

ベネフィット・リスク評価  
情報の不確かさを考慮した  
市販後のベネフィット・リスク再評価  
—PBRER のフレームワークとしての応用—

日本製薬工業協会  
データサイエンス部会  
タスクフォース 1

Ver 1.0

2014年3月

## 目次

<b>EXECUTIVE SUMMARY</b> .....	<b>4</b>
<b>1 評価の準備</b> .....	<b>5</b>
1.1 承認時のベネフィット・リスク評価と市販後のベネフィット・リスク再評価 .....	6
1.2 データソースのエビデンスレベル .....	7
1.3 ベネフィット・リスク再評価が必要な状況 .....	9
1.4 承認申請時に残るデータの不確かさ .....	11
<b>2 PBRER をフレームワークとして利用する</b> .....	<b>13</b>
2.1 背景の定義 .....	15
2.2 ベネフィットとリスクの項目の列挙 .....	16
2.3 データソースの特定 .....	18
2.4 項目の優先順位付け, 重み付け .....	20
2.5 結果のまとめ .....	21
<b>3 データの不確かさへの対処</b> .....	<b>23</b>
3.1 必要な情報に関して症例数が少ない .....	24
3.2 対照群がない .....	25
3.3 交絡因子が存在する .....	26
3.4 バイアスがある .....	26
3.5 そもそも情報がない .....	28
<b>4 おわりに</b> .....	<b>29</b>
<b>5 別添：用語解説</b> .....	<b>31</b>
<b>6 REFERENCES</b> .....	<b>35</b>

## はじめに

ベネフィット・リスク評価は決して新しい話題ではない。ベネフィット・リスク評価についての議論は、21 世紀に入りさらに欧米の規制当局、製薬企業および各研究グループで活発に行われているが、興味深いことに、15 年も前に公開された CIOMS のガイダンス[1]は、市販後のベネフィット・リスク再評価を対象としたものであった。

これまで、われわれ日本製薬工業協会データサイエンス部会のタスクフォース 1 では、『ベネフィット・リスク評価入門』（2013 年 3 月公開）[2]、『ベネフィット・リスク評価 中級編 多基準決定分析への招待—その理論と事例—』（同 5 月公開）[3]、『ベネフィット・リスク評価 中級編 定量的手法に関する考察』（同 10 月公開）[4]と、ベネフィット・リスク評価に関する日本語の報告書を発表する共に、パイロット調査[5]やワークショップ[6]、説明会を開催するなど活動を続けてきたが、市販後のベネフィット・リスク「再評価」に関する関心とニーズが高いことを肌で感じてきた。Periodic benefit-risk evaluation report，すなわち PBREER の提出も始まったが、市販後のベネフィット・リスク再評価の基本的な考え方や方法論についての分かりやすい解説は、これまでに本邦では出版されていない。

そこで、市販後のベネフィット・リスク再評価について、本報告書をまとめることにした。承認申請時のベネフィット・リスク評価と市販後の再評価とでは何が異なるのであろうか？ベネフィット・リスク評価はさまざまな専門知識を持った担当者からなるチームワークであることから、本報告書は臨床開発、薬事、生物統計、疫学、安全性担当者など、幅広い読者層を対象としている。ベネフィット・リスク評価を行うに当たり、理論的背景や限界点を理解した上で、適切に応用したいと考える読者の参考になれば幸いである。

### **Executive summary**

- 市販後だけに応用可能な特別なベネフィット・リスク評価法が存在するわけではない
- ベネフィット・リスク再評価は PBRER 作成のために実施するのではない。PBRER を単なる報告フォーマットとしてではなく、ベネフィット・リスク再評価のフレームワークとして積極的に活用してみよう
- まずは承認申請時のエビデンスレベルの高いデータソースをもとに、ベネフィット・リスク評価をおこない、再評価が必要になる状況に備えて準備をしておく。一方で、承認申請時点での情報不足やデータの不確かさについて明らかにしておく
- エビデンスレベルの異なるさまざまなデータソースについてよく理解し、それらから得られるデータを正しく評価しよう
- エビデンスレベルの低いデータソースも、他の情報と組み合わせたり、その不確かさの原因を吟味することで、それを活用する糸口がつかめることがある。ただし、そうするためには、相応の経験と知識が必要である。
- 特に定量的手法を応用する場合には、数値の増減に一喜一憂せず、その奥にある臨床的価値をよく考えよう

## 1 評価の準備

本章のポイント：

- 市販後だけに応用可能な特別なベネフィット・リスク評価法が存在するわけではない
- まずは承認申請時のデータでベネフィット・リスク評価をおこなう
- 再評価が必要になる状況を考えて準備しておく
- 承認申請時点でのデータの不足や不確かさを確認する
- データのエビデンスレベルを考える

ベネフィット・リスク評価の実際について、日本製薬工業協会データサイエンス部会のタスクフォース1 (TF1) では、これまで『ベネフィット・リスク評価入門』、『ベネフィット・リスク評価 中級編 多基準決定分析への招待—その理論と事例—』、および『ベネフィット・リスク評価 中級編 定量的手法に関する考察』と3本の報告書を公開してきた[2-4]。本報告書では市販後のベネフィット・リスク再評価に焦点をあてているが、再評価に際し、あるいは再評価をはじめめる以前の問題として、**承認申請時のベネフィット・リスク評価はすでに終わっていることが前提**である。えっ、そうなの？と思われた方は、まず『ベネフィット・リスク評価入門』をご覧ください、まずは承認申請時の情報（臨床試験の情報）を使ってベネフィット・リスク評価を行ってみたい。なお、**太字\***で示した用語は別添に解説を付けたのでご参照いただきたい。

また、市販後の再評価の特徴として、安全性に関する重要な情報のアップデートがあった場合に、意思決定までの時間が十分に取れないことが予測される。だからこそ、事前の十分な準備が必要なのである。情報の不確かさを加味した上で、可能な限り速やかな意思決定をサポートするためにベネフィット・リスクの再評価を行うとすれば、どのような工夫が可能であろうか？

TF1は、以下の3点を提案する。

- 個々のデータソースに対するエビデンスレベルについて整理しておく
- PBRER を応用したフレームワークを利用する
- 情報の不確かさを明確にし、その種類に応じた対応を織り込むようにする

エビデンスレベルについては1.2でさらに解説するが、とりあえずは、得られた医学情報の科学的根拠（エビデンス）の強さ、あるいは確からしさを適切に評価したうえでレベル分けしたもの、と考えておこう。

## 1.1 承認時のベネフィット・リスク評価と市販後のベネフィット・リスク再評価

これまでの報告書でも繰り返し述べた通り、ベネフィット・リスク評価は意思決定をサポートするために行うものであり、その本質は定性的な評価にある。これは承認申請時のベネフィット・リスク評価でも、市販後のベネフィット・リスク再評価でも同じである。市販後だけに応用可能な特別なベネフィット・リスク評価法が存在するわけではなく、承認申請時でも市販後でもベネフィット・リスク評価の基本的なコンセプトや方法は同じである。この前提に立った上で、承認申請時の評価と市販後の再評価との違いを考えてみると表 1-1 のようになる。

表 1-1 申請時と市販後のベネフィット・リスク評価の違い

承認申請時	市販後
<ul style="list-style-type: none"><li>エビデンスレベルの高い臨床試験が主なデータソースとなる</li><li>データ収集や解析に十分に時間をかけられることが多い</li></ul>	<ul style="list-style-type: none"><li>メタアナリシスの結果やその他臨床研究の研究報告、自発報告、企業がスポンサーとなる調査など、データソースのエビデンスレベルにばらつきがある</li><li>意思決定までの時間が限られる場合がある</li></ul>

すなわち、承認時はエビデンスレベルの高い臨床試験のデータを主に対象としていたのに対し、市販後では論文等で報告される臨床研究の結果や、自発報告、企業がスポンサーとなる調査などから、さまざまな情報が五月雨式にもたらされる（図 1-1）。その中には必ずしもエビデンスレベル（信頼性）の低いデータも含まれていることだろう。

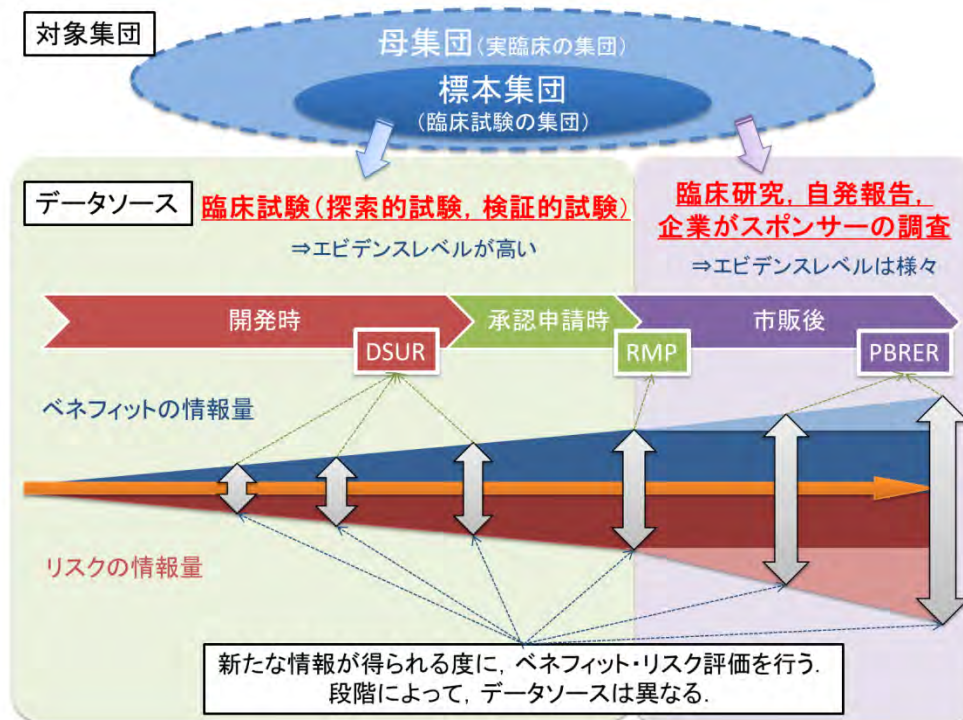


図 1-1 医薬品のライフサイクルとベネフィット・リスク評価

疾患集団から得られるデータの情報源（データソース）およびそのエビデンスレベル，開発が進むとともに，ベネフィットやリスクに関する情報量が増していくことを示している．情報量について，水色は市販後に新たに得られるベネフィットの情報，ピンクは市販後に新たに得られるリスクの情報を示す．

## 1.2 データソースのエビデンスレベル

エビデンスレベルについては，表 1-2 に米国医療政策研究局（現在の医療研究品質庁：AHRQ）による分類を参考までに示した[7]．

表 1-2 米国医療政策研究局（現在の医療研究品質庁）によるエビデンスの分類

エビデンスレベル	エビデンスの種類
Ia	複数のランダム化比較試験のメタアナリシス
Ib	少なくとも一つのランダム化比較試験
IIa	少なくとも一つがよくデザインされた非ランダム化比較試験
IIb	少なくとも一つ他のタイプがよくデザインされた準実験的研究*
III	比較試験，相関研究，症例対象研究など，よくデザインされた非実験的記述的研究．
IV	専門家委員会の報告や意見，権威者の臨床経験

米国医療政策研究局：Agency for Health Care Policy and Research,  
 医療研究品質庁：Agency for Healthcare Research and Quality (AHRQ)  
 \*準実験的研究：対照群またはランダム化のいずれかを欠く実験的研究．

先に「承認申請時のベネフィット・リスク評価はすでに終わっていることが前提」と述べたが、これは実際の評価実施時期を問題にしているのではなく、承認申請時までに得られる可能性のあるデータソースは信頼性が高いと考えられるため、この時まで利用可能な情報に基づいてまずはベネフィット・リスク評価を行った方が、いきなり市販後のエビデンスレベルのばらつくデータを用いてベネフィット・リスク再評価を行うよりハードルが低いだらうと考えてのことである。

すなわち、承認申請時に用いられる主要なデータはランダム化比較試験（表 1-2 では最も高いエビデンスレベルとして示されている）から得られていることがほとんどであり、データソースの信頼性についてはそれほど心配せずに、データを解釈しベネフィット・リスク評価を行うことができる。このようにしてまずは評価の基本的な土台を作っておいてから、それを利用して再評価するとよい。というのは、こうすることで、市販前と後のベネフィット・リスク評価の一貫性・整合性が保てるうえ、市販後のベネフィット・リスク評価それ自体も、既に評価の基本的な土台があることで、はるかに容易に行うことができるからである。

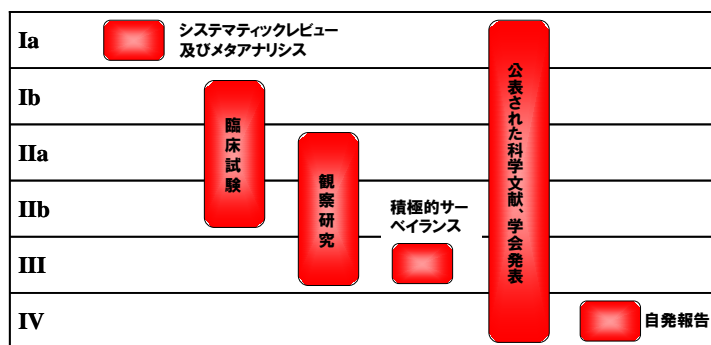


図 1-2 市販後のデータソースのエビデンスレベル

一方、市販後に利用可能と考えられる代表的なデータソースについて、それぞれのエビデンスレベルを考えてみると、開発段階と異なりデータソースそのものが多種多様である。このため、たとえ同じ研究デザインに分類される研究であったとしても、個々の調査研究の信頼性・妥当性を慎重に評価する必要がある（図 1-2）。

例えば、妊婦の薬剤への曝露とある先天性欠損を有する児の出産との関連を調査する症例対照研究が 2 つあるとしよう。一方の研究では妊娠期間中の薬剤の服薬状況を直接出産後の母親から面接で情報を得ており、他方の研究では母親の診療録から薬剤の処方状況を確認していたとする。研究デザインの観点からどちらの研究も同じ症例対照デザインでありエビデンスレベルは同じである。しかしながら、前者の研究では、先天性欠損を持った児を産んだ母親は、そうでない児を産んだ対照の母親より薬剤の曝露に関して詳細な記憶までより深く思い出そうとするために（想起バイアス：情報バイアス\*の一種）、症例・対照間で薬剤曝露情報の質に差が生じる懸念がある。このような場合には、薬剤曝露と先天性欠損の関連を正確に評価することは困難である。一方、後者の研究では、症例群、対照群



ともに同じレベルの曝露情報が得られるので、少なくとも人の記憶に依存する際に生じがちな問題点は回避できるであろう。このように、個々の研究に対する批判的な評価・吟味はその研究結果のエビデンスレベルを決定するということを忘れてはならない。

さらにデータに含まれる可能性のあるバイアス\*や不足している情報についても精査する必要がある。例えばAHRQは、エビデンスの強さを評価するのに重要な要素として、以下の3点を挙げている[8]。

- **質**：バイアスをどれほど小さくすることができたかという観点に基づいて個々の調査の質を評価したうえで、それらを総合したもの
- **量**：効果の強さ、調査の数、症例数または検出力
- **一貫性**：同等の、あるいは異なるデザインを用いた調査から同じような結果が報告される程度

すなわち、前述の症例対照研究の例であれば、先天性欠損の発生率が曝露の有無により大きく異なる場合、また調査症例数が多い場合には、それなりにエビデンスが強くなると考えられる（バイアスの影響が消えるわけではない）。また、同様の報告が多方面からなされていればさらにエビデンスは強くなる。しかし、ベネフィット・リスク再評価において目的とする対象集団や項目に関して十分な情報が受動的に得られることは稀である。

市販後には、不確実な情報を確かなものにする努力や、不足する情報を補う努力を続けていくことになる。その詳細は原則的にはリスク管理計画（RMP）に記載されることになるが、本報告書ではそのための具体的な方策に関しては対象外とする。あくまでもベネフィット・リスク評価に着目し、市販後の再評価に向けてどのような準備ができるか、またするべきか、十分に準備した上でどのように再評価を行うかについて述べる。

### 1.3 ベネフィット・リスク再評価が必要な状況

ベネフィット・リスク評価／再評価は、必要に応じて適宜行うものであり、事前に定めたスケジュールに従って義務的に行うものではない。例えば以下のような場合にベネフィット・リスク再評価が必要であり、実際これまでも（明示的かどうかは別として）行ってきたはずである（図 1-3）。

- ・ 未知で重篤な安全性懸念事項が発生した場合
- ・ 安全性懸念事項の頻度に変更があった場合
- ・ 安全性懸念事項のプロファイル（重篤性、予防可能性、可逆性等）に変更があった場合

- ・ 対象疾患を取り巻く状況が変化した場合（ガイドラインの変更で、診断基準が変更になる等）
- ・ 対象医薬品の相対的地位が変化した場合（ガイドラインの変更で治療パラダイムが変化した、新しい治療手段が登場した等）

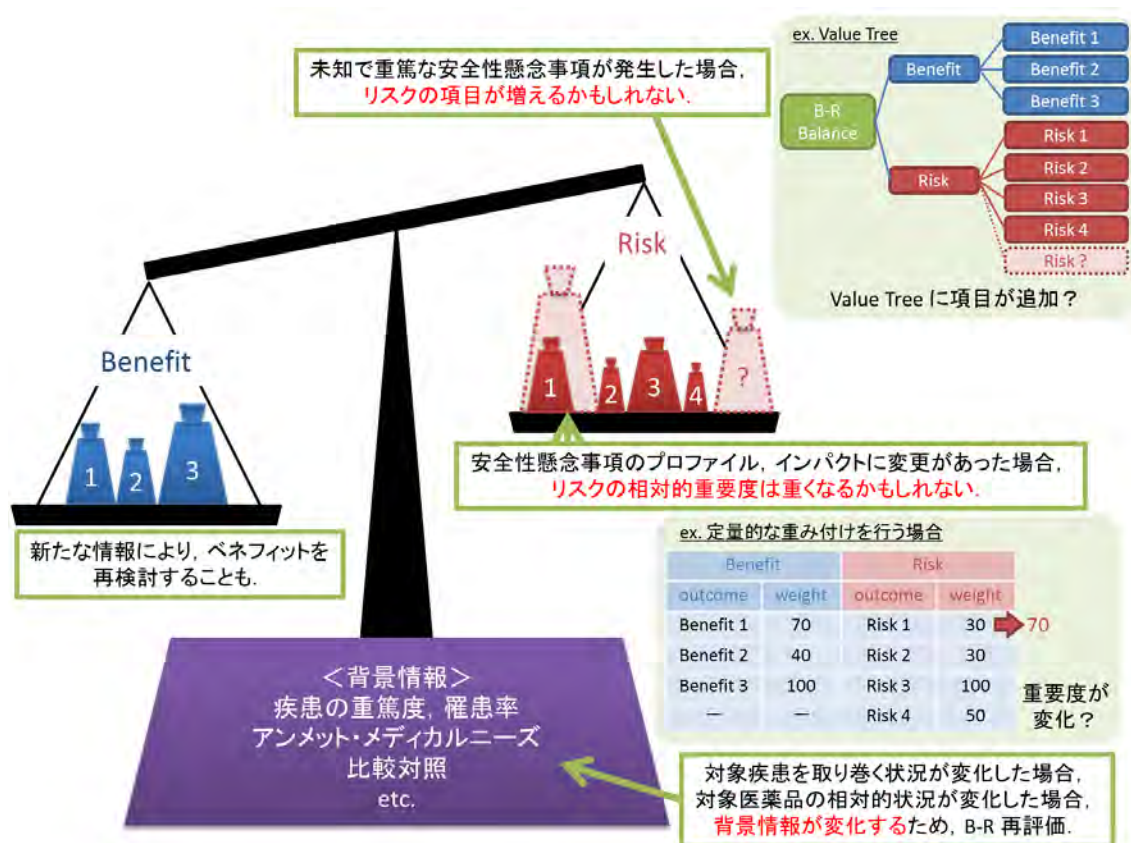


図 1-3 ベネフィット・リスク再評価が必要な場合とその対応の例

例えば、未知で重篤なリスクイベントが起きた場合には、その新たなリスクの項目をベネフィット・リスク評価に追加すべきかを考慮する必要がある。また、リスクイベントのプロファイルが変化した場合、例えば思ったより重症であったり、頻度が高かった場合には、リスクの相対的な重要度（重み）を変更する必要があるかもしれない。その重みの変更がベネフィット・リスク評価にどのような影響を与えるか、評価する必要があるかもしれない。さらに対象の薬剤自体ではなく、対象疾患を取り巻く状況、例えば新しい診断法や治療法の開発によって、診断ガイドラインや治療ガイドラインが改訂された場合には、対象医薬品の相対的な位置づけが変化する可能性がある。これにともない、ベネフィット・リスク評価の根底をなす背景情報そのものも変わるようなら、その新たな背景情報に基づいて、あらためてベネフィット・リスク評価を行う必要があるだろう。

よくある誤解の一つに、新たな適応が追加承認されるとベネフィットの項目に追加できるというものがあるが、新たな適応ということは対象集団（背景情報）が異なるということであり、対象集団が異なればベネフィット・リスク評価の最初のステップである背景のところから、あらためてベネフィット・リスク評価を行うことになるはずである。

#### 1.4 承認申請時に残るデータの不確かさ

承認申請時には、ベネフィットがリスクを上回ることを明確に述べる必要があり、臨床試験パッケージを通して計画的にエビデンスを構築している。特に有効性は承認申請までに検証しなければならず、また評価の妥当性を高めるため、エビデンスレベルの高い二重盲検のランダム化比較臨床試験を行うことが多い。それでもベネフィット・リスク評価を行うに当たり、不確実な情報や不足する情報はあつた。古くから指摘されている通り、臨床試験には一般的に：

- Too Few：集積症例が少ない
- Too Simple：合併症や併用薬が少ない
- Too Brief：曝露期間が短い
- Too Median-Aged：低年齢、高齢者が少ない
- Too Narrow：絞つた対象疾患

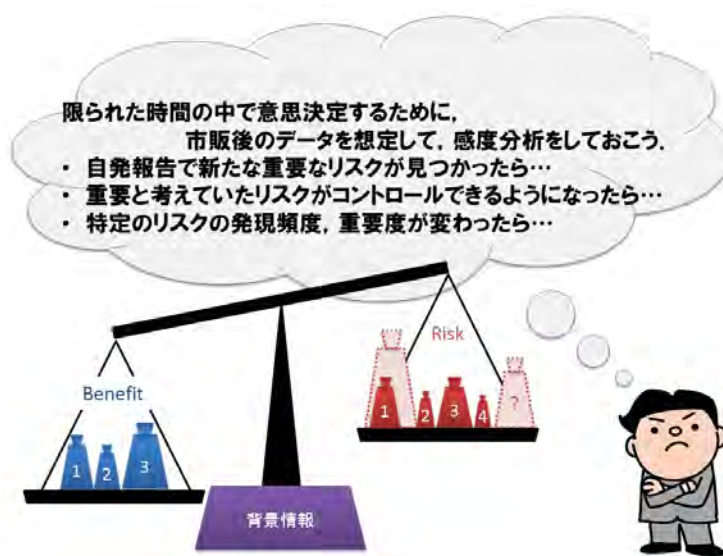
のような特徴があり、これらの不確実な情報を明らかにすることが重要であることは、先の報告書でも述べた通りである。

例えば、代替エンドポイントで承認を取得する場合、真のエンドポイントでの有効性にはある程度の不確かさが残る。臨床試験の被験者は患者集団全体からのランダムサンプリングではないため、試験で証明した医薬品の効果に関して**一般化可能性\***（外部妥当性）に不確かさが残る。また、症例数が限られていることから、往々にしてリスクイベントの評価指標に関する**精度\***には不確かさが残るし、サブグループでのベネフィットやリスクに関する検討についても同様である。慢性疾患の場合には、長期投与のエビデンスに不確かさが残るかもしれない。

このような情報の不確かさは、ベネフィット・リスク評価の結論に影響を与える可能性がある。そこで、現在の情報についてだけでなく、近い将来に得られるであろう市販後のデータを想定した感度分析が必要となる。例えば自発報告で新たな重要な有害事象が見つかった場合にベネフィット・リスク再評価したとしたら結論は変わるのだろうか？

重要さと発現頻度を仮に定めてみて、この程度の発現頻度であれば、重要な有害事象が見

つかったとしてもベネフィット・リスク評価の結論には影響を与えないだろうといった検討を事前におけば、受け身のまま新たな情報が得られるのを待つだけでなく、先手を打った市販後情報の収集が可能になる場合もあるだろう。他にも、重要と考えていた有害事象のコントロールが容易になって重要性が低下したらどうなるのか、一部の有害事象の発現割合が臨床試験の結果と異なっていたらどうなるのか、といった様々な疑問に対して事前に検討しておくことは、情報が更新された場合に限られた時間の中で意思決定する必要があるベネフィット・リスク再評価において特に有用である。



得られた情報によりタイムリーに対応するために、例えば MCDA (Multi-Criteria Decision Analysis : 多基準決定分析) のモデルを承認申請時に作成しておく方法もある。その有害事象がすでにモデルに含まれていたならば、新たな情報が得られるたびにその頻度やインパクトのパラメータを更新する。あるいは含まれていなかったが重要だと考えるのなら、新たにリスクの項目に追加して結果に影響を与えるかを見してみる。モデルを準備しておくのは大変だが、市販後の再評価が必要になった際に、意思決定までの時間を短縮したり、申請時の意思決定との整合性を保つことができるなどのメリットがある。

## 2 PBRER をフレームワークとして利用する

本章のポイント：

- PBRER を単なる報告フォーマットとしてではなく、フレームワークとして活用しよう
- ベネフィット・リスク再評価は、PBRER 作成のために実施するのではないことを肝に銘じよう
- 各データソースのエビデンスレベルを考えてデータを解釈しよう
- リスクの情報が増えるからと言って、天秤がリスクの方に傾くわけではない。情報の精度を考えよう

ベネフィット・リスク評価にはフレームワークを利用するのが便利である[2]。すなわち、①背景や意思決定の目的を整理し、②ベネフィットとリスクの項目を特定し、③各ステークホルダーの定性的な相対的重要性の評価を加味し、④利用可能なデータソースのエビデンスレベルを精査した上で必要なデータを抽出して、⑤全体的な評価を試みる。こうした基本的なステップをたどることで、ベネフィット・リスク評価に必要な材料を揃え、各ステークホルダー間で考え方やその道筋をある程度そろえることができる。これはベネフィット・リスク再評価でも同じである。

定期的ベネフィット・リスク評価報告（Periodic Benefit-Risk Evaluation Report, PBRER）は、「製品の全体的なベネフィット・リスク評価を可能にするため、医薬品のリスク及び承認適応に対するベネフィットに関する新しい情報又は明らかになりつつある情報の包括的、簡潔かつ重要な分析を示す」ものであり、医薬品のベネフィット・リスク評価を定期的に報告するための共通の基準である[9]。以前は Periodic Safety Update Report（PSUR）として主に安全性の新規情報を評価し、安全性に関する全体像を示すものであったところを、PBRER では、累積情報に基づいて、リスクのみならずベネフィットやベネフィットとリスクのバランスについても定期的に評価し報告することになった。

日本では、医薬品のベネフィット・リスク評価についての議論が高まる前に、PBRER 導入に関する情報もたらされた形になり、まるで「PBRER に記載するためにベネフィット・リスク評価を行わなければならない」かのように誤解されている風潮もあるようである。しかし、これは目的と手段を取り違えた誤った認識である。ベネフィット・リスク（再）評価は意思決定の必要に応じて行うものであり、**PBRER 作成を目的に行うものではない**。PSUR について考えてみるとよい。PSUR を使って安全性定期報告を行っている会社も、PSUR で報告しなければならないという理由で安全性情報の収集や各種の安全性評価・対策を行っているわけではないはずである。上記のような理由により、本報告書は「PBRER の書き方」を説明することを目的としていない。

PBRER は単なる報告のフォームでなく、その内容や考え方の筋道は、むしろベネフィット・

リスク評価のフレームワークそのものである。タイトルにある「PBRER をフレームワークとして利用する」とはこのような意図を示したものである。

実際、PBRER はベネフィット・リスク評価で一般的に使用されるフレームワークとよく対応している (図 2-1)。

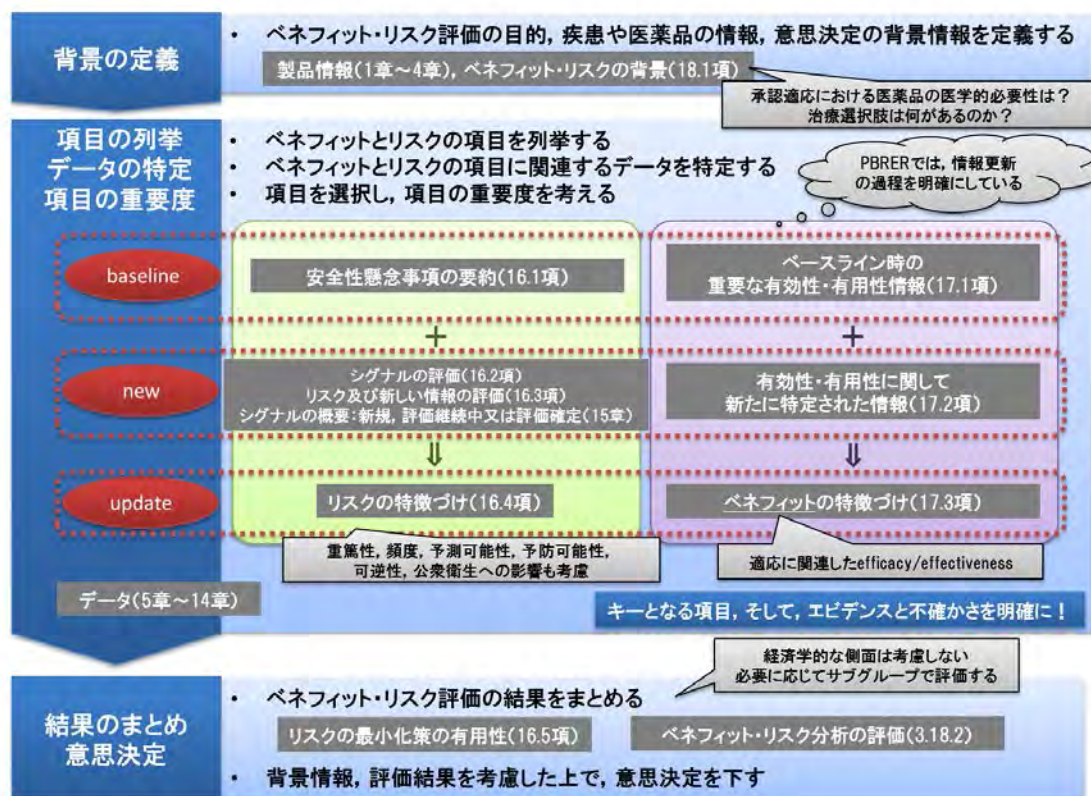


図 2-1 一般的な定性的フレームワークと PBRER の対応関係

青色は一般的な定性的なフレームワークであり、灰色は PBRER の記載事項である。フレームワークと PBRER の記載事項は対応している。

PBRER では、製品情報の詳細を記載し、ベネフィットとリスクに関してそれぞれ事前に得られていた情報を要約したうえで、新たに得られた情報を評価し、最新のプロファイルを明確にするなど、意思決定のプロセス及び根拠を明確にする工夫がなされている。以下では、フレームワークのそれぞれのステップを PBRER の各章に照らしながら復習したい。

PBRER は全 19 章で構成されており、大まかに分類すると、1~5 章が医薬品の使用状況などの背景情報、6~14 章が実際の各種データに基づく情報、15~19 章が評価および結論となっている (表 2-1)。

表 2-1 PBRER の構成

表紙	
エグゼクティブサマリー	
目次	
1. 緒言	} 背景情報
2. 世界各国における販売承認の状況	
3. 安全性上の理由で調査期間内に実施された措置について	
4. 安全性参照情報の変更	
5. 推定使用患者数と使用実態	
6. サマリーテーブルのデータ	} 各種データ
7. 調査期間中の臨床試験で認められた重大な安全性情報の要約	
8. 非介入研究からの知見	
9. 他の臨床試験及び情報源からの情報	
10. 非臨床データ	
11. 文献	} 評価および結論
12. 他の定期報告	
13. 比較臨床試験における有効性の欠如	
14. データロックポイント後に入手した情報	
15. シグナルの概要：新規，評価継続中又は評価確定	
16. シグナル及びリスクの評価	} 評価および結論
17. ベネフィットの評価	
18. 承認適応に対する包括的なベネフィット・リスク分析	
19. 結論及び措置	
20. PBRER の添付資料	

## 2.1 背景の定義

ベネフィット・リスク評価／再評価で最も大切なステップであり，すべての意思決定の礎となる情報である．ベネフィット・リスク評価の目的（何のために，誰が，どんな選択肢の中から意思決定するのか），対象疾患の情報，対象となる医薬品の情報（剤型，用量，使用状況），得られている情報（不確定情報も含め），リスク最小化策など，意思決定の背景情報を定義する．このような背景情報の定義は以降のステップの土台となるだけでなく，関係者間における情報の共有を促進することにもなる．手を抜かずに丁寧に作業することが適切なベネフィット・リスク評価を行うためには肝要である．

ベネフィット・リスク評価の目的として最も重要なのは，ICH E2C (R2) ガイドライン「定期的ベネフィット・リスク評価報告（PBRER）」に記載されているとおり，「医薬品のリスク及び承認適応に対するベネフィットに関する新しい情報又は明らかになりつつある情報の包括的，簡潔かつ重要な分析を行い，適切な場合には，ベネフィット・リスクプロファ

イルの最適化を目的とした措置を提案する」という意思決定を行うことである。

対象集団の情報としては、疾患の疫学情報（重症度，罹患率等），診断や治療のガイドライン状況，アンメット・ニーズ，他の治療の選択肢，治療集団の特性などの基本情報を整理しておく。先にも述べたように，診断基準や治療手段はベネフィット・リスク評価に影響を及ぼすものであるため，対象疾患および対象医薬品の相対的状況などはこの時点でまとめておくことになる。なお，PBRRER においては，18 章（承認適応に対する包括的なベネフィット・リスク分析）でこれら情報を記載することになるが，これはベネフィット・リスク評価をまとめる報告書として便宜上最終段階で記載するよう求められているだけであり，実際の評価時には最初になされるべきステップであることには留意されたい。

対象医薬品の情報としては，作用機序，用法・用量，剤型，投与経路などをまとめておく。なお，PBRRER では，1 章において同様の情報を記載することになる。

また，医薬品の使用実態に関する情報も重要である。臨床試験の患者集団は年齢，性別，合併症などが限られた集団であるが，市販後ではより幅広い患者層に用いられる可能性を考慮する必要がある。もし，ベネフィット・リスクプロファイルに重要な差異が存在するような対象集団があるなら，ベネフィット・リスク評価を対象集団別を実施するべきである。PBRRER では，これら医薬品使用患者の情報を 5 章の「推定使用患者数と使用実態」に記載することになっており，特殊な集団（小児，高齢者，妊婦，肝障害を合併する患者等）に対する承認後の使用なども可能な限り記載することになっている。

## 2.2 ベネフィットとリスクの項目の列挙

ベネフィットおよびリスクの項目を列挙するステップである。ここで列挙するベネフィットやリスクの項目としては，臨床試験の主要・副次評価項目や臨床アウトカム，重要な特定されたリスクや潜在的リスクとなるような項目で，全体像を把握するために網羅的に列挙する。Value Tree などを活用するのも有効な手段と考えられる（図 2-2）。上述の通り，すでに承認申請時の情報に基づいてこれらの項目の列挙は終わっているはずなので，具体的にはその後，どのような追加情報が得られたかを考えることになる。





図 2-2 Value Tree の例

まずベネフィットの項目については、承認申請時より特に新たな情報が得られていない場合には、実質的に臨床試験データを持ち越して使用することになる。PBRER の報告対象となるベネフィットは ICH E2C (R2) ガイドラインで「承認適応に関するもの」と定められており、PBRER ではそれ以外を対象外であるが、ステークホルダーにとって重要なベネフィットの項目について新たな知見が得られているのであれば、もちろんベネフィット・リスク評価に含めることを検討するべきである（ベネフィット・リスク評価は PBRER 作成のために実施するのではないことを思い出してほしい）。ただし次章から述べるように、市販後のデータソースのエビデンスレベルにはばらつきが大きいので、適切に管理された臨床試験から得られたデータでない場合には、そのデータの信頼性についてはよく検討すべきである。PBRER では、図 2-1 で示したようにベースラインのベネフィット情報については 17.1 項が、PBRER 調査期間中に得られた新たなベネフィット情報については 17.2 項が、ベネフィットのエビデンスの確からしき等のレビュー内容については 17.3 項が対応する。

次にリスクの項目については、承認申請時に明らかとなり、医薬品リスク管理計画 (RMP) にまとめられている安全性の懸念事項（重要な特定されたリスク、重要な潜在的リスク、重要な不足情報）が中心となるが、それらの既に知られているリスクのアウトカムに加え、ベネフィット・リスクを再評価する時点までに得られた安全性シグナル評価の結果（「シグナル」についてはコラム参照のこと）、新たにリスクとして判断されたものや既知のリスクの新規情報についても考慮する必要がある。なお、日本では RMP が導入されたばかりで対象製品もまだ限られており、特に PBRER 報告対象となる古い製品では RMP が整備されていないこともあるだろう。そのような場合でも、実質的には RMP 作成と同じ手順で当該医薬品の安全性情報をまとめて安全性の懸念事項と同様のものを整備することが必要になる。PBRER では、15 章の「シグナルの概要：新規、評価継続中又は評価確定」および 16 章「シグナル及びリスクの評価」が対応する。

コラム：シグナルとリスク

CIOMS VIII のシグナルの定義は、「単一あるいは複数の情報源（観察や実験）から得られた情報であり、それらは、介入と事象もしくは関連した事象の組み合わせ、あるいは有害もしくは有用な事象の間に示唆された新たな潜在的な因果関連や、既に知られていた関係での新たな側面を示すものであり、検証に足りる十分な可能性があると判断されたもの」である[10].

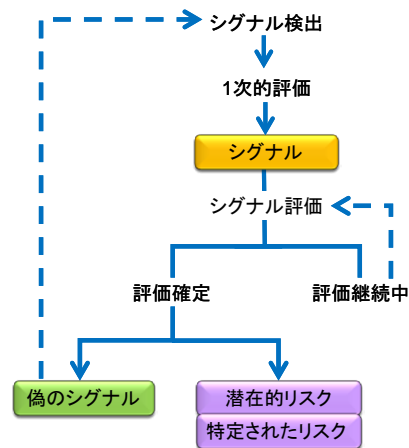


図 2-3 シグナル検出・評価のフロー

E2C (R2) ガイドラインにおいても基本的には CIOMS VIII の定義を採用しているが、有害な事象のみに焦点が絞られている。図 2-3 のように、検出されたシグナルのうち、医薬品製造販売承認取得者が一次的の評価を行い、追加で評価が必要と判断されたシグナルが報告対象となっている。シグナルは、医学的判断および情報の科学的な評価に基づき「偽の」シグナルとして棄却されるか、潜在的リスクまたは特定されたリスクとして分類され、その評価が確定する。「偽の」シグナルとして棄却されたシグナルは、新規情報の追加などにより再度シグナルとして取り上げられ評価される可能性がある。なお、リスクの定義は RMP と同様である。

### 2.3 データソースの特定

ここでは、列挙したベネフィットやリスクの項目についてどのようなデータソースから追加情報が得られるかを整理する。すべての利用可能なデータを特定し、各データソースのエビデンスレベルを精査し、必要に応じて取捨選択する。市販後では、開発段階とは異なり多彩なデータソースが考えられることに加え、同じ項目に対して複数のデータソースから情報が存在することも十分にあり得るので、すべてのデータをリストアップし、主要内容を要約して、各項目のエビデンスの不確かさのレベルを確認することが重要である。

ICH E2C(R2)ガイドラインに、「PBRER 作成時に使用する可能性があるデータソースの例」(表 2-2) が添付資料として示されている。

ベネフィットのデータソースは、新たなベネフィット情報が得られていない限り、臨床試験の情報を市販後でもそのまま使用することになる。一方リスクに関しては、自発報告が現在でも市販後におけるシグナル検出及び安全性情報の重要なデータソースである。

表 2-2 PBRER 作成時に使用する可能性があるデータソースの例

1	システムティックレビュー及びメタアナリシス
2	未承認の適応や集団に関する研究を含む臨床試験
3	登録制度 (registries) などの観察研究
4	製品使用データ及び医薬品使用実態に関する情報
5	積極的サーベイランス (active surveillance) システム (例: 拠点医療機関 [sentinel sites])
6	公表された科学文献又は学会で発表された情報を含む抄録に由来する報告
7	自発報告 (例: MAH [Marketing Authorization Holder] の安全性データベース上での報告)
8	非臨床試験
9	MAH が運営するウェブサイト*
10	未発表の原稿
11	ライセンス供与先, 他の治験依頼者又は学術機関/研究ネットワークからの情報
12	患者支援プログラム
13	製品の品質に関する調査
14	規制当局からの情報
*詳細は ICH E2D ガイドライン「承認後の安全性情報の取扱い: 緊急報告のための用語の定義と報告の基準について」を参照	

近年は自発報告の個別評価, 集積評価に加え, 統計的なシグナルの検出なども実施されている。しかし, エビデンスの分類から考えると, 自発報告のエビデンスレベルは決して高くはなく, 自発報告で集められたデータは, リスクの定量化に用いることは出来ないと考えられている。これは, 自発報告が, 曝露情報 (分母データ) の不足, 過少報告, 報告バイアスなど, 様々な欠点を伴っているためである。

「市販後はリスクの情報が多く積み上がっていくため, それに合わせて, 市販後で得られたベネフィット情報も追加しなければ, ベネフィット・リスクバランスを保てないのではないか?」という疑問をよく耳にする。しかしリスク情報が増えることは, 様々なエビデンスレベルの報告により, その情報の精度が上がる, 不確実性が減るという側面もあり, 一概にリスクの方に天秤が傾くわけではない。したがって, リスク情報が増えるからとの理由でやみくもにベネフィット情報を収集する必要は全くない。

日本では, 製造販売後調査も重要なデータソースである。これは表 2-2 の 3 あるいは 5 に該当するだろう。従来の製造販売後調査の多くは特定の医薬品に曝露された集団のみの観察研究であり, ベネフィット・リスク評価において肝心な比較対照データについては外部に求めざるを得ないという限界がある。

個々の調査・研究等のエビデンスレベルを整理する手段として, 表 2-3 のようなチェック表を利用して, エビデンスレベルを判断した理由を記録することで透明性や論理性の向上を図るのもよいだろう。BRAT のフレームワークでもデータソースを整理するステップで,

試験デザインの特徴と共に試験の質を評価することを提案しており[11]，市販後でもこれは同様である。

表 2-3 エビデンスレベルチェック表サンプル

データソース	ベネフィット/ リスク項目	エビデンスレベル				判断理由等
		I	II	III	IV	
メタアナリシス	ベネフィットA	X				多数のRCTを対象としたメタアナリシス
臨床試験	ベネフィットB	X				大規模第III相試験の主要評価項目
公表論文 (後ろ向きコホート研究)	リスクC		X			対象者の選定・規模が妥当．重要な交絡因子*の調整や感度分析有り
製造販売後調査	リスクD			X		3000例1年の1群調査結果(発生頻度変化)
自発報告	リスクE				X	自発報告からの集積

それぞれの項目に対して重要なデータソース（根拠データ）として選択された試験・研究の特性をより明確に把握するために，それぞれの試験や研究の概要をテーブルなどでまとめるとよい（表 2-4）．具体的には，選択理由，試験・研究デザイン，症例数，比較対照，用量，評価指標などを盛り込む．なお，評価指標としては，リスク比，オッズ比，リスク差などが扱いやすいだろう．当然のことながら，これらの指標を算出するためには対照群が必要である．また，データソースの幅が広い市販後では，普段から各項目に対するデータソースのエビデンスレベルをどのように判定するか，適宜整理しておくことが望ましい。

表 2-4 データソーステーブルの一例

研究番号	項目	研究デザイン	評価指標	点推定値	信頼区間	実薬群			対照群		
						N	イベント	%	N	イベント	%
AAA	ベネフィット1	RCT,DB,Placebo	リスク比	X.XX	X.XX-X.XX	XX	XX	X.X	XX	XX	X.X
BBB	ベネフィット2	RCT,DB,Placebo	リスク差	X.XX	X.XX-X.XX	XX	XX	X.X	XX	XX	X.X
CCC	リスク1	RCT,DB,Placebo	リスク比	X.XX	X.XX-X.XX	XX	XX	X.X	XX	XX	X.X
DDD	リスク2	コホート研究	ハザード比	X.XX	X.XX-X.XX	XX	XX	X.X	XX	XX	X.X

PBRER では，7～14章において章毎に各種データソース別の情報を取りまとめるようになっており，切り口は反対であるが同様の作業が求められている．さらに，16章及び17章においてそれぞれリスク及びベネフィットについて評価指標などを含めた特徴づけを行うことになっている。

## 2.4 項目の優先順位付け，重み付け

本ステップでは，リスクまたはベネフィットの各項目について，データの内容をレビュー

すると共に相対的な重要度を考える。すなわち、前ステップまでに挙げられたベネフィットやリスクの各項目のうち、ベネフィット・リスク評価を行う際に重視すべきベネフィット及びリスクを選択するステップである。立場により優先順位は異なるため、判断の根拠を明確にすることが重要である。

RMPの安全性の懸念事項やPBRERにおいては、特定されたリスクや潜在的リスクを「重要であるか否か」に分類することが求められており、それを判定する際は以下の要因を考慮する[9]。

- ◇ リスクの医学的重篤性（個別患者に対する影響など）
- ◇ リスクの頻度、予測可能性、予防可能性及び可逆性
- ◇ 公衆衛生に対する潜在的影響（頻度、治療対象集団のサイズ）
- ◇ 一般市民がリスクを認識した結果として、予防的ベネフィットのある医薬品を回避する可能性

項目の優先順位付けや重み付けは定性的でも定量的でもよいが、上記の通り、具体的な基準があるわけではなく、相対的かつ主観的にそのリスクが「重要か否か」が決定されることがわかる。このように、相対的な重要性（重み付け）については、意思決定者を含む各ステークホルダの考え方による違いや、他の治療手段の登場などの治療環境の変化による違いが予測されるので、事前にこの考え方について医療従事者や患者等から積極的に選好データを収集するなど、できる限りの準備をしておくことが重要である。とはいえ、すべての事象を事前に予測できるわけではないため、項目によって、またベネフィット・リスクの再評価が必要となった条件によっては、これらの選好情報を収集する時間的余裕がない場合もあるだろう。したがって、どのように優先順位付けや重み付けを行ったのか、その判断基準はどのようなものだったのかについて記録しておくことが、透明性のある意思決定のために必要不可欠である。

PBRERでは、本ステップも16章「シグナル及びリスクの評価」、17章「ベネフィットの評価」および18章「承認適応に対する包括的なベネフィット・リスク分析」に含まれている。

## 2.5 結果のまとめ

最後に、データを要約し結果を解釈する。データの要約は、MCDAなどの定量的手法を用いて総合評価値のようなものを算出する場合もあれば、比較対照とのリスク比やリスク差を表の形式でまとめるなど半定量的方法を応用する場合もある。個々の手法よりも、その評価に至った道筋やその根拠を論理的に説明できることが重要であり、意思決定者やその時の状況に応じて、適切なデータの要約を行えばよい。これまでに整備してきた情報を包

括的かつ論理的にレビューし、ベネフィット・リスク評価の結論を裏付ける前提、考慮事項及び判断を明確にする。加えて、ベネフィット及びリスクの情報の不確かさがどのように最終的な評価に影響するのかを説明し、評価の限界を考察することで、透明性（説明可能性）の向上に繋がる。

医薬品が販売承認されたということは、承認取得時点において（承認された製品情報に準拠して使用する場合には）医薬品のベネフィットがそのリスクを上回るという結論に達したことを意味するはずである。市販後に医薬品に関する新しい情報が追加されても、ベネフィットが引き続きリスクを上回るか否かを評価し、適宜必要な判断を下すことが必要になる。1章でも述べた通り、ベネフィット・リスク評価は、決まったタイミングで定期的実施するものではなく、影響を及ぼすような情報が得られた場合に迅速に実施されるべきである。したがって、ベネフィット・リスク評価の結論においては、前回の評価以降に得られた新しい情報が全体的なベネフィット・リスクバランスに及ぼす影響を明示することで、継続的なベネフィット・リスク評価が可能となる。

市販後においては、新たに得られる情報は上述の通り安全性情報であることが多いだろう。上述の過程を経て十分に評価したベネフィット・リスクバランスも、突然もたらされた重要なリスク情報によって瞬時に崩れてしまうかもしれない。しかしここで報告数や、分母情報が正しく得られていると信じられるエビデンスの高い情報源における発現率などの数値の増減だけにとらわれるのではなく、その情報もたらす臨床的重要性をよく考える必要がある。特に定量的手法を利用する場合には、総合評価値のようなものの変化やその大小に目を奪われがちになるかもしれないが、MCDAに関する報告書[3]でも繰り返し述べた通り、その裏にある臨床的価値に関する質的な解釈や判断こそが、ベネフィット・リスク評価の鍵であることを忘れてはならない。

PBRER では、18章の「承認適応に対する包括的なベネフィット・リスク分析」が該当する。PBRER でのベネフィット・リスク分析の評価の留意点については ICH E2C(R2)ガイドラインに記載されているのでここでは割愛する。

### 3 データの不確かさへの対処

本章のポイント：

- エビデンスレベルの低いデータソースも、他の情報と組み合わせたり、その不確かさの原因を吟味することで、それを活用する糸口がつかめることがある。ただし、そうするためには、相応の経験と知識が必要である。

前述の通り、市販後のベネフィット・リスク評価に利用可能なデータソースは様々でありそこからもたらされるデータの質、すなわちエビデンスレベルも一律ではない。データソースごとのエビデンスレベルを理解し整理した後、エビデンスレベルの高い情報については承認申請前の臨床試験と同様な扱いでベネフィット・リスク評価に用いることができるかもしれない。しかしその一方で、エビデンスレベルの低い情報はどのようにベネフィット・リスク評価に用いればよいのであろうか？

例えば、得られた情報間で結果やアウトカム指標に一貫性が無い場合、データの統合やその解釈に不確かさが残る。また、対照群がない場合は、同時比較対照が無いことによる不確かさが解釈を難しくするかもしれない。また、ランダム化を行わない（行えない）研究や調査では様々なバイアスが生じやすく、結果を正しく解釈するにはそれなりの医学的、疫学的知識が必要となる。ここで、「ランダム化比較試験のメタアナリシス」のエビデンスが高い理由を改めて考えてみると、次の3点が浮かび上がる[12]。

- 複数の試験を統合することで症例数が増え、**精度**が高くなる。
- ランダム化比較試験を対象とするため、**比較可能性\***が保証されている
- 複数の試験間で効果の方向や大きさの異質性（heterogeneity）を調べることで、**一般化可能性**を検討できる

逆に言えば、市販後のベネフィット・リスク再評価に利用したいデータソースについて、これらの点に疑問が残るようであれば、その情報のエビデンスレベルは低いかもしれない。これからデータを取得するなら、もちろんエビデンスレベルの高いデータが得られるようによく計画すべきである。しかし以下では、市販後に非自発的に集積されるデータや、すでに手元に得られているデータについて考えよう。すでに手元に得られているデータのエビデンスレベルが高くないことが想定される場合には、エビデンスとして上記3点の何に問題があるのか、またそのエビデンスレベルの低下をもたらすデータソースの不確かさの種類は何なのか、を理解することが必要となる。データソースの不確かさの種類がわかれば、その不確かさの種類に応じた対処法を検討することで、それらのデータソースを含めたベネフィット・リスク評価が可能となる。

市販後のベネフィット・リスク再評価に用いるデータソースにどのような種類の不確かさ

が含まれているのか、あるいは含まれている可能性があるのかを整理し、それらの不確かさがエビデンスにどのような影響をもたらしているかを表 3-1 に示した。

**表 3-1 データソースの不確かさの種類とエビデンスに与える影響**

不確かさの種類	エビデンスに与える影響
症例数が少ない	精度
対照群がない 交絡因子が存在する 情報バイアスが存在する	比較可能性
情報が足りない	比較可能性, 一般化可能性

以下の節ではそれらの例ごとに整理した上で具体的にどのようにエビデンスレベルの低下をもたらすかを考察し、その不確かさの種類に応じて対処できる可能性のある手法をいくつか紹介する。

### 3.1 必要な情報に関して症例数が少ない

一般的に市販後には、市販前と比較して幾何級数的に曝露症例数が増加する。ところが、ベネフィット・リスク評価に必要なデータ、あるいはエビデンスレベルが高いデータとなると、症例数が限られているかもしれない。例えば多くの患者に使用されている薬剤について、ある有害事象に関する自発報告がない場合、本当にその有害事象の発生がない（あるいはかなり少ない）ため報告がないのか、疑問が残る場合がある。一定の手順を踏んで積極的な追跡を受けた症例数が実は限られており、ある程度の経験や知識、検査等を経ないと発生の確認が難しい有害事象の場合など、特にこうした問題が起こりやすい。あるいは希少疾患のように、市販後であっても曝露症例数が少ない場合もあるだろう。

必要な情報に関して症例数が少ない場合には、情報の精度が低く、はっきりした結論が得られないだろうことは容易に想像がつく。しかし市販後においては、ごく少数の報告例からのデータのみに基づいて、何らかの意思決定を迫られる場合もあるだろう。このような場合、症例数が増えるまで待つことのリスクと、安全性マージンを考えて早期に対策を打つことで結果的にこうむるかもしれないデメリットについて、天秤にかける必要にせまられる。

ベネフィット・リスク評価やそれに基づく意思決定は本質的に主観的なものであり、例えばある未知で重篤な有害事象の発生の医学的・社会的インパクトが誰の目にも明らかなのであれば、症例数が不足して因果関係に不確実性があったとしても、比較的容易に意思決定ができるかもしれない。このような場合に、わざわざ時間をかけてモデルを作る等



の回り道をして再評価を行う意義は薄い。どのデータに基づいて誰がいつどのような議論がなされ、その結論に至ったのかを明示的に記録しておけばそれで充分であろう。

一方で、意思決定に迷う場合もあるかもしれない。その有害事象のインパクトや頻度を考慮すると、ベネフィット・リスク・バランスに結果的に影響を与えない可能性もあるかもしれないと考えられるような場合である。このような場合には、それぞれのベネフィットやリスクの項目に関するばらつきを把握し、明示することが第一歩となる。具体的には、点推定値だけでなく信頼区間も求め、また図示することで、ばらつきの範囲を把握する。適切であれば、ベネフィット・リスク評価に用いる他の評価項目の信頼区間も一緒に図示した、いわゆる **Forest Plot** を作成し、あわせて検討するとよい。特に信頼区間の端に注目し、ベネフィット・リスク評価の観点からもっとも好ましくない場合を想定して検討することも有意義かもしれない。

また、定量的なベネフィット・リスク評価を行っている場合には、ばらつきが大きい評価項目の値を、例えば信頼区間の範囲内で動かしてみて、ベネフィット・リスク・バランスがどの程度変化するか調べることができる。こうした検討を、一般に感度分析と呼ぶが、いろいろ値を変えてもベネフィット・リスク評価の最終的な結論が変わらないことが確認できれば、症例数が少ないことによる不確かさがあっても一定の結論は得られるだろう。

さらに、市販後あらたにもたらされた情報を用いて、既存の情報を更新する際、**ベイズ\***的な手法も有用かもしれない。症例数が少ない場合であっても、事前分布として医療従事者の意見や論文情報等を適切に考慮することで、事後情報の精度を上げることができるかもしれない。事前分布についてはいろいろな状況を想定し、感度分析を行う必要があるだろう。ベイズ的な情報更新の詳細は別添を参照されたい。

### 3.2 対照群がない

そもそもベネフィット・リスク評価は選択肢からどちらかよいものを選ぶという意思決定をサポートするものなので、何らかの比較対象は必須であることを念頭に置くべきである。しかしながら、自発報告には本質的に対照群の情報は含まれていないし、従来型の製造販売後調査にも対照群を置かない場合がほとんどである。このため、いずれのベネフィットやリスクのアウトカムについても、それが対象医薬品によるものなのか、もっと別の原因で生じたものなのか、解釈することができない。

このように対照として比較すべき情報がない場合には、何らかの形で、ヒストリカルコントロール等の外部情報に頼らざるを得ないかもしれない。例えば、ごく稀で重篤な有害事象の場合、その患者集団で一例でも発生したら、それだけで問題である、といった場合があ

る。これなども、そのような患者集団でそのような事象が発生する頻度は0であるはずだ、という外部情報と暗黙のうちに比較したうえでの判断になっている。未知の患者集団で、疫学データの蓄積が乏しいほど、こうした「暗黙の」対照すら想定しにくくなる。

また、再発も考えられるようなリスクイベントで、曝露の影響が速やかに出現し、かつ消失するものについては、ケース自身の曝露とイベントの履歴を詳細に調べ、いわばケース自身から得られた情報を対照とみなす方法が適応できる場合があるかもしれない[13, 14].

さらに、自発報告のデータベースに対して統計的手法を適応するシグナル検出では、問題となる有害事象以外の有害事象を生じた人を対照とみなすことが多いが、事情が許せば、こうした方法の適応も考えられる[15].

いずれにせよ、一見対照が存在しないように思える状況下であっても、実際にベネフィット・リスク評価を行おうとすると、例外なく、何らかの対照が必要となる。

### 3.3 交絡因子が存在する

ランダム化が行われていない研究や調査では、曝露とアウトカムの両方に影響を与える要因（交絡因子と呼ばれる）が、結果の解釈に影響を与える場合がある。例えば、より安全な医薬品が背景リスクの高い患者に投与されやすく、見かけ上の有害事象発症率が高くなる現象がある。一方、より有効であると考えられる医薬品は、重症度の高い患者に投与されやすいので、見かけ上の有効性が低くなる場合もある。サブグループ解析の結果を参照したり、個々の症例レベルのデータが利用可能な場合はこうした要因の影響を調整した解析を行うことで、ある程度対処することが可能である[16, 17].

コホート研究やケース・コントロール研究等、交絡の影響を完全には制御できない懸念がある研究から得られた結果を解釈する際には、さまざまな注意が必要である。上述した交絡要因の影響を調整した解析、例えば傾向スコアを用いた解析や多変量解析等の結果にも注意を払うべきである。また論文によっては明記されていないかもしれないが、研究それ自体がもつ限界に関する記載にもよく注意して結果を解釈する必要がある。さらに感度解析等、主論文の解析内容を補足する解析の結果が別論文として公表されている場合もあるので、専門家の協力を仰ぎつつ、補足的な情報にも目を配ることが必要である。

### 3.4 バイアスがある

バイアスのないデータはなく、バイアスに事後的に対応するための決定的な方法も存在しない。とは言え、そもそもどのようなバイアスの懸念があるかを吟味することは、どのような医学情報を解釈する際にも常に必要である。どんな場合にも適応できる魔法のような

方法はないが、問題となる疾患や薬剤、（薬剤）疫学等に対する基本的な理解が必須であることは間違いない。代表的なバイアスの例について巻末で紹介を試みたが、紙幅も限られており、あくまでも参考である。

バイアスの有無について吟味する際、多くの場合、本当はどのような研究を行うことが理想的か、なるべく具体的にイメージすることが出発点となる。というのは、その理想と比較したときに欠けていること、実現できないことが、そのままバイアスの原因となりうるからである。例えば、多くの医学研究では、①どのような集団で、②どのような曝露や介入に対する、③どのようなアウトカムを、④どのように比較することで知りたい疑問にせまることができるのかを具体化することが不可欠である。そして、このような形に構造化した研究上の問いを **research question** と呼ぶことがある。なるべく具体的な **research question** を念頭に置きながら医学情報を吟味することで、自分が何を知りたいのか・知る必要があるのか、より具体的に検討を進めることができる。

バイアスとは、一言でいえば理想的な研究を行えば知り得たであろう真の結果とは異なる結果が観察されることである。やや別の言い方をすると、バイアスとは、理想的とは言い難い不適切な研究や調査により、真実とは異なる結果が得られることであり、適切な研究や調査を行ったにもかかわらず被験者数の不足等により真実とはやや異なる結果が得られることがあるという確率論的な現象、すなわち「ばらつき」とは別の概念である。

このためバイアスについては、「理想的な研究」なるものの内容が具体的であればあるほど検討がしやすい。例えば、①については A という患者集団について調べたいのに文献では B という集団について調べていた。②についてはある薬剤の単剤投与について知りたいのだが併用も許していた。③についてはイベントの定義が自分が知りたい内容とはやや異なっており、④については類薬との群間比較ではなくある薬剤を使用する前後の比較が行われていた。こうした、本当に自分が知りたい・知るべきデータからの偏差が、そのままバイアスの原因となりうる。こうした違いが結果にもたらす影響が、既存の医学情報等から大きくないことがわかれば、バイアスの懸念は小さいかもしれないが、そうでなければそのデータを鵜呑みにすることはできないだろう。

文献や医学情報を批判的に吟味するというのは、研究の細部にまで目を凝らして揚げ足取りをすることではなく、自分が知りたい **research question** と照らし合わせたときに、その情報が自分の問いに本当の意味で答えてくれているかを問うことである。この種の検討には、それなりの経験と知識が必要であり、一朝一夕に解決できる問題ではない。しかし、さまざまなデータソースからもたらされるデータを評価するうえで、避けて通れない問題であることも確かである。

なお、Turner らが提案するように各種のバイアスが存在しうるかどうかを整理した上で、そ

それぞれのバイアスの影響を評価し、効果の推定に組み込むことが一つの対処法となりうる場合もある[18]。詳細は参考文献を参照いただきたいが、ごくかいつまんで説明すると、「比較される介入群は同じ時期に登録されているか？」などのチェック項目に「Yes / No / Unclear」といった回答を行ってバイアスの有無の可能性について整理し、その回答結果をまとめることで各種のバイアスがどの程度影響をもたらしているのかを「High / Medium / Low / None」の4段階で半定量的に評価する。その後それらのバイアスの影響を、効果を推定するモデルの因子として組み込むことで、効果の推定値や分散に対するバイアスの影響を考慮した結果が得られる、というような方法である。ベネフィット・リスク再評価内でもそれらの推定値を用いることで、バイアスの影響を加味することがどのように結論に影響を与えるかを検討することが可能な場合もあるだろう。

### 3.5 そもそも情報がない

測定されていない評価項目は評価出来ないし、論文情報を用いる場合など、サマリデータのみしか入手できない場合は、被験者レベルのデータが入手できる場合と比較して検討できることを選択肢が狭くなる。観察期間が短い場合は、長期投与に関する情報は得られない。そもそもデータソース内にベネフィット・リスク評価に必要な情報が存在しないことが原因なので、根本的な対処法はデータを取得すること以外ない。とは言え、項目間の相関に関する情報や、各項目を重み付けするための選好情報など一部の情報不足については、既に『ベネフィット・リスク評価 中級編 定量的手法に関する考察』で紹介したようなLyndらによるシミュレーションの方法や[19]、SMAAなどの方法[20, 21]による対処が参考になる場合もあるかもしれない。ここで重要なことは、必要になってからデータ収集を開始するのではなく、ベネフィット・リスク評価で必要になるデータを見越して、試験を組んでおくことである。特に特定集団における使用経験や選好の情報などは、あらかじめ計画し、取得しておかなければ必要な時に情報は得られないままとなる。

## 4 おわりに

本報告書では、市販後だけに応用可能な特別なベネフィット・リスク評価法が存在するわけではないこと、市販後のベネフィット・リスク再評価の特徴は、エビデンスレベルの異なる様々なデータソースから玉石混交のデータが得られ、その評価を求められること、および意思決定までの時間が限られるかもしれないことであることを述べた。したがって、まずは承認申請時のエビデンスレベルの高いデータソースをもとに、ベネフィット・リスク評価をおこなっておくこと、情報不足やデータの不確かさについて明らかにすること、データソースのエビデンスレベルを検討しておくことなどで、再評価が必要となるタイミングに備えて十分に準備しておくことを推奨した。

市販後のベネフィット・リスク再評価のポイントは、エビデンスレベルの異なるさまざまなデータソースについてよく理解し、それらから得られるデータを正しく扱うことである。しかし、エビデンスレベルを評価できるようになるためには、それなりの経験と基礎知識が必要であり、一朝一夕に解決できる問題ではない。しかし、市販後のさまざまなデータソースからもたらされるデータに基づいてベネフィット・リスク再評価を行うためには、避けて通ることができない問題でもある。

また、ベネフィット・リスク再評価は PBRER 作成のために実施するのではないことも述べた。ベネフィット・リスク評価は必要に応じて随時更新すべきものである。PBRER はその時点での情報を報告するツールであるが、一方で単なる報告フォーマットとして扱うにはもったいないほど、ベネフィット・リスク再評価のフレームワークをよく示している。そこで、PBRER をガイドとしてベネフィット・リスク再評価を行うことを提案した。

治験薬の臨床開発段階においても、ICH E2F ガイドライン「治験安全性最新報告」より、臨床開発段階における定期的な安全性情報の分析を行う治験安全性最新報告（The Development Safety Update Report : DSUR）の作成が必要となり、その時点での累積安全性データから特定されたリスクと、予測される有効性/ベネフィットのバランスについて簡潔に述べることとなっている。今後は、臨床開発段階から、ベネフィット・リスク評価を随時実施し、その結果が承認時の医薬品リスク管理計画（RMP）に反映されることになり、さらに市販後のベネフィット・リスク再評価に引き継がれていくというように、開発段階からの市販後までを通して、ベネフィット・リスク評価を行うという考え方が重要になるであろう。

再評価のさらに先を考えると、ベネフィットを最大化し、リスクを最小化するための活動や、その活動の評価が提案されることもあるだろう。これら措置は RMP と密にリンクしているため、（RMP が整備されている医薬品においては）必然的に RMP も合わせて改訂されることになるだろう（図 4-1）。

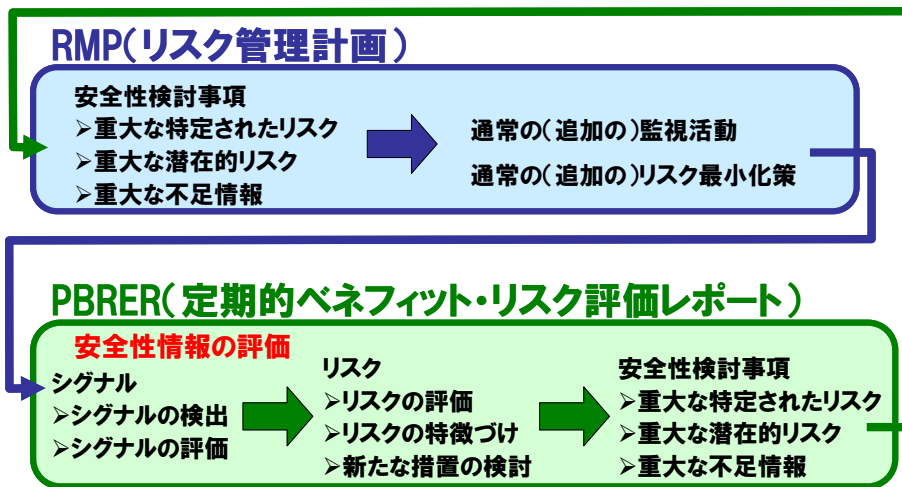


図 4-1 RMP と PBRER の関係

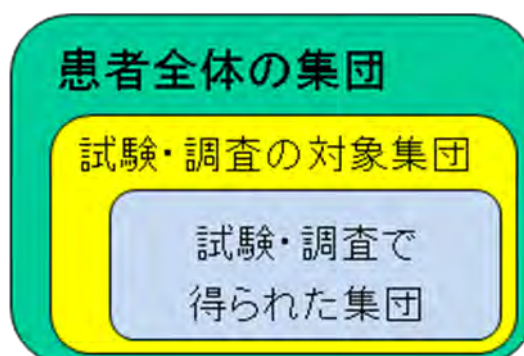
RMP で活動を計画し，実行した結果，再びベネフィット・リスク再評価を実施することになる．このように，ベネフィット・リスクバランスの最適化は，ライフサイクルを通じて何度も繰り返し行われるのである．

市販後には対象医薬品に関する情報が更新されるだけでなく，医薬品を取り巻く環境もダイナミックに変動し続ける．起こることをすべて想定することはできないが，柔軟かつ迅速に対応できるよう，できる限りの準備をし，ベネフィット・リスク再評価に備えることが肝要である．

## 5 別添：用語解説

### 一般化可能性

試験や調査で得られた集団（標本）から導かれた結果を患者全体の集団（母集団）へ外挿することについての信頼の程度のこと。標本と母集団が大きく乖離している場合は、一般化可能性の問題が生じることになる。



### 比較可能性

試験治療の受けた群の被験者と対照群の被験者を比較検討する上で、それぞれの被験者の性質が同質である程度のこと。その性質には既知の性質（測定されている性質）の他に、未知の性質（測定されていない性質）も含まれるため、高い比較可能性を得るためには、ランダム化割付が最適である。

### 精度

試験治療の効果の推定値と真の値の間に生じる偶然の誤差の程度を精度（推定精度）と呼ぶ。偶然誤差はその名の通り、偶然に発生するため、真の値の周りに正負の傾向をもたずに発生する。偶然の差が小さくなることで、推定精度は高くなる。この偶然の差は標準誤差と呼ばれ、試験における調査例数を増やすことで誤差は小さくなり、推定値は真の値に近づく。

### バイアス（偏り）

試験治療の効果の推定値と真の値の間に生じる系統的な誤差は偏りもしくはバイアスと呼ばれる。この誤差が小さい場合に、正確度が高いと言われる。バイアスは、試験の計画、実施、解析及び結果の解釈と関連した因子の影響により、真の値から一定の傾向で偏った誤差を生じる。

各種のバイアスは、調査デザインやデータソースの種類によって混入のしやすさが異なり、複数の調査間での結果の一貫性の欠如を引き起こし、解釈を困難にする。また、一度生じたバイアスは、一部の例外を除いて統計的手法では取り除くことが出来ない。そのため、バイアスが発生する原因を事前に排除する努力が求められる。

### 選択バイアス

母集団から標本を抽出する際や、群の割付け時に発生するバイアスである。前者のバイア

スがあれば、母集団と標本に乖離が生じ、一般化可能性に問題が残る。後者のバイアスがあれば、試験治療群と対照群で被験者背景が異なってしまうことになり、比較可能性に問題が生ずる。

### 情報バイアス

観察者がアウトカムの測定や評価を行う際に発生するバイアスである。主なものは表 5-1 を参照されたい。例えば自発報告は重症度の高い副作用や新規薬剤の副作用が報告されやすいという報告バイアスが起りやすい。先に発売された同種の薬剤で安全性の懸念があれば、後で発売する薬剤に関しては十分な注意喚起を行うことで報告者の認識レベルが上がり、逆に報告されやすくなる現象もある。

表 5-1 情報バイアスの例

名称	内容	例
測定バイアス	観察者や、測定機器の質によるデータの違い	尿路結石の診断において、画像検査毎に感度、特異度が異なることが知られている[22]
観察バイアス	観察者が対象となる治療内容を知っていると、評価に影響がでてしまう	プラセボを服薬していることが判っている患者に発生した有害事象の因果関係判定は、「関連なし」になりやすい。
報告バイアス	イベントの有無についての報告の正確さが異なる	自発報告では重症度の高い副作用や新規薬剤の副作用が報告されやすい。
想起バイアス	過去の経験や事象を調査する際の思い出しやすさが異なる	後ろ向き研究において、ケース群の方が過去の曝露経験を思い出しやすい。

### 解析バイアス

当初予定されていた解析方法以外の方法を事後的に用いることで、試験結果を有利に見せることができる。俗に言う後付け解析であり、これを解析バイアスという。特に検証的な側面を持つ試験であれば、検証に関係する事項を事前に計画書として作成しておくべきである。

### 出版バイアス

ネガティブな結果が出た研究は、ポジティブな結果が出た研究よりも公開されにくい。これを出版バイアスという。そのため、公表データを用いてメタアナリシスを行うと、ポジティブな結果に偏りがちになる。出版バイアスの有無は、ファンネル・プロットの形で確認でき、出版バイアスが存在する場合には trim and fill 法[23]、Copas Selection model [24-26]等の方法によりバイアスを調整する手法が提案されている。

### 交絡

バイアスの一種で、ある因子が曝露とアウトカムの両方に影響することにより、曝露とア



ウトカムの関係がはっきりしなくなること。例えば、より安全な医薬品が背景リスクの高い患者に投与されやすく、見かけ上の有害事象発症率が高くなる現象が知られている。

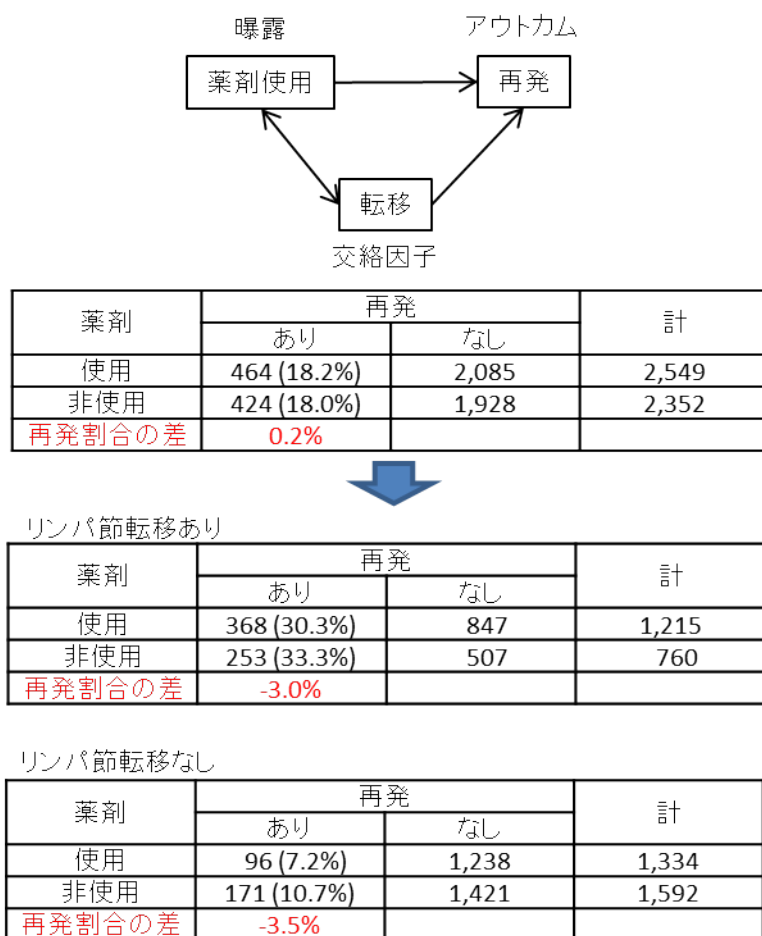


図 5-1 処方による交絡の事例[27]

図 5-1 の事例では、薬剤使用群と非使用群とでイベント（この例では再発）発症割合にほとんど差が無いように見える。薬剤の使用とイベントの発症の両方に関連のある因子（この例ではリンパ節転移）でサブグループ解析を行うと、イベント発症率がサブグループ間で大きく異なっており、いずれのサブグループにおいても薬剤使用群の方がイベント発症割合が低いことがわかる。これは、リンパ節への転移がある（再発率が高いことが知られている）患者に優先的に薬剤が投与される交絡が原因である。

## ベイズ統計

EMA の報告書（work package 2）は、規制当局の意思決定において、ベイズ統計の有用性を High と評価している[28]。ベイズ統計では、まず事前に得られている情報を事前分布という確率分布で表す。そして、ベイズの定理を用いて、得られた情報（データの尤度）に基づき、事前分布を事後分布に更新していく（図 5-2）。詳細は成書を参考されたい[29]。

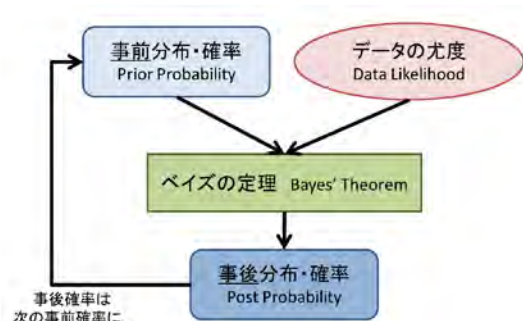


図 5-2 ベイズ流の情報更新

ベイズ流アプローチを用いてリスク情報を更新した例として、回転性粥腫切除術（rotational

atherectomy; RA) 後のリスクモニタリングを紹介する[30]. RA 後の院内合併症に関して、ベイズ流アプローチを用いて術後の MACE (major adverse cardiac event ; 死亡, 心筋梗塞, 緊急の冠動脈バイパス手術の必要性) のリスクを評価した.

事前分布は, 事前に十分な情報が無かったため, 図 5-3 の①のように設定した. MACE イベント発現割合の期待値は 50%とし, 事前分布は左右に大きく裾を引いており, 不確かさが大きいことがわかる. 表 5-2 のデータが経時的に得られ, その都度, 情報を更新していくと, 事後分布は①→②→③に更新された. イベント発現割合は時間が経つにつれて, 想定より低いことがわかり, 不確かさも小さくなったことがわかる. なお, 論文では事前分布をいくつか想定し, 感度分析を行っている.

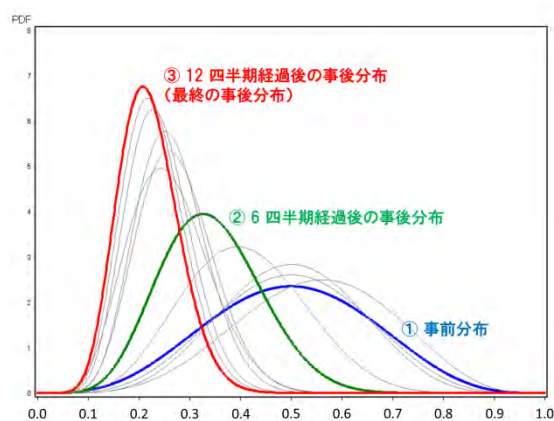


図 5-3 事後分布の更新経緯

表 5-2 MACE イベント発現頻度

年	時点		患者数	発現数	
	四半期				
1997	1		1	1	
	2		0	0	
	3		1	0	
	4		2	1	
1998	1		3	0	
	2		6	1	②
	3		7	0	
	4		6	2	
1999	1		5	1	
	2		4	0	
	3		2	0	
	4		2	0	③
Final			39	6	

ベイズ流アプローチを用いてベネフィット・リスク評価を行った論文としては, 他にも更年期障害に対するホルモン療法のベネフィット・リスク評価[31], 低出生体重児の慢性肺疾患に対するヒドロコルチゾン[32]などがある.

## 6 References

1. Council for International Organizations of Medical Sciences (CIOMS) Working Group IV, Benefit-Risk Balance for Marketed Drugs: Evaluating Safety Signals. 市販薬のベネフィット・リスク・バランス、安全性シグナルの評価. 1998.  
[<http://www.cioms.ch/publications/g4-benefit-risk.pdf>]
2. 日本製薬工業協会データサイエンス部会, ベネフィット・リスク評価 入門. 2013.  
[<http://www.jpma.or.jp/about/board/evaluation/allotment/pdf/benefitrisk.pdf>]
3. 日本製薬工業協会データサイエンス部会, ベネフィット・リスク評価 中級編 多基準決定分析への招待—その理論と事例—. 2013.  
[[http://www.jpma.or.jp/information/evaluation/allotment/pdf/benefitrisk\\_mid.pdf](http://www.jpma.or.jp/information/evaluation/allotment/pdf/benefitrisk_mid.pdf)]
4. 日本製薬工業協会データサイエンス部会, ベネフィット・リスク評価 中級編 定量的手法に関する考察. 2013.  
[[http://www.jpma.or.jp/information/evaluation/allotment/pdf/benefitrisk\\_mid2.pdf](http://www.jpma.or.jp/information/evaluation/allotment/pdf/benefitrisk_mid2.pdf)]
5. 中島章博, et al., 日本におけるベネフィット・リスク評価の現状と今後への期待. 医薬品医療機器レギュラトリーサイエンス, 2013. **44**(3): p. 216-223.
6. 松下泰之, Topics 「ベネフィット・リスク評価体験ワークショップ」を開催  
[[http://www.jpma-newsletter.net/PDF/2013\\_154\\_08.pdf](http://www.jpma-newsletter.net/PDF/2013_154_08.pdf)]. 製薬協ニューズレター, 2013. **154**: p. 22-23.
7. 森實敏夫, 臨床医のための EBM アップグレード. 医学書院, 2002.
8. Agency for Healthcare Research and Quality, Systems to Rate the Strength of Scientific Evidence. [<http://archive.ahrq.gov/clinic/epcsums/strengthsum.htm>]
9. 日米 EU 医薬品規制調和国際会議, ICH 調和三極ガイドライン 定期的ベネフィット・リスク評価報告 (PBRER) E2C (R2) 2013. [[http://www.pmda.go.jp/ich/e/e2cr2\\_13\\_5\\_17.pdf](http://www.pmda.go.jp/ich/e/e2cr2_13_5_17.pdf)]
10. Council for International Organizations of Medical Sciences (CIOMS) Working Group VIII, Practical Aspects of Signal Detection in Pharmacovigilance. ファーマコビジランスにおけるシグナル検出の実践 CIOMS VIII (レーダー出版センター). 2010.
11. Levitan, B.S., et al., Application of the BRAT framework to case studies: observations and insights. Clin Pharmacol Ther, 2011. **89**(2): p. 217-24.
12. 浜田知久馬, メタアナリシスの功罪 -MIXED プロシジャによるメタアナリシスと公表バイアスへの対応. SAS Forum ユーザー会 学術総会 講演資料, 2007.
13. Greenland, S., A unified approach to the analysis of case-distribution (case-only) studies. . Stat Med, 1999. **18**: p. 1-15.
14. Whitaker, H., et al., Tutorial in biostatistics: the selfcontrolled case series method. Stat Med, 2006. **25**: p. 1768-1797.
15. Rothman, K., S. Lanes, and S. Sacks, The reporting odds ratio and its advantages over the proportional reporting ratio. Pharmacoepidemiol Drug Safe, 2004. **13**: p. 519-523.

16. Little, R.J. and D.B. Rubin, Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu Rev Public Health*, 2000. **21**: p. 121-45.
17. Weitzen, S., et al., Principles for modeling propensity scores in medical research: a systematic literature review. *Pharmacoepidemiol Drug Saf*, 2004. **13**(12): p. 841-53.
18. Turner, R.M., et al., Bias modelling in evidence synthesis. *J R Stat Soc Ser A Stat Soc*, 2009. **172**(1): p. 21-47.
19. Lynd, L.D., et al., A quantitative evaluation of the regulatory assessment of the benefits and risks of rofecoxib relative to naproxen: an application of the incremental net-benefit framework. *Pharmacoepidemiol Drug Saf*, 2010. **19**(11): p. 1172-80.
20. Lahdelma, R., J. Hokkanen, and P. Salminen, SMAA - Stochastic multiobjective acceptability analysis. *European Journal of Operational Research*, 1998. **106**(1): p. 137-143.
21. Tervonen, T. and J.R. Figueira, A survey on stochastic multicriteria acceptability analysis methods. *Journal of Multi-Criteria Decision Analysis*, 2008. **15**(1-2): p. 1-14.
22. 日本泌尿器科学会・日本泌尿器内視鏡学会・日本尿路結石症学会編, 尿路結石症診療ガイドライン 第2版. 金原出版, 2013: p. 37-39.
23. Duval, S. and R. Tweedie, A Nonparametric “Trim and Fill” Method of Accounting for Publication Bias in Meta-Analysis. *Journal of the American Statistical Association*, 2000. **95**(449): p. 89-98.
24. Copas, J., What works?: selectivity models and meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 1999. **162**(1): p. 95-109.
25. Copas, J. and J.Q. Shi, Meta-analysis, funnel plots and sensitivity analysis. *Biostatistics*, 2000. **1**(3): p. 247-62.
26. Copas, J.B. and J.Q. Shi, A sensitivity analysis for publication bias in systematic reviews. *Stat Methods Med Res*, 2001. **10**(4): p. 251-65.
27. 佐藤俊哉 and 松山裕, 交絡という不思議な現象と交絡を取りのぞく解析 -標準化と周辺構造モデル-. *計量生物学*, 2011. **32**(Special Issue): p. S35-S49.
28. European Medicines Agency, Benefit-risk methodology project. Work package 2 report: Applicability of current tools and processes for regulatory benefit-risk assessment. 2010. .  
[[http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Report/2010/10/WC500097750.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Report/2010/10/WC500097750.pdf)]
29. 渡部洋, ベイズ統計学入門. 福村出版, 1999.
30. Resnic, F.S., et al., Exploration of a bayesian updating methodology to monitor the safety of interventional cardiovascular procedures. *Med Decis Making*, 2004. **24**(4): p. 399-407.
31. Minelli, C., et al., Benefits and harms associated with hormone replacement therapy: clinical decision analysis. *BMJ*, 2004. **328**(7436): p. 371.
32. Shaffer, M.L. and K.L. Watterberg, Joint distribution approaches to simultaneously quantifying benefit and risk. *BMC Med Res Methodol*, 2006. **6**: p. 48.

2014年3月

日本製薬工業協会 データサイエンス部会 タスクフォース1

神浦 俊文	(日本新薬)
川口 源太	(キッセイ)
木村 友美*	(ヤンセンファーマ)
小森 哲志	(バイエル薬品)
阪口 元伸	(武田薬品工業)
杉谷 康雄	(中外製薬)
中島 章博	(帝人ファーマ)
松下 泰之*	(第一三共)
山下 智広	(参天製薬)

副部会長 酒井 弘憲

\*タスクフォースリーダー