

FDA による企業のためのガイダンス：ヒト用医薬品と生物学的製剤の有効性に関する臨床的エビデンスの提出

米国 FDA 医薬品評価研究センター・生物製剤評価研究センター
1998 年 5 月 臨床 6

Guidance for industry: Providing clinical evidence of effectiveness for human drug and biological products

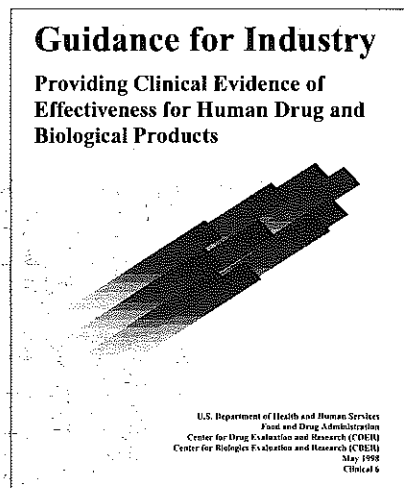
U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
May 1998 Clinical 6

訳：津谷喜一郎¹⁾，臼杵浩志²⁾

[訳者 はじめに] 1998 年 4 月 23 日に開催されたセミナー「医薬品の適応の現状と今後の課題」で、米国 FDA の Laurie B. Burke により本ガイダンスのドラフトが紹介された(臨床評価 1999; 26(Supplement) 医薬品適応外使用のエビデンス 特集: 153-190)。その後同年 5 月にガイダンスは final 版となった。このガイダンスは、外挿可能性のトピックなど、医薬品の適応外使用やその他の領域に関係の深い内容であり、翻訳紹介することとした。本ガイダンスは、<http://www.fda.gov/cder/guidance/>より“evidence”と入力しサーチすると pdf 形式で見ることができる。

目次

- I. 序論
- II. 有効性を支持するのに必要なエビデンスの量
 - A. 医薬品と生物学的製剤の法的基準
 - B. 法的基準の科学的ベース
 - C. 有効性を支持するエビデンスの量
- III. 有効性の表示を支持するのに必要なエビデンスの質に関するドキュメント
 - A. 臨床データまたは詳細な研究報告書への通常のアクセスが不十分な場合に依拠するもの
 - B. 品質管理/オンサイト・モニタリングが他の方法や十分に行われていない研究への依拠



1) 東京医科歯科大学難治疾患研究所・情報医学研究部門(臨床薬理学)

2) 日本製薬工業協会医薬品評価委員会 PMS 部会

FDA 企業のためのガイダンス⁽¹⁾

ヒト用医薬品と生物学的製剤の有効性⁽²⁾に関わる臨床的エビデンスの提出

I. 序 論

この文書は、新薬申請書(new drug application: NDA)、生物学的製剤の認可申請書(biologics license application: BLA)、または適応(indication)追加についての一部変更(一変)申請書の提出を計画している申請者に、有効性(effectiveness)証明のために提出すべきエビデンスについてガイダンスを提供するものである。

この文書はまた、1997年の食品医薬品局近代化法(the Food and Drug Administration Modernization Act, 近代化法)サブセクション403(b)(1), (2)のヒト用医薬品と生物学的製剤に関する要件(PL 105-115)を満たすことにもある⁽³⁾。サブセクション403(b)(1)は、公表文献がどのような状況下で新しい適応症の一変申請書の基礎となるか、そのガイダンスをFDAに求めている。本ガイダンスのセクションIIIは、公表文献が追加適応症の承認を部分的または全体的に支援するかを記述してその要請に込えている。また、サブセクション403(b)(2)は、原申請書の裏付資料を一変申請資料として利用することを認め、FDAに提出資料の重複を避けるためのガイダンスをつくるよう求め、本ガイダンスのセクション

IIは、原申請書またはその他既存の関連データを一変申請資料として用いられる範囲を示し、その要請を満足するものである。

1962年、議会は、連邦食品医薬品化粧品法を改正して、新しい要件を付け加え、製造者は、販売承認を得るために、「適切でよくコントロールされた研究」(adequate and well-controlled studies)を実施して、その製品の有効性を実証しなければならないこととなった。それ以来、何が有効性のエビデンスとして十分たりうるかが争点として浮上し、FDA、科学界、産業界その他によって議論されてきた。有効性についての健全な(sound)エビデンスは、FDAが新製品または新用途の効果/リスク評価を行う際に、その諾否を分ける重要な要因である。同時に、有効性の証明は、医薬品開発の期間と費用に関わる主要な要素であり、必要なエビデンスの量と性格は、新しい治療法が一般に利用できるようになるかどうか、あるいはなるとすれば何時からかを決定する重要な要件となる。有効性について信頼性あるエビデンスが効率的な形で明らかにされれば、公衆衛生への最大の寄与となる。

医薬品開発と臨床評価の理論と実践は、医薬品の有効性についての証明が求められるようになって以来、顕著な進展を遂げ、有効性を支持するた

注1: 本ガイダンスは、ヒト用医薬品と生物学的製剤の有効性を示す臨床的エビデンスの提出について、FDAが現在いかに考えているかを示すものである。単なる見解の表明であって、いかなる者にも、またいかなる者のためにも権利を創出、あるいは付与せず、FDAまたは国民を拘束する効力もない。適用可能な法または規則もしくはその両方の要件を満たすならば、代替案を用いることもできる。

注2: 本ガイダンスにおいて、「効果」(efficacy)とは、適切でよくコントロールされた臨床試験またはそのような試験の実施意図における知見(finding)、また「有効性」(effectiveness)とは、臨床的効果(clinical efficacy)その他のデータをもとに下された規制当局の決定に基づくものをいう。

注3: 近代化法セクション403の規則はまた、動物用医薬品と医療用具にも適用される。かかる製品は別のガイダンスで取り扱う。

めに必要なデータの量や種類に少なからず影響を与えた。医学が進歩し、病因論と病期についての理解が深まるにつれ、医薬品の臨床研究の範囲が狭められ、例えば、より具体的な病期に、また臨床的に明瞭な対象小集団に、焦点が絞り込まれる傾向がますます強くなってきた。その結果として、製品の適応症は狭められることがしばしばおこるようになった。一方、考えられる適応症の裾野はより広がった。そして密接に関連した複数の適応症を対象とした多数の研究から、新しい用途についての有効性判定に影響を及ぼすデータが入手できるようになった。同じように、対象集団の異なる研究、ひとつの医薬品だけまたは併用する場合の研究、そして異なる用量・剤形の研究と、そのどれもが医薬品の特定の新しい用途を支持するものとなった。同時に、臨床評価と臨床薬理学が進歩した結果として、臨床効果試験 (clinical efficacy trial) のデザインと実施が一段と厳密になり、複数の臨床施設で試験を実施するのが普通になった。厳密さが増し、範囲が広がったことは、ひとつの研究の信頼性 (reliability)、一般化可能性 (generalizability)、有効性 (effectiveness) を実証 (substantiate) する能力という面で大きな意味がある。

FDA は、かかる進展を踏まえ、医薬品と生物学的製剤の有効性の実証の量的・質的基準について、同局がいかにか考えているかを明確に表明した方がよいと判断した。FDA は、スポンサーが、本ガイダンスによって、有効性の確立に十分な医薬品開発プログラムを立案できるようになることを願っている。ただし範囲は広げすぎべきでない。本ガイダンスはまた、医薬品の有効性を支持するために必要な臨床試験データを FDA が評価する際に、その評価をより一貫性あるものにし、予測性を強化するはずである。

本ガイダンスには今ひとつの主要なゴールがある。すなわち、承認済み医薬品のラベリング (labeling) に新用途を付け加える一変申請書の提

出を奨励することである。FDA は、ひとつの医薬品の新用途承認を裏付けるために必要なエビデンスの量と質について、現在いかに考えているかを明確化することによって、新用途についての一変申請書の提出が煩雑になりすぎないように願っている。

II. 有効性を支持するのに必要なエビデンスの量

A. 医薬品と生物学的製剤の法的基準

医薬品：医薬品承認の有効性要件は、1962年に連邦食品医薬品化粧品法 (Federal Food, Drug, and Cosmetic Act, 以下、「同法」または「FDC法」という) に付け加えられた。1938年の同法成立から1962年の修正までの間、医薬品製造者はその医薬品が安全であることを示ささえすればよかった。有効性要件を追加しようという動きにまずはずみをつけたのは、医薬品の価格が高い上に、製薬会社の表示 (claim) が誤りを招きやすいものであり、適正に支持されたものでないという議会の懸念であった。2年間にわたって公聴会が開かれた後、議会は1962年に修正法を採択し、この修正によって、医薬品製造者に「実質的なエビデンス」 (substantial evidence) によりその有効性の証明を義務づける規定を盛り込んだ。「実質的なエビデンス」とは、同法セクション 505(d) に「当該医薬品の有効性評価のための科学的訓練および経験を経て適切な能力を有する専門家による、臨床研究を含む「適切でよくコントロールされた研究」 (adequate and well-controlled studies) から成るエビデンスで、当該医薬品が、ラベリングまたはラベリング案に基づき処方、推薦、提案された条件のもとで、所期の効果 (effect) を有しまたは有することが公正かつ責任を以て結論することができるもの」と定義されている。

この規定を追加した1962年の修正以来、有効性証明のために必要なエビデンスの量と質をめぐっ

て議論が重ねられてきた。量の面については、議会在有効性証明のために、それぞれ独立してそれだけで、有効性を証明できる「適切でよくコントロールされた研究」を2つ以上要求しているというのが、これまでのFDAの見解であった(例えば、Benylin についての最終的決定 44FR51512, 518(1979年8月31日), Warner Lambert Co. 対 Heckler, 787F. 2d 147(3d Cir. 1986)参照)。FDAのこの見解は、同法の文言⁽¹⁴⁾と1962年修正が成立した経緯に基づくものである。上院報告書の文言は、「適切でよくコントロールされた研究」という語句の狙いが、必要なデータの質だけでなく、必要なエビデンスの「量」(quantum)を説明することにあると暗示している(上院報告書 No.1744, パート2, 第87議会, 第2会期6, 1962年)。

にもかかわらず、FDAは、議会の枠組み(scheme)が課した限度内で弾力的な姿勢をとり、個々の医薬品のデータが納得のいくものである限り可能な範囲内で法の要件を幅広く解釈してきた。FDAは、ある場合には、新しい用途についての有効性を示す唯一の「適切でよくコントロールされた研究」を支持するために、異なる用量・療法(regiment)・剤形、異なる病期、異なるポピュレーション(population, 患者集団)、異なるエンドポイントといった他の「適切でよくコントロールされた研究」からの関連する情報を活用している。そのような場合には、厳密な意味での新用途の研究はひとつだけしかないにもかかわらず、新用途を裏付ける研究が実際には多数あるということになり、専門家は、それらをひとまとめにして、有効性の実質的エビデンスがあると結論することができる。また、FDAは、唯一の「適切でよくコントロールされた研究」だけで承認を認めることもあったが、そういった場合は、一般には、その単一の研究が、優れたデザインの多施設研究であり、

生存率への影響など重要な臨床的利益(clinical benefit)について、信頼性が高く、統計学的に強いエビデンスがあったり、確認のための研究が倫理的な理由から難しい場合に限られていた。

議会は、近代化法セクション115(a)の中で、同法セクション505(d)を修正し、FDAが有効性の確認に十分であると判断しさえすれば、「一つの適切でよくコントロールされた臨床研究データとそれを確認するためのエビデンス」によって実質的エビデンスとみなすことができることを明確にした。この明確化にあたって、議会は、承認の法的要件についてのFDAの解釈を追認しており、医薬品開発についての科学の進歩が臨床試験データの質を高めているとのFDAの見方を認めている。

生物学的製剤：生物学的製剤は、公衆衛生サービス法(PHS法)セクション351(42U.S.C.262)の下で承認される。1944年以来有効であるセクション351のもとで、生物学的製剤のライセンスは、製品の「継続的安全性、純度、力価」(continued safety, purity, and potency)を保証するための規格に適合していることを示しさえすれば、発行されてきた。「力価」(potency)とは、有効性を含むものと長く解釈されてきた(21CFR600.3(s))。1972年に、FDAは、以前にライセンスを付与した生物学的製剤すべての安全性と有効性に関するレビューを開始した。FDAは、有効性のエビデンスとは、新薬について「適切でよくコントロールされた研究」(21CFR314.126)と定義されているように「コントロールされた臨床研究」であると述べている。しかし、生物学的製剤には適用できないとか、有効性を実証するには別の方法が適切で試験の妥当性に不可欠であるとして適用を控える場合はこの限りではない(21CFR601.25(d)(2))。例え

注4：同法セクション505(d)は、「実質的エビデンス」を「臨床研究を含む適切でよくコントロールされた研究」(adequate and well controlled investigations, including clinical investigations)と定義するにあたって、複数形を用いている。「研究」(investigation)については、また、新薬申請書の内容を列挙している同法セクション505(b)をも参照されたい。

ば、臨床上の有効性と相関する血清学的反応データがよい例である。非生物学的医薬品の場合と同じように、FDA は、確固とした結果の出ている単一の多施設研究をもとに、生物学的製剤を承認してきた。

近代化法セクション 123(a)が、PHS 法セクション 351 を修正して、生物学的製剤とその製剤が製造される事業所について別個のライセンスを取得する必要がないことを明確化したものの、生物学的製剤の有効性のエビデンスとなる基準は変わっていない。すなわち、製剤が「安全で、純度よく、力価を有するもの」(safe, pure, and potent)であることが示されなければならない(修正 PHS 法セクション 351(a)(2))。議会はまた、近代化法の中で(セクション 123(f)), PHS 法セクション 351 のもとで BLA の承認を受ける必要がある製剤、および FDC 法セクション 505(b)(1) のもとで NDA の承認を受ける必要がある製剤の「レビューと承認における差異を最小限にする」(minimize differences in the review and approval)措置を講じることを、FDA に要求している。

B. 法的基準の科学的ベース

複数の「適切でよくコントロールされた研究」が通常必要であるということは、実験結果のそれぞれについて「独立した実証」(independent substantiation)が必要であることを意味する。効価に関する単一の臨床試験は、他の独立したエビデンスの支援がなければ、通常は有効性ありと結論を適切に科学的に支持するものではないと考えられてきた。その理由のいくつかをあげれば、次の通りである。

- 臨床試験には必ず、予期できず、検出できない、システマティックなバイアスがつきまとう。スポンサーや研究者にいかにか誠意があっても、かかるバイアスが作用して、欠陥のあ

る結論につながる恐れがある。さらに、研究者によっては、意識して評価にバイアスを持ち込む恐れがある。

- 生物系特有の多様性に基づき、単なる偶然によってプラスの試験結果を生む恐れがある。この可能性は、単一の効果試験の結果についての統計学的な評価においてはある程度認知し、数値化することができる。しかし、FDA に肯定的な結果を提出することを目的として、毎年数百のランダム化臨床試験が実施されていることをも忘れてはならない。試験された医薬品がすべて無効の場合でも、試験が 40 件実施されれば、そのうちのひとつくらいは、単なる偶然から通常の統計学的有意性⁽⁵⁾を以って効果を「示す」(demonstrate)こともある。したがって、偽のポジティブ(false positive)な所見が出て(すなわち、有効性のない医薬品に効果があるとの結果が偶然出たということ)、有効性のエビデンスとして FDA に提出される恐れがある。それぞれ独立した研究によって、正の(favorable)結果を実証することとすれば、単一の研究で起こるような誤った結論になる偶然の可能性を防止することができる。
- ひとつのセンターで得られる結果は、施設や研究者ごとの特殊な要因(例えば、病気の定義、併用療法、食事など)に依存する。そのような場合には、結果はたとえ正しくとも、意図した患者群(population)について一般化可能とすることはできない。この可能性こそ、実証的研究がそれぞれに独立していなければならないことが強調される最大の理由である。
- 稀ではあるが、正の結果が科学的不正行為(scientific fraud)の産物の場合がある。

これらの問題に対処するためには、統計学的、

注 5 : p 値=0.05, 両側の意味するところは、片側の有効率(偽陽性)の誤差率が 0.025, すなわち 40 回に 1 回ということである。

方法論的,あるいはその他の安全弁があるものの, 往々にして, 単一の試験だけでこの問題に取り組むのは不適切であることが多い. 複数の実験結果について独立した実証を行うことにより研究間の一貫性(consistency)をもたらし, 薬剤の有効性について誤った結論に導くバイアス, 偶然, 施設要因, 不正といった要因を格段に少なくするのである.

独立した実証の必要性は, 知見の反復(replication)の必要性とされる場合が多い. しかし, 別の研究者が, 別の患者を対象として同じ実験を正確に繰り返すのが, 結論を実証する唯一の手段であると思われるので, 「反復」というのは, 最良の用語ではないであろう. 臨床知見の独立した実証には多くの方法が考えられ, 試験の精確(precise)な反復はそのひとつにすぎない. 研究デザインにつきもののシステムティックバイアスによって, 結論が左右されることもあり, 時としては, 最適とは言い切れない場合もある. デザインが異なり, それぞれ独立して実施される研究, 恐らく対象となるポピュレーション, エンドポイント, 剤形も異なっていると思われるが, その結果から得られる結果は, 同じ研究を反復した場合と少なくとも同じ程度に納得できる, あるいはもっと説得力のあるエビデンスとなりうる.

C. 有効性を支持するエビデンスの量

以下の3節は, 特定の状況下における有効性についての実質的エビデンスを確立するためのエビデンスの量についてガイダンスを与えるものである. 第1節は, 新用法の有効性を既存の効果試験によって全面的に外挿する場合である. 第2節は, 特定の新用法についての単一の「適切でよくコントロールされた研究」を, 異なる病期, 類似した疾患, 異なる使用条件(異なる用量, 使用期間, 療法), 異なる投与量, 異なるエンドポイントの研究など, 別の関連した「適切でよくコントロールされた研究」情報によって支持, 裏付ける場合であ

る. 第3節は, 単一のマルチセンターによる試験にて, 他の「適切でよくコントロールされた研究」による情報の支持, 裏付けなしに, その使用が有効であるとのエビデンスとなしうる場合である.

どの場合においても, 有効性を支持, 裏付ける研究はいずれも, 21CFR314.126の「適切でよくコントロールされた研究」の要件に適合するものと仮定する. また, 所定の用法について単一の研究に依拠することは, 単一であろうと, 関連試験データによる援用であろうと, 研究が不完全である恐れ, あるいは情報に矛盾がある(裏付けとならない)恐れが残ることを, 正しく認識しなければならない. いずれにせよ, その単一の研究が適切にデザインされていること, ベースラインの不均衡, 非盲検, 解析法の事後変化その他の要因に起因するバイアスの可能性が最低限であると判断できること, およびプロトコルに記録された明確な事前仮説を結果が反映していることが前提となる. さらに, 類似した数回の試みにおいて有効性が裏付けられないのに, ひとつの研究でプラスの結果が出たとしても, 有効性が認められなかった研究結果を度外視してよいという強力な理由がない限り, その製品の用途について説得力あるエビデンスとはならない(例えば, 研究が明らかに不適切あるいは分析感度(assay sensitivity)が不足していて, 3群比較で証明されたような既知の活性物質の有効性が示せない場合).

有効性の判断を行うにあたって, 単一の研究に頼るか否かは, 今日の医薬品開発ではあまり問題とはならない. 医薬品開発の場では, 適切な用量を見いだすこと, 病気の複雑さや重症度の異なる患者を研究すること, その医薬品を他の療法と比較すること, 安全性目的で適切な数の患者を研究すること, その他, 上市前に知るべきことを知る必要があり, その結果として, 有効性の決定の基礎とすべく, 複数の「適切でよくコントロールされた研究」を実施することとなる.

本ガイダンスの意図は, 既存の効果データがそ

それぞれの表示について独立した実証となるような状況を残らず列挙することにあるのではない。むしろ、どのような推論法(reasoning)を用いることができるか、その例を示すことにある。かかる例は、NDA または BLA 申請の際の表示か、一変申請の際の表示かにかかわらず、あてはまるものである。

1. 既存の研究からの外挿

承認済み医薬品についての新しい適応症に対する有効性、または新医薬品の有効性は、場合によっては、新たに「適切でよくコントロールされた臨床試験」なしに適切に実証することができる。それは、通常、別の種類のデータによって、既知の有効性を新しいポピュレーションまたは異なる用量、療法、剤形などにあてはめて行い、以下は、別の薬効表示または製品について、効果データから有効性を外挿できる場合の例である。

a. 小児科での使用

製品表示(product labeling)について小児科領域部分を改定した規則(21CFR201.57(f)(9)(iv))は、小児における使用についてコントロールされた臨床試験がなくても表示に小児科用情報を含めることを認めている。そのような場合には、スポンサーは、小児科用途を支持する情報を提示しなければならず、病気の経過とその医薬品の効果が、小児と成人において十分に似ており、成人の効果データを小児患者に外挿できるとの判断をFDAが下すことが前提となる。成人・小児ポピュレーションにおいて病気の経過と医薬品の効果がともに似ているとの結論を裏付けることができると思われるエビデンスとしては、成人・小児ポピュレーションの病態生理学および病気の自然経過が共通であることを示すエビデンス、各ポピュレーションにおける共通の薬物代謝および類似の濃度-反応関係のエビデンス、当該疾病・病状または関連疾病・病状に用いた場合のその医薬品または同じ治療効果を有する他の医薬品に関する経験があげ

られる。従来から、小児用の表示に関する情報が成人の効果データから外挿されてきた例としては、疼痛用のイブプロフェン、季節的アレルギー性鼻炎用のロラチジンがあげられる。

b. 生物学的同等性

代替処方と新用量力価の有効性は、生物学的同等性のエビデンスを基礎に評価される。

c. 修正放出剤形

修正放出剤形(modified release dosage)は、以前に研究された速放出剤形に新しい剤形を結びつける薬物動態データを基礎に承認される場合がある。修正放出剤形と速放出剤形の薬物動態パターンが同一でないため、速放出剤形データを修正放出剤形にあてはめるためには、時間経過による変動を含む血中濃度と反応性の関係にある程度理解することが一般に重要である。

d. 異なる用量、療法または剤形

用量-反応関係は一般に連続的であり、ひとつの用量、療法または剤形の有効性についての情報は、他の用量、療法または剤形と密接に関連している。血中レベルと曝露(exposure)が大して違わない場合には、薬物動態データだけに基つき、新しい用量、療法、剤形も有効であると結論することが可能である。また、たとえ血中レベルがまったく異なっても、血中濃度と反応の関係が、時間の経過にしたがってどのように変化するかを含み、十分に理解されていれば、新たな臨床効果試験がなくても、薬物動態データをもとに、新しい用量、療法、剤形が有効であるとの結論を下すこともできる。このような状況においては、十分に規定された薬物動態/薬力学(PK/PD)とともに薬物動態データが、ひとつの用量、療法、剤形から別の新しい用量、療法、剤形へ、コントロールされた臨床試験を解釈する場合に用いることができる(第II節C.2.aも参照されたい)。

2. 新用途についての単一の研究による有効性の証明：関連研究データからの独立した実証

以下では、新用途についての単一の研究が、関連する用途についての研究データからの独立した援用によって、有効性のエビデンスとする例を考察する。このような事例においては、新用途の研究と関連研究によって、当該医薬品が所期の効果を有すると結論することができる。関連研究が新用途についての単一の研究を実証する能力があるか否かは判断の問題であり、研究の質と成果、そして新用途との関連の程度如何にかかわっている。

a. 異なる用量、療法または剤形

第II節c.1.dで検討した通り、血中レベルと曝露が大して異ならない場合には、新たな臨床効果試験を行うことなく、薬物動態データをもとに、新しい用量、療法、剤形が有効であるとの結論を下すことができる。また、たとえ血中レベルと曝露がまったく異なっているとしても、血中濃度と反応の関係が十分に理解されていれば、やはり可能である。血中濃度と反応の関係があまりよく理解されておらず、新用量、療法、剤形の薬物動態が以前のものとは異なる場合には、新しい療法の有効性の裏付けとしての臨床効果データが必要になる可能性が大きい。その場合、通常は、単一の追加効果研究で十分なはずである。例えば、リスペリドンの1日1回の投与が最近承認された際には、その裏付けに単一のよくコントロールされた試験が必要とされた。というのは、1日1回の療法と1日2回の療法の薬物動態が異なり、リスペリドンのPK/PD関係が十分に理解されていなかったためである。

b. 他の病期の研究

多くの場合、効果の大きさと効果対リスクの関係は、病期(phase)が異なれば異なるけれども、ひとつの病期に有効な療法は、他の病期にも有効である。例えば、ある医薬品が、特定の癌の不応期

の患者に有効であることが分かっているとすれば、同じ腫瘍のもっと早い段階についても、通常は、単一の「適切でよくコントロールされた研究」だけで、新用途を裏付ける十分な有効性のエビデンスとなる。

c. 他のポピュレーションの研究

特定のポピュレーションのサブセットにおける反応は、ポピュレーション全体と質的に同一であることが多い。人口統計的に抜き出されたサブセットにおける有効性については、ほぼ例外なく、別の研究の必要がない(第II節c.1.aの小児科ポピュレーションの研究も参照されたい)。ただし、新たな研究が必要な場合もある。ある条件のもとで一般的に有効であったり、ひとつのポピュレーションで有効であることがすでに分かっている医薬品については、年齢、人種、性別、合併症その他のサブセットにおける有効性を裏付けるには、普通ならば、単一の研究で十分である。例えば、男性の乳癌にタモキシフェンを用いることを裏付けるには、単一の研究だけで十分である。

d. 併用または単一療法の研究

単一療法(monotherapy)で有効と分かっている医薬品については、他の療法と組み合わせ(多剤併用の一部としてまたは固定用量併用法として)その医薬品の有効性を証明するには、通常は、単一の「適切でよくコントロールされた研究」だけで十分である。同じように、併用療法の一部として医薬品の有効性が分かっている場合(すなわち、併用への効果が分かっている場合)において、モノセラピーとしてのその用途または同一用途について異なる組み合わせで用いる場合は、通常は、適切にデザインされた単一の研究だけで実証できると思われる。例えば、免疫応答の実証を狙った新しい組み合わせワクチンの場合、単独もしくは組み合わせにおいて有効であることがすでに証明されている製品もしくは抗原が含まれていさえすれば、単一の研究だけで有効性のエビデンスとしては十分である。このような状況は腫瘍薬や

抗高血圧薬によく見受けられるが、他の場合にもまたありうる。

e. 密接に関連した病気の研究

病因論的または病態生理学的に関連した病状の研究、もしくは複数の病気に共通の症状(例えば、疼痛)の研究は、互いに支持する場合があって、いくつもの使用が当初から承認されており、単一の「適切でよくコントロールされた研究」に基づき新たな表示追加される場合がある。例えば、ある抗凝固薬または抗血小板療法は、不安定アンギナ/急性心臓冠状静脈症候群および血管形成術後症状についての個別研究に基づき、2つの異なる環境での使用が承認されることが考えられる。研究のエンドポイントと抗凝固薬または抗血小板剤使用の理論的基盤が類似しているため、各研究が各有効性表示につき互いに支持するのである。同じように、複数の疼痛症状における単一の鎮痛作用の研究は、一般的な鎮痛の適応症または多数の具体的な適応症のいずれを支持するにも十分である。最近ラモトリジンがLennox-Gastaut症候群(稀であり、大抵は小児に発生する全身化癇攣障害)治療について、単一の「適切でよくコントロールされた研究」に基づいて承認されたが、これは、ひとつには、部分的に始まる成人の癇攣発作にこの医薬品が効くことを示す関連データがあったためであった。

f. 疾患は密接に関連していないが、治療法の一般的目的が類似している研究

抗菌薬や抗腫瘍薬など、ある種の薬物療法は、異なる疾患にまたがる適切な介入となる。この種の治療法では、ひとつの病気における有効性のエビデンスが、まったく異なる病気における有効性の独立した実証となることがある。例えば、ある感染症における抗菌薬の有効性を示すエビデンスを以って、病原体、病気の進行に影響を及ぼす感

染部位の特徴(例えば、構造、免疫学)、患者群と似たものが類似している場合には、単一の研究であってもそれを支持し、有効性を証明しようと主張することが可能である⁽¹⁰⁾。同じように、腫瘍薬も、ひとつ以上の種類の腫瘍において有効であるとのエビデンスがあれば、異なる種類の腫瘍、とくに発生学的に同系統の腫瘍型のものであるならば、単一の研究だけを根拠に有効性を判断してよいと考えられる。

g. 異なる臨床エンドポイントの研究

臨床的意味のある2つの異なるエンドポイントについての異なる研究で、有益な結果が実証されれば、それぞれの結果の有効性表示(claim)は相互に実証される。例えば、心不全の治療薬であるエナラプリルの当初の表示は、数ヵ月にわたる症状改善を示したひとつの研究と、もっと重症のポピュレーションにおいて生存率の改善を示した第二の研究によって裏付けられた。それぞれ「適切でよくコントロールされた研究」から来た2つの異なる知見が、症状改善と生存率改善の両方に有効であるとの結論につながったのである。

h. 薬理的/病態生理学的エンドポイント

病気の病態生理学と治療法の作用メカニズムが非常によく理解されているときには、特定の薬理学的効果と臨床的な有効性を関連づけることができる。バリデーションの済んだ代替エンドポイント(surrogate endpoint)として認められる薬理学的効果は、通常の承認を支持する(例えば、血圧に対する効果、コレステロール低下作用)。また、臨床上的有用性を合理的に予測する可能性ある薬理学的効果は、21CFR314 サブパートHおよび21CFR601 サブパートEに規定されている条件のもとで、迅速承認の支持となる場合がある(例えば、CD4 カウントおよびウィルス負荷効果はHIV感染に対する抗ウィルス薬の有効性を支持する)。

注6：「抗感染症医薬品の部：抗感染症医薬品の臨床開発とラベリングにおける留意点」(Division of Anti-Infective Drug Products: Points to Consider in the Clinical Development and Labeling of Anti-Infective Drug Product, October 1992)を参照。

薬理効果が、有効性エンドポイントとは認められないが、臨床結果とのつながりが濃厚である場合(理論的基礎からだけでなく、以前の治療的経験やよく理解されている病態生理学をもとに考えて)には、臨床効果を示す単一の「適切でよくコントロールされた研究」を、関連薬理効果を示す「適正でよくコントロールされた研究」における説得力のあるデータによって実証することができる。

例えば、凝固因子などの入れ替え療法の承認を裏付けるためには、明らかにプラスの結果が出た単一の試験だけで十分な場合があり、その場合には、治療対象の病状がその要因の欠如によって引き起こされたことを示す明確なエビデンスと組み合わせられる。欠如因子の物理的代替、または消失した生理作用の回復を実証すれば、臨床効果の強力な実証となる。先天性代謝異常の是正治療の場合も、同じように見ることができる。予防ワクチンの場合には、「適切でよくコントロールされた試験」を、強制動物攻撃/保護モデル(compelling animal challenge/protection model)、ヒト血清学データ、受動抗体データ(passive antibody data)、または病因情報によって裏付けることができる。薬理的エンドポイントを病気の改善または予防に結び付けるエビデンスが多ければ多いほど、単一の臨床効果研究だけでよいという主張は説得力あるものになる。しかしながら、有益な薬理効果がありそうでも、臨床効果との相関関係がない場合も多く、有効性のエビデンスに薬理効果を利用する場合は、常に慎重でなければならない。

例えば、1型抗不整脈剤の不整脈抑制効果や、ホスホジエステラーゼ抑制薬やベータアドレナリン作用性筋変力剤による心拍出量増強などの薬理効果は、期待されたように、突然死を減らしたり、心不全において結果を改善するよりもむしろ、死亡率の増大につながる。薬理効果と臨床効果の間に、期待されたような相関関係がない理由は多様である。例えば、薬理効果と臨床効果の関係が十

分に理解されておらず、測定され有益と考えられた効果の他にも、その薬物に帰因するその他の薬理効果があるからである。一般的に、独立した実証を行うにあたって、薬理的結果の効用(utility)が最大となるのは、当該薬理分野においてすでに経験がある場合である。しかしながら、かかる場合でさえ、臨床効果と相関関係がある薬理効果が、すべての臨床上有益であると確信することは難しい。往々にして、他の効果があって、当面の問題に密接に関連している可能性を否定できないのである。

3. 単一の研究からの有効性のエビデンス

有効性が要件として最初に求められるようになったのは1962年であるが、その頃の効果研究モデルは、ひとつの施設、一人の研究者によるもので、盲検手続についても比較的ゆるやかな小規模な試験が普通で、プロスペクティブな研究デザインとか、結果や分析についての検証についてあまり配慮はされていなかった。現在では、主たる臨床効果研究は大抵が多施設(マルチセンター)による研究であり、かかる研究には、プロスペクティブにデザインされた明確な臨床・統計分析基準がある。このような研究は、バイアスの影響を受け難く、一般に普遍化しやすく、非常に説得力ある統計結果を導き、と同時に、評価にあたり、サブグループ、施設、多重エンドポイントにまたがって内的一貫性(internal consistency)を追求できる場合が多い。

今日においては、臨床試験の厳密さと規模が増したために、一定の状況下においては、別の比較試験による独立した実証を行わなくとも、唯一の「適切でよくコントロールされた研究」だけで、承認のための十分な科学的・法的根拠とすることができるようになった。例えば、チモロールの梗塞形成後の死亡率低下についての承認は、とくに説得力がある(p値が低い)、内的に一貫した唯一のマルチセンター研究に基づいて、死亡率と再発

率に対する大きな効果が実証された。この研究は、倫理的な理由から、反復実施できないと考えられた。生物製剤評価研究センター(CBER)もまた、単一の説得力ある研究をもとに多数の製品を承認している。FDAは、1995年に、単一のマルチセンター研究がどのような場合に十分であるかについての一般声明を提示した(60FR39181, 1995年8月1日)が、単一の「適切でよくコントロールされた研究」が有効性表示の適切な裏付けとなる状況については包括的に説明していない。

また、単一の研究にどのような特徴があれば、有効性表示の適切な裏付けとなるかについての総合的な説明もない。従って、単一の「適切でよくコントロールされた研究」に頼ることができるか否かということは、必然的に判断(judgment)の問題となる。2つの説得力ある研究をもとにした結論は常に、同様な単一の研究をもとにした結論よりも確実である。この理由から、単一の研究のみに頼ってもよい場合は一般に、死亡率、不可逆的罹病率(irreversible morbidity)など臨床的に意味のある効果、あるいは重大な結果につながる恐れがある病気の予防、さらには第二の研究による確認が技術的あるいは倫理的に不可能であるような場合である。例えば、梗塞後死亡率の減少、骨粗しょう症骨折の予防、百日咳の予防などを実証する試験を連続的に反復することは、いくら確固たるプラスの効果が出るとしても、重大な倫理的問題を提起する。症状に対して効果があることだけを証明する試験の場合には、反復しても、倫理的問題とはならないのが普通である。

以下の考察では、「適切でよくコントロールされた研究」について、どのような特徴があれば有効性表示の裏付けになるか、その特徴を取り上げる。かかる特徴はいずれも、必ずしも決定なものではなく、ひとつの研究の中にかかる特徴がひとつ以上あれば、その研究が有効性表示を適切に支持するとの結論に寄与する。

a. 大規模マルチセンター研究

(1)ひとつの研究施設に異常に大きな患者の塊がない、(2)肯定的結果に異常に責任を有している研究者または施設がない、といった場合の大きなマルチセンター研究においては、研究の内的一貫性のために、知見の普遍化可能性やひとりの研究者の行為のみに帰因する説明不能な結果を心配する必要が少なくなる。分析の結果、ひとつの施設が効果について大きく寄与していることが分かった場合は、マルチセンター研究の信頼性は薄らぐ。

b. 研究対象のサブセットにまたがる一貫性

大規模共同研究の場合、被験者の参加基準が比較的幅広く、研究ポピュレーションが同時期または以前の治療、病期、年齢、性別、人種などの重要な共変量(covariant)の面で、多様である場合が多い。そのような試験の結果についての解析は、キーとなる患者のサブセットにまたがる一貫性により、多様なポピュレーションに対する知見の普遍化につきまとう懸念を多少なりとも解消する(小規模試験や患者組入れ基準がもっと狭い試験では、そうはいかない)。例えば、チモロールの梗塞後研究は、3つの重症度層内でそれぞれ、患者を別々にランダム化した。この研究は、各層の生存にプラスの効果があることを示し、この薬の効用が特定の病期(例えば、比較的重症の病期や軽症の病期)に限られないとの結論を裏付けた。

c. 単一の研究内の多重研究

適正にデザインされた要因研究(factorial studies)は、対になる一連の比較研究として解析することができる。すなわち、単一の研究の中にあっても、モノセラピーあるいは別の医薬品と組み合わせることで当該医薬品の作用を別々に実証することができる。このモデルは、ISIS IIで用いられ、成功を収め、アスピリンとストレプトキナーゼの両方が、単独の場合と併用の場合ともに、心筋梗塞患者の生存に有利な効果を及ぼすことを示した(アスピリン単独およびストレプトキナーゼ単独

の場合はそれぞれプラセボより勝っていた。アスピリンとストレプトキナーゼの併用の場合は、アスピリンとストレプトキナーゼそれぞれ単独の場合より勝っていた。これはすなわち、アスピリンとストレプトキナーゼの有効性についての別個の実証(ただし、完全に独立してはいない)と考えられる。

d. 異なる事象に関連した多重エンドポイント

単一の研究が、プロスペクティブに決められた複数の重要な一次、二次エンドポイントを含み、そのそれぞれが、有益であるが異なる効果を表す場合がある。統計的に説得力のあるエビデンスにより、そのようなエンドポイントの少なくともひとつに効果があることが示された場合には、その研究の示したエビデンスの意味は内的に重くなる。例えば、多発性硬化症の病状再燃の予防薬であるベータ・インターフェロン(Betaseron)の承認は、単一のマルチセンター研究に基づいていた。その理由は、少なくともひとつには、病状再燃の減少率とMRIによる病状の緩和があり、これらは、まったく異なるが、論理的に関連した2つのエンドポイントである。

同じように、脂質減少、血管形成術後、または梗塞後研究において死亡および非致死的心筋梗塞の両方に有利な効果のあることが分かれば、それは別個ではあるが一貫性ある有効性の実証ということになり、死亡率減少という知見が偶然の産物である可能性は非常に少なくなる。例えば、複雑な血管形成術またはアテレクトミーを受けている患者の補助療法としてのabciximabの承認は、単一の研究を基礎としていた。この研究では、組み合わせエンドポイント(死亡数、新たな梗塞、および緊急介入の必要の減少の合計症例数)についての全般的結果が確固としており、組み合わせエンドポイントのうちの2つの構成要素(新たな梗塞の減少、緊急介入の必要の減少)を別々に評価して、統計的に意味のある効果が出た。それとは対照的に、梗塞後に2つの異なる鬱病尺度または

SGOT・CPKレベルで測定された気分の変化など、本質的には同じ現象の評価であり、かかる現象相互の強力な関係付けである多重エンドポイントにプラスの効果があっても、単一の研究からのエビデンスの内的意味はあまり重くはならない。

単一の研究内における2つの一貫性ある知見は通常治療効果が偶然のものではないと考えられるが、研究行為自体におけるバイアスや解析上のバイアスを防ぐことはできない。例えば、割付け(treatment assignment)が、重要な予後可変要素の面で十分にバランスがとれていないと、両方のエンドポイントに明らかな効果が出る恐れがある。したがって、研究デザインと実施行為を綿密に調べるか否かが、この型の研究の評価においては、その適否を決する。

e. 統計学的に極めて説得力ある所見

マルチセンター研究では、p値が非常に低いと、治療効果がないというゼロ仮説と大きく喰い違ふこととなる。研究によっては、複数の施設のデータから、わずかでも統計学的意味のある結果があれば、それを検出することができる。たとえそれができない場合でも、極端な結果や意味レベルが出れば、それは全般的なものであり、多施設共同研究であるからには、大半の研究施設の知見が同じようなのである。例えば、ストレプトキナーゼの血栓崩壊試験(ISISII, GISSI)の治療効果は非常に大きく、p値が非常に低かった。そのために、その説得力(persuasiveness)がいよいよ増したのである。効果率が高い感染症予防ワクチン(例えば、効果の推定値が80%以上であり、95%の信頼区間が合理的な狭さである)は、単一の「適切でよくコントロールされた試験」を基礎に承認された。

4. 単一のマルチセンター研究への依存 — 警告

単一の内的一貫性ある確固たるマルチセンター研究に説得力があることは、なるほど確かであるが、その確固たる結果でさえ、孤立したまたはバ

イアスのあるものである恐れがあることは、正しく認識しなければならない。とくに、その研究が、類似の研究の中で、効果を示唆する唯一の研究である場合には、注意が必要である。最近、筋変力剤である vesnarinone の心不全に対する一見非常に有利な結果(非常に優れたデザインのプラセボ対照マルチセンター研究と思われるもの-p値が極端で、死亡率の60%低下)が、反復不能であることが分かった。この所見を実証しようという試みの中で、前の研究で生命を救うように思われた用量が、死亡率を大幅に高め(26パーセント)、それより少ない用量でも、生存に有害な作用を及ぼすようであった。第二の研究のポピュレーションの病気が、全体として、第一の研究よりも重かったものの、各研究中の同じような重症度の患者における結果は矛盾していた。従って、この要因が結果の矛盾の原因ではないと思われる。

単一のマルチセンター研究に依存するか否かを検討するにあたって、結果が正しくない可能性を考慮すること、および入手できるデータを残らず吟味し、単一のマルチセンター研究に依存できることを示すか否かを熟考することが、非常に重要である。vesnarinone の場合には、マルチセンター研究の驚くべき好結果と矛盾する他のデータがあった。このデータは、逆の用量-反応関係を示すように見え、症状面の効果を示唆するものではなく、血行力学的エンドポイントへの影響は示さなかった。かかる矛盾があったため、FDAは、心腎諮問委員会の助言を得たうえで承認を拒否した一以後の研究結果から得られた決定であった。

この例は、薬理学的合理性がない、重大な結果を伴う所期の効果が他にないなど、データが不適切で矛盾していれば、単一の研究の説得力がいかにか低くなるかを示している。第二の比較試験は、説明のつかない原因から第一の有利な研究結果を実証できなかったのであるが、そのことは第一の有利な研究が誤りであったことの証明とはならない一有効な薬剤の研究が、多様な原因からその薬

剤に効果があると示さないこともある一このことこそ、多くの場合、単一の有利な研究だけを依拠とはしない理由である。

III. 有効性の表示を支持するのに必要なエビデンスの質に関するドキュメント

新製品または既承認の製品の new use (new use) についての承認を支持するために必要とするデータを提出する際に、スポンサー(試験依頼者)は試験が適切にデザインされたことを証明する文書を添付しなくてはならない。「適切でよくコントロールされた試験」の基本的性格については、21CFR 314.126を参照されたい。有効性を裏付けるための試験が適切でよくコントロールされたものであることを証明するために、通常、試験の計画、プロトコル、実施およびデータ処理に関する広範な文書を当局に提出し、詳細な患者記録が臨床試験の場で得られるようにしておく。

しかし、科学的見地から、必要とされる文書の範囲は、個々の試験、提出データの性格、および表示を支持するその他のエビデンスによって異なるものと思われる。したがって、科学的エビデンスの適切性が保証される限り、当局はさまざまな質的レベルのデータ文書を受理している。本章では、既承認医薬品の new use を評価するための試験に重点を置いて、提出文書の範囲にかかる要因について検討する。

本章の目的に関して、「エビデンスの質をめぐる文書」(documentation of the quality of evidence)の句(phrase)は、(1)文書の完全性、(2)エビデンスとして提出されたデータの証明を目的とした一次試験データおよび原試験関連記録(被験者のカルテ、薬剤説明記録(drug accountability record)などへのアクセス可能性を意味する。相互に関連するこれらの要素は、試験が「適切でよくコントロールされた試験」であるか否かの判定に関わるものである。

高レベルの文書を作成するためには、現実的に、表示を支持するための試験は通常 GCP に則って行われる。スポンサーは臨床試験施設を定期的モニターし、FDA は原プロトコル、一次データ、オンサイト審査 (on-site audits) のための臨床試験施設の基礎文書、および完全な試験報告書に、定期的にアクセスする。

しかし、医薬品の有効性を評価する試験で上述の文書が揃っていなかったり (患者記録が欠けているなど)、商業ベースで依頼された試験で通常行われるよりも不十分なモニタリングで試験が行われる状況がよくある。このような状況は効能追加 (supplemental indication) の場合にはよく見られる。承認後の臨床試験は当該薬品のスポンサー以外の研究グループによって行われることが多く、モニタリング手順やデータ収集手順は、スポンサーのそれよりも範囲が限定されることが多い。特定の状況下では、通常レベルを下回る文書作成や監視体制であっても、スポンサーはそのような試験により有効性を裏付けることができる。つぎにそのような例を示す。

A. 臨床データまたは詳細な研究報告書への通常のアクセスが不十分な場合に 依拠するもの

FDA による一次データへのアクセスが、薬事行政上の決定をするのに重要であることが知られている。また、公表された試験報告の結論に懐疑的となるのも当然のことである。これまでの経験から、試験報告が必ずしも試験の計画、実施および結果を完全かつ正確に伝えるものでないことが明らかになっている。明らかな不正 (outright fraud, 用意周到な虚偽報告など) はまれである。しかし、資料の不備、不明瞭、事前に計画された解析計画からの逸脱の秘匿、またはエンドポイントに関する判断あるいは評価方法に関する説明不足はよく見られる不正である。典型的な例として、学術専門誌の論文審査員は、限られたデータセットや解

析資料しか入手せず、プロトコルやその変更資料には目を通さないため、試験者 (investigator) が評価不可能と断定したポピュレーションの実態を知ることもなく、結局情報不足のため致命的な省略や問題点を探知することができない。専門家による論文審査 (peer review) の有用性は、審査員の研究経歴や専門領域がまちまちであることにより影響を受けている。FDA は、Anturane による再梗塞試験、tacrine および抗敗血症用 HA-1A 抗体の効果の報告を審査した経験から、公表された医学文献への信頼性に懸念を表明する。

このような懸念はあるものの、以下に述べるような要素の存在により、FDA は、有効性を支持するデータにアクセスできなかったり、あるいは詳細な試験報告が部分的または全的に不十分であっても (いわゆる「ペーパーファイル」によって)、有効性の表示を指示するための資料として認めている。有効性を支持するものとしての文献報告の活用は、詳細な重要試験データの追加提出によっていっそう高まる。以下のセクション 1 は、有効性表示を支持するために、追加情報の提出によって、研究への依拠の可能性が高まる場合について記載する。セクション 2 には、公表文献だけで有効性に関する知見を説得できる場合の要件を述べる。セクション 2 において述べる要件は、通常、単独またはセクション 1 で述べる他の重要情報との組み合わせにかかわらず、その文献報告の信頼性評価に関連している点に注意されたい。

1. データの信頼性を高める重要情報とともに、 公表文献や他の報告書を提出する場合

スポンサーが別のグループによる試験の結果に依拠することを希望しているが、その一次データが得られない場合は、適切に行われた試験であれば、試験全体の立案内容を示すプロトコル、試験の実施および手順記録、重要変数についての患者データリスト、統計解析記録など、その他の重要情報を入手することが可能である。FDA は米国お

よびヨーロッパ諸国の政府機関(米国国立衛生研究所 NIH, 英国医学協議会 BMRC)や民間組織 (ISIS 試験, SAVE 試験)をスポンサーとする膨大な多施設試験結果を評価した豊富な経験を持っていて、それに関する一次試験データには限りがあるものの、その他の重要情報は入手できた。試験に関する下記の重要情報のうち、できる限り多くの情報を公表報告に添付して提出すれば、当該試験が有効性表示を支持する可能性を増すことができる。

- a. 当該試験の Protokol ならびに試験実施中に行った重要な計画変更、およびそれと試験拡大とランダム化との関係。
- b. 事前の統計解析計画および試験実施中または試験後に行われた当初計画の変更、ならびに解析を解錠前か後に (pre- and post-unbinding) 行った旨の注記。
- c. 割付け表 (randomization codes) と被験者の報告試験への組入れ年月日に関するドキュメント。
- d. 解析から除外された被験者と投薬データとの照合および除外理由ならびに試験に参加したすべての被験者を用いた結果の解析を含む全被験者の詳細な説明。
- e. 重要な変数および関連ベースライン特性に関する各被験者のデータの電子記録または書面記録。個々の被験者の反応が重要な変数 (癌患者の客観的反応、感染症における臨床的治癒および病原菌除去、特定原因による死亡例など) である場合は、症例報告、病院記録および陳述などの評価に関する詳細な基礎資料を可能な限り提出する。
- f. 安全性が重要テーマである場合は、全死亡例と毒性による脱落例に関する完全な情報。しかし、承認後の追加適応症の場合は、通常、検査試験結果または詳細な有害事象報告の必要性は低く、はるかに限定された文書 (予期されなかった死亡例、それまで報告のなかった

重篤な有害反応など)の提出で足りる。この件に関する適用除外要件として、追加適応症のポピュレーションが既承認適応症の場合と大きく異なり、既存の安全性情報の適用範囲が限定される場合 (脳卒中患者と心筋梗塞患者における血栓溶解作用など)、または新たにポピュレーションとなる患者層に安全上の深刻な問題 (予防接種の範囲を幼児から乳児に拡大する場合など) がある場合をあげることができる。

2. 公表文献報告のみの提出

下記の要素は、新製品または新適応症の承認を支持するにあたって、公表報告書のみに依拠する信頼可能性を高める。

- a. 異なる試験グループによって行われた多数の試験の結果。ただし、個々の試験が適切にデザインされたことが明らかで、試験による知見に一貫性があること。
- b. 統計解析計画、解析方法 (事前に決定された)、試験のエンドポイントの明確かつ適切な記載や、全登録患者の詳細な説明を含む公表報告書詳細
- c. 客観的に評価され、試験者 (investigator) の判断に左右されない明確で適切なエンドポイント (全体死亡率、血圧、病原菌除去など)。かかるエンドポイントは原因別の死亡率や症状の緩解といった主観的なエンドポイントよりも解釈が容易である。
- d. 有効性に関する一貫した結論を導き、共変数の調整などの選択的事後解析、サブセット (subset) 解析、またはデータセットの縮小 (反応患者やコンプライアンスの高い患者のみの分析、または「選択基準にあった」 (eligible) もしくは「評価可能な」 (evaluable) サブセットのみの解析など) を必要としない、試験計画書に明示した解析法に則った頑健 (robust) な解析結果。

e. 適切に文書化された試験手順に従った試験グループによる試験の実施および手順を有効に実施した旨の経時報告。

これまで主としてあるいは公表報告のみを基礎として承認を行った事例がある。膵臓機能評価にかかる secretin の初回承認、悪性肋膜浸出治療における bleomycin と滑石 (talc) およびマラリア治療における doxycycline の最近の承認がその例である。

B. 品質管理/オンサイト・モニタリングが他の方法や十分に行われていない研究への依拠

企業をスポンサーとする試験では、データの質的保証を目的とした広範な実地 (on-site) あるいは中央 (central) のモニタリングと監査 (audit) を典型的に行う。他のスポンサーの支援を受けて行う試験では、手順は厳密さを欠き、オンサイト・モニタリングはまったく行われなこともある。最近国際的に合意された ICH の GCP^(註) では、モニタリングの範囲は、個々の試験に特有な要因 (デザイン、複雑性、規模、試験のアウトカム項目の種類など) にもとづくべきであり、さらにオンサイト・モニタリングにはきまざまなレベルがあつてしかるべきだとの主張が見られる。最近では、信頼性が高くかつ重要な試験、それもしばしば死亡率に関する重要なものが、政府や独立した試験グループによって実施されているが、これらはほと

んどオンサイト・モニタリングが行われていない。それらの試験では、試験文書の密接な管理とレビュー、試験者 (investigator) とともになされる広範なガイダンスや計画といった別な方法でデータの品質管理が行われているのである。医薬品の当初の承認および追加適応症の承認のためにそのような試験の結果に依拠してきた長い歴史がある。モニタリング体制が限定されているかまたはまったくない状況で行われた試験に依拠できるかどうかは、次の要因がかかわっている。

1. データの品質保証のための事前計画が存在すること。
2. バイアスの恐れが本来的に少ない試験であること。例えば、手順が比較的単純で、患者組入れ基準が容易でクリティカル (critical) なものでなく、アウトカムが容易に評価できる試験であること。
3. 重要なデータのサンプリングができ、支持する記録 (病院記録など) との比較が可能であること。
4. 確立した試験実施手順を持ち、その実施手順を有効に用いて試験を行った実績のあるグループによる試験であること。

[本日本語訳の一部は、平成12年度創薬等ヒューマンサイエンス総合研究「難治疾患・稀少疾患を主とした医薬品の適応外使用のエビデンスに関する調査研究」による]

* * *

注7: International Conference on Harmonization Guidance for Industry E6, *Good Clinical Practice: Consolidated Guideline*, April 1996.