する症例やシグナル検出指標値の時系列的な変化などが表示されるなど、探索的な分析を支援する。



出典： https://phaseforward.com/resource/documents/CTSD_Data_Sheet072506.pdf

図 4-1 セクターマップ

このように、海外規制当局では、アナリストのような主に非定型業務を行う担当者に対して、データマイニング（シグナル検出計算）の指標値を提供することによる支援が行われている。

4.3.2 関連研究

特定の副作用に影響を与えている要因の分析に、データマイニングを利用している例の1つとして、デンマーク・University of Aarhus と アメリカ・University of California, Los Angeles のグループによる研究がある[9]。

この論文では、デンマークの大規模出生コホート研究で得られたデータを対象として、計算機科学分野の代表的なデータマイニング手法の1つである相関ルール分析手法とその指標である支持度、確信度を用いて要因分析を行った事例が紹介されている。

(1) 相関ルール

相関ルールとは、事例中で頻出する（共起する）要因の組を相関(Association)があるとして If-Then 形式で表現したものであり、X、Y を要因組として、If X Then Y、あるいは $X \rightarrow Y$ のように表現される。X、Y の出現頻度が表 4-3 のような場合、支持度、確信度は次のように定義される。

- 支持度： n_{11}/n_{++}
- 確信度： n_{11}/n_{1+}

確信度は、要因 X が見られる報告のうち、要因 Y も見られる報告の割合を表す。

表 4-3 要因組 X、Y の出現頻度

	Y	Y以外	合計
X	n11	n12	n1+
X以外	n21	n22	n2+
合計	n+1	n+2	n++

(2) 分析例

適用事例として、抑うつ剤の服用と服用時期、および喫煙の有無の早産への影響の分析例が示されている。

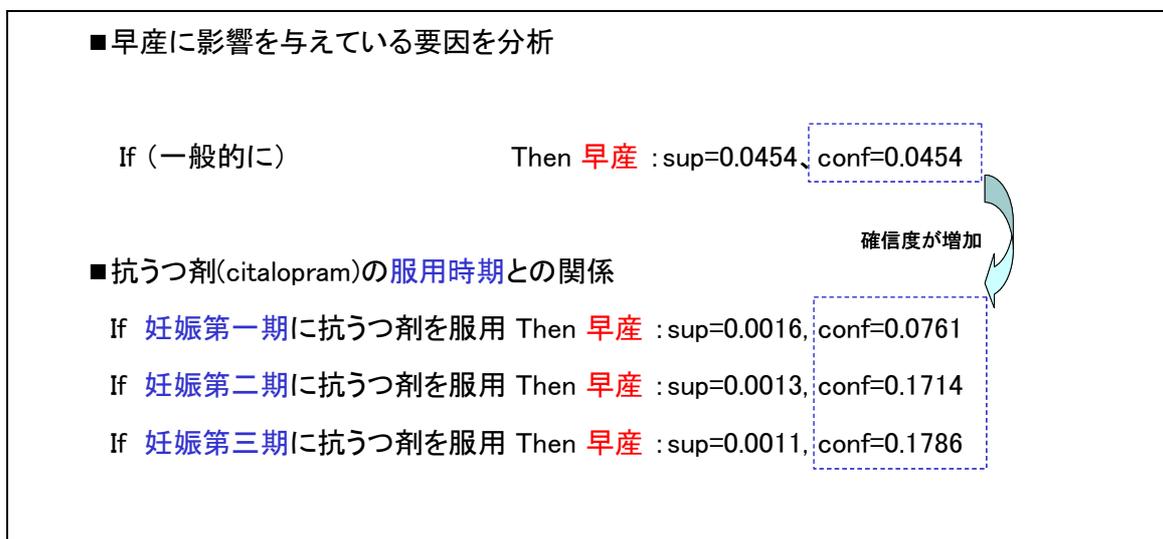


図 4-2 抗うつ剤の服用時期との関係

図 4-2 は、抗うつ剤の服用時期の要因について分析した例である。確信度の変化から、抗うつ剤を服用しない場合に比べて、早産となる割合が高くなることがわかるとしている。次に、図 4-3 はさらに喫煙の要因を考慮した場合である。この例でも、それぞれ確信度が増加しており、喫煙も要因として注目すべきであるとしている。

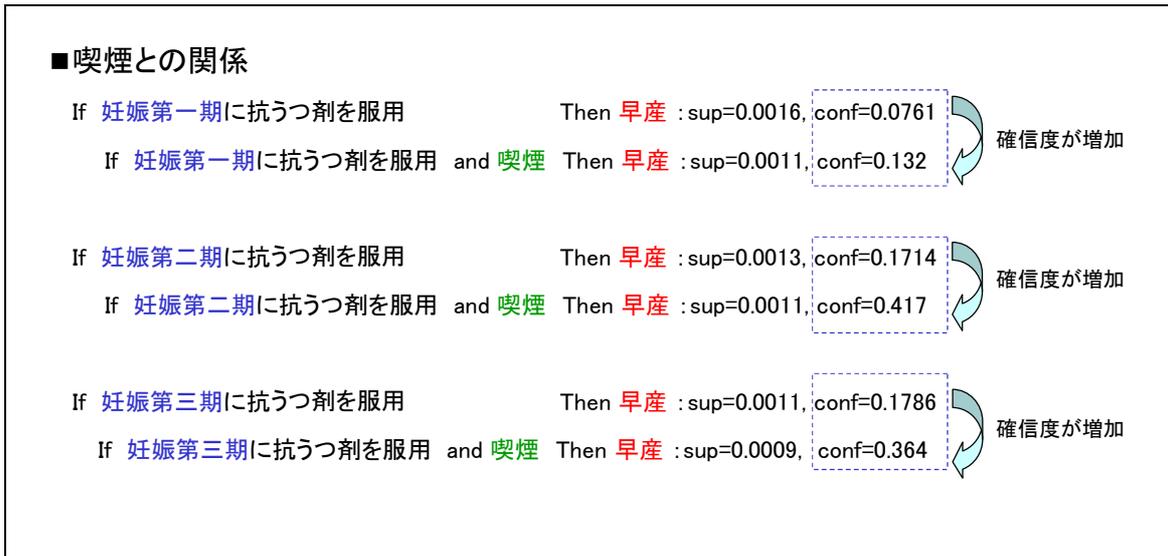


図 4-3 抗うつ剤服用時期に加えて喫煙を考慮した場合

この研究では、対象としているデータがコホート研究で得られたデータであるため、自発報告を対象とするデータマイニングでの有効性については検討が必要であるが、アナリストが分析を進める中で各種の統計的指標を提示することは、分析の支援として有効であると考えられる。

4.4 分析支援ツールのプロトタイプ

アナリストにデータに基づく各種指標値を提示し、分析を支援するツールのプロトタイプを準備した。分析支援ツールは大きく分けて 2 とおりの機能を持つ。

(1) リスト表示機能

指定された医薬品（群）、あるいは副作用（群）について、多く報告されている医薬品、副作用を頻度が高いものから順にリスト表示する機能である。注目している医薬品、副作用に関する報告の状況を把握することを支援する。

例えば、医薬品を指定し、頻度が高い副作用をリスト表示することにより、当該医薬品に関する副作用の報告状況がわかる。また、同様に頻度が高い医薬品をリスト表示することにより、併用されている医薬品の状況がわかる。

ツールのインターフェースイメージを図 4-4 に示す。

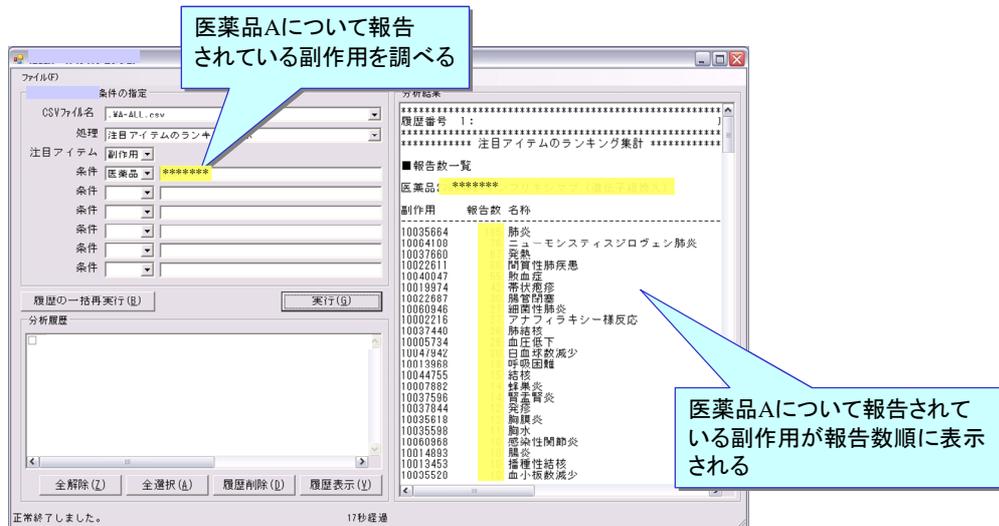


図 4-4 リスト表示機能

(2) 指標値算出機能

指定された医薬品（群）と副作用（群）の組について、各種の指標を表示し、因果関係の有無の判定を支援する機能である。

次の3種類の統計的指標を表示する。

- 頻度
表 4-3 のように 2×2 表形式で頻度を表示する。
- 割合
相関ルールの確信度、支持度に相当する割合を表示する。
- 統計的指標（シグナル検出指標値）
ROR 等のシグナル検出指標値を表示する。

ツールのインターフェイスイメージを図 4-5 に示す。アナリストは各種の指標を参照しながら、因果関係の有無について判定することができる。

また、このツールは、医薬品、副作用の集合をそれぞれ1つの仮想的な医薬品、副作用とみなしグルーピング機能を有している。

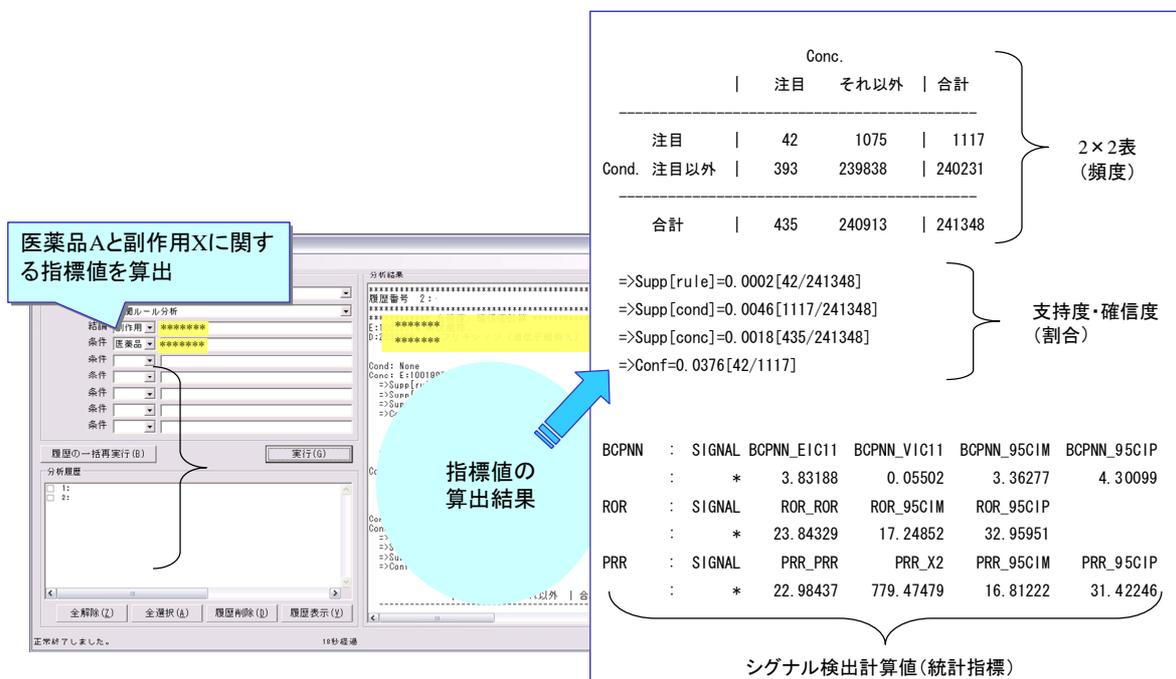


図 4-5 指標値算出機能

(3) ツールを利用した分析イメージ

分析支援ツールを利用した分析イメージを図 4-6 に示す。リスト表示機能を用いた俯瞰と深い分析対象の絞り込みと指標値算出機能を用いた注目している医薬品・副作用組の因果関係の判定を繰り返しながら、探索的な分析を効率的に進めることを想定している。

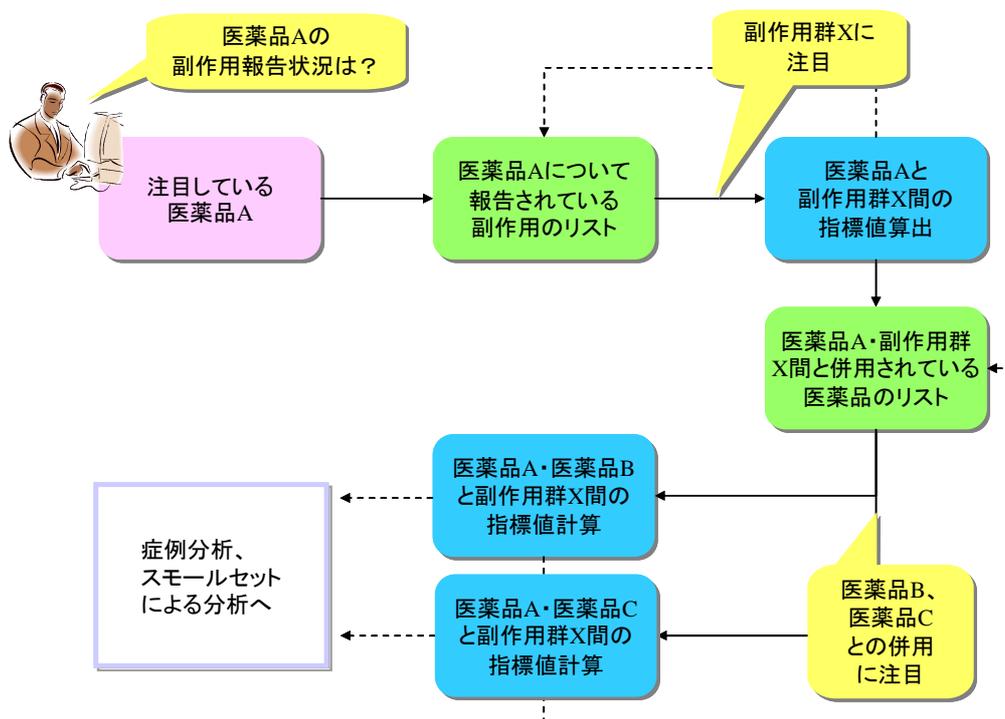


図 4-6 分析支援ツールを用いた分析イメージ

4.5 今後の課題

非定型業務におけるアナリストの分析について、データマイニングを用いて支援する方法として、シグナル検出指標値を含む各種の指標値の表示について検討を行った。その有効性について評価するため、指定された医薬品、または副作用の組について指標値を表示する機能を持つツールを開発し、試用した。各種の指標値を参考指標として提示することは有効であると考えられる。

今後の課題として、次の点が挙げられる。

(1) 基本機能の拡充

層別分析、スモールセットによる分析、MedDRA の各階層における分析等、各種分析のための集計や指標値計算ができるように基本機能を拡充する。

(2) 指標値の参照に関するガイドライン

現在は指定された内容に関する指標値を計算し、提示するのみであり、その参照の方法については示していない。アナリストに対して頻度、割合、統計的指標の使い分けや、判定するための数値の見方についてガイドラインを作成、提示することが望まれる。

(3) ナビゲーション機能の検討

探索的な分析を行う中で探索の方向付け、あるいは示唆を与える機能（ナビゲーション）機能の搭載が望まれる。

(4) 可視化インタフェースの搭載

アナリストの分析結果の把握を支援するために、可視化インタフェースの搭載が期待される。

5. 調査のまとめ

本年度の調査では、(1) データマイニングを用いた業務システムの仕様作成、(2) データマイニング導入後の新業務フローの検討、および(3) データマイニングの高度化に関する検討を実施した。

データマイニングを導入した業務システムの仕様を作成するにあたり、Lareb、EMEA 海外規制当局に対してヒアリングを行った結果、両者では、データマイニング（シグナル検出指標値）をいわゆる機械的なフィルタリングには用いず、参考値としてアセッサーに提示していることがわかった。このデータマイニングの位置づけは、PMDA において検討しているものと類似しており、大いに参考となった。

一方、Lareb では、アセッサーのための業務システムにシグナル検出指標値（ROR 値）のみでなく、各度数なども参考情報として提示しているということであった。これらの数値を正しく利用するためには、シグナル検出指標値の内容に関する理解が必要であるが、アセッサーはこれについて一定レベルの知識を有しているとのことであった。データマイニング値を機械的なフィルタリングではなく、参考値として用いるためには、参照するアセッサーのシグナル検出指標値に関する理解の向上が必要であり、PMDA においてもシグナル検出指標値への理解を深めるための取組が行われることが期待される。

Lareb の業務システムは、業務フローに沿ったものとなっており、アセッサーにとってユーザビリティの高いものであった。

これら海外規制当局へのヒアリングから得られた情報は、業務システムの仕様、および今後の開発に反映される。

データマイニングを導入した安全対策業務フロー（新業務フロー）に関して、フロー中の判定基準の検討と新業務フローにおける業務量評価の試行を行った。今回の評価では顕著ではなかったが、各種の判定基準を変更することでスクリーニング結果等が変化し、業務量は変化する。また、今後、副作用報告の年間報告数や累積報告数が増加するなどデータベースの状況が変化することにより、同じ判定基準であってもスクリーニング結果が質、量ともに変化し、その結果、業務量が変化する。よって、新業務フローの導入までに、新しい業務フローに沿った試行と量、内容の両面からの判定基準の評価を行うことが期待される。また、新業務フローの導入後も定期的に判定基準を見直し、スクリーニング結果等を適切なものとする必要がある。

また、新業務フローの導入に向けて、未知・既知データ等のデータベースの整備が進められることが望まれる。

さらに、データマイニング高度化検討、および将来導入するデータマイニング検討の一

環として、海外の規制当局におけるアナリスト相当の役割である担当者による非定型な分析を支援するデータマイニングの検討を行った。

今年度は、初期の取組として注目している医薬品、副作用と多くの症例で報告されている他の医薬品や副作用を多いものから表示するリスト表示機能と、医薬品群と副作用群の組み合わせについて、度数、割合、シグナル検出指標値等の指標値を提示する指標値計算機能を実装したツールを開発し、試用した。

今後は、スモールセット、層別、PT以外の MedDRA 階層など、他の分析での指標値計算ができるように機能を拡充しながら、有望な分析の切り口を提示するなど積極的な分析支援の方法について検討していくことが期待される。

また、相互作用分析など、その他の高度化検討、将来導入するデータマイニングの検討についても検討を進めていくことが望まれる。

参考文献

- [1] 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, 三菱総合研究所, 平成 17 年
- [2] 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, 三菱総合研究所, 平成 18 年
- [3] データマイニング手法の導入に関する検討結果報告書, 独立行政法人 医薬品医療機器総合機構, 平成 19 年
- [4] William Dumouchel, “A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions”, *Pharmacoepidemiology and Drug Safety*, Vol.11, pp. 3-10, (2002)
- [5] The European Medicines Agency in 2006
- [6] Guideline on the use of statistical signal detection methods in the EudraVigilance Data Analysis System, EMEA, 2007.10.11
- [7] Ekins-Daukes, S; Kauser, S; Wise, L, Signal Detection in the UK: The Use of Quantitative Methods at the MHRA, *Drug Safety*, Vol.30, No.10, pp. 963(2007).
- [8] W. Patrick, H. Emma, M. Jane, Impact Analysis of Signals Detected from Spontaneous Adverse Drug Reaction Reporting Data, *Drug Safety*, Vol. 28, No.10, pp. 843-850(8), (2005).
- [9] Yu Chen, Lars Henning, Pedersen Wesley, W. Chu, Jorn Olsen, “Drug Exposure Side Effects from Mining Pregnancy Data”, *ACM SIGKDD Explorations Newsletter*, Vol.9 , No.1, pp. 22-29 (2007)