

## 小児集団における医薬品の臨床試験に関するガイダンスに関する質疑応答集 (Q &amp; A)

	質問	回答
1 1.4	「成人の疾患や病態に対する医薬品の開発が行われている段階において、当該医薬品が小児集団で使用されると推定される場合には、小児集団を医薬品の開発計画に組み入れるべきである」とは、小児で使用されると想定される医薬品にはすべて適応されるのか。	基本的には、小児で使用されると推定されるすべての医薬品について、開発計画の中に組み入れるべきである。ただし、疾患の重篤性や医薬品の有効性、安全性および医療現場での必要性を勘案し、医薬品開発全般のなかで総合的に判断することとしている。
2 2.1	「小児臨床試験の計画の時期と方法についての正当性は、初期段階及びその後定期的に規制当局との間で明らかにする必要がある」とあるが、規制当局との相談部門はどこか。 また 2.3.2 に該当する医薬品については「初期の安全性データと有益性の可能性を示す合理的な根拠を示した後、開発を早期に始める」とあるが、具体的な相談時期について示されたい。	規制当局との相談は、医薬品機構での治験相談を利用されたい。医薬品の特性により開発時期は異なるが、2.3.2 に該当する医薬品の場合、成人の安全性データが集積された後に小児治験が開始されるため、遅くとも成人での第Ⅱ相試験終了時の治験相談が適切と考えられる。
3 2.3.2	「小児臨床試験の成績は、申請データベースの一部となるべきものである。これが可能でない場合、小児の成績が得られていないことの妥当性を詳細に示すべきである」とあるが、申請資料での記載場所と記載内容を示されたい。	記載場所については特に決められていない。 CTDでは、臨床概括や臨床概要（3.3.3 部分集団における結果の比較）に小児患者集団に関する記載項目が設定されている。
5 2.6.3	インフォームドアセントとは何か。	本ガイダンス及びGCPにあるとおり、小児が臨床試験に参加することの同意（インフォームドコンセント）は法的な保護者から得ることとなっている。しかしこの場合であっても小児の人権を尊重し、被験者の理解力に応じて説明を行なうことが必要であ

		<p>り、さらに適切と考えられる被験者からはアセント文書または同意文書への署名と日付の記入が望まれる。</p> <p>(参考)</p> <p>米國小児科学会によるインフォームドアセントの定義：研究対象者として参加する場合、未成年者が与える積極的な合意。但しコンセントとは同格のものではない。</p>
5 2.6.3	アセントを取得すべき「治験への参加を理解できる知的レベルにある被験者」の年齢を示されたい。	<p>個人の知的成熟度には個人差があり、年齢は特定できない。しかしながら一般的に中学入学以降であれば内容が理解できる年齢と考えられる。</p>
6 2.6.3	Q5に該当しない小児治験者に対する同意説明の考えを示されたい。	<p>「すべての被験者は、彼らが理解できる言葉や用語で臨床試験について可能な限り十分な説明を受ける必要がある」とガイダンス中にも記載されている。</p> <p>概ね7歳以上であれば、簡単な説明に対し理解可能と考えられる。しかしながらそれ以下の年齢であっても、臨床試験に関し被験者が理解できると思われる事項があれば説明すべきである。</p>
7 2.6.3	コンセントとアセントの関係はどのようになっているのか。コンセント取得時であってもアセントが取得できなければ治験は実施できないのか。	<p>代諾者（保護者）からの同意（コンセント）とともに被験者本人からアセントも取得した上で治験を開始すべきである。法的規制を受けない小児被験者からの同意であるアセントを取得する年齢について、米國小児科学会を参考（注1）に、現時点での一応の目安を脚注（注2）に示した。</p>
8	本ガイダンスと市販直後の安全対策改善案や適応外使用に関する通知との関係を示されたい。	<p>小児での有効性、安全性が確立されていない医薬品については、当ガイダンスに基づいた臨床試験を実施し、用法・用量を設定すべきである。ただし既存薬であって適応外使用の通知（適応外使</p>

		<p>用に係る医療用医薬品の取扱いについて（平成11年2月1日付研第4号、医薬審第104号）に記載された条件に合致する場合は、臨床試験を実施せず信頼できる文献等を用いた申請も可能である。</p> <p>また小児での用法・用量を有する医薬品に対し、その承認時に追加情報が必要と判断された場合には、市販後臨床試験等を実施することとなる。</p>
--	--	--

注1) Guidelines for the ethical conduct of studies to evaluate drugs in pediatric populations.

Committee on Drugs, American Academy of Pediatrics. Pediatrics – 1995 Feb; 95(2): 286-94

注2) コンセントとアセントの関係

	対象	根拠
同意文書（コンセント）	代諾者（法的保護者）	GCP 省令 50 条
アセント文書	小児被験者（概ね中学生以上） <sup>1)</sup>	法的根拠なし（IRB・責任医師の判断）
アセント <sup>2)</sup>	小児被験者（概ね7歳以上） <sup>2)</sup>	法的根拠なし（IRB・責任医師の判断）

1) アセント文書あるいは同意文書に、同意の署名と年月日を小児被験者本人が記入すべきである。

2) 中学生未満の小児に対してもできる限り小児被験者本人が同意の署名と年月日をアセント文書に記入することが望ましい。

本人からの署名が得られない場合、あるいは文書を用いずに口頭でアセントが取られた場合は、代諾者に署名された同意文書に、本人からアセントが取られたことを記載するべきである。

薬食審査発第0928010号  
平成19年9月28日

各都道府県衛生主管部（局）長 殿

厚生労働省医薬食品局審査管理課長

### 国際共同治験に関する基本的考え方について

従来、我が国においては、ICH-E5ガイドラインに基づく「外国臨床データを受け入れる際に考慮すべき民族的要因について（平成10年8月11日医薬審第762号 厚生省医薬安全局審査管理課長通知）」により、いわゆる「ブリッジング」による海外臨床試験成績を承認申請資料として活用することを認めており、また、欧米諸国における市販後調査等の結果についても必要に応じ承認審査に際して活用しているところである。

他方、総合科学技術会議報告書「科学技術の振興及び成果の社会への還元に向けた制度改革について（平成18年12月）」においては、新規医薬品開発の効率化・迅速化の観点から外国との国際共同治験を推進すべき旨指摘しており、また、厚生労働大臣の検討会報告書「有効で安全な医薬品を迅速に提供するための検討会報告書（平成19年7月）」においては、「ドラッグ・ラグ（欧米で承認されている医薬品が我が国では未承認であって国民に提供されない状況）」解消のためには、国際共同治験の推進を図る必要があり、承認審査の観点から必要な国際共同治験実施に当たっての基本的考え方を明らかにする必要がある旨、指摘している。

このような状況を踏まえ、今般、独立行政法人医薬品医療機器総合機構における対面助言等の国際共同治験に関する現時点の知見について、別添「国際共同治験に関する基本的考え方」としてとりまとめたので、貴管下製造販売業者等の業務に活用するよう、周知方願います。

なお、本通知の写しを日本製薬団体連合会他関連団体宛てに発出していることを申し添える。

## (別添) 国際共同治験に関する基本的考え方

### はじめに

我が国では、ICH-E5 ガイドライン（「外国臨床データを受け入れる際に考慮すべき民族的要因について」平成10年8月11日医薬審第672号厚生省医薬安全局審査管理課長通知）が公表されて以来、ブリッジング開発戦略を通じて、医薬品の種類、対象疾患領域及び臨床開発の国際的進展状況に応じ、国内外の治験データを活用する知識及び経験が着実に蓄積しつつある。このような知識及び経験を生かし、近年、開発初期の段階から国際共同治験の実施を含めた開発戦略を採用するケースが増加しつつあり、日本における医薬品の開発戦略は、今後さらに多様化するものと考えられる。

我が国では、現在、国内での新薬承認時期が諸外国よりも数年遅いという問題（「ドラッグ・ラグ」）が深刻化しており、この問題を本質的に解消するためには、我が国における医薬品の開発時期を諸外国と同調させる必要がある。このための有効な手段の一つとして、日本が国際共同治験に早期から参加することが考えられる。これにより日本での医薬品開発が促進され、「ドラッグ・ラグ」が解消できれば、日本の患者が有効で安全な医薬品を諸外国に遅れることなく使用できるようになると考えられ、日本における薬剤治療レベルの向上及び公衆衛生の向上に大きく寄与するものと考えられる。

日本を含む国際共同治験を推進するため、独立行政法人医薬品医療機器総合機構（以下「総合機構」という。）は、平成18年度より国際共同治験に関する対面助言の予約申込みに際して優遇措置を講じている。国際共同治験の実施を前提とした治験デザイン、治験データ等の取扱いについて、総合機構と企業との間で、個々のケースに応じた検討を行うことは重要であり、これまでも対面助言を通じて対応してきたところであるが、今般、これまでの個々の事例等を踏まえ、国際共同治験を計画・実施する際の基本的な考え方をとりまとめることとしたものである。

本文書は、国際共同治験に関する現時点における基本的な考え方をまとめたものであり、企業における検討を促進し、日本の積極的な国際共同治験への参加を推進することに資するものとする。

なお、本文書に挙げた各事項は、現時点における科学的知見に基づいて検討されたものであり、今後の状況の変化、科学技術の進歩等に応じて随時見直され、改訂されるべきものであることに留意する必要がある。

### 適用範囲

本文書は、主に、新規の医薬品を今後開発する場合を想定しているが、既に諸外国において第Ⅱ相又は第Ⅲ相臨床試験が終了し、外国人に関する一定のデータが得られている場合にも適用可能な事項も含まれている。

### 基本的考え方

国際共同治験は、国内臨床試験とは異なり、様々な地域及び民族にまたがって臨床試験が実施されるため、その治験計画に際しては、民族的要因を考慮して計画することが必要である。したがって、ICH-E5ガイドラインで述べられている事項を検討することは、国際共同治験を計画する場合にも有用である。ブリッジングの考え方については、諸外国で開発が先行している場合のみならず、国際共同治験のように同時期に実施する場合にも適用可能である。この考え方は、ICH-E5ガイドラインのQ&A（「外国臨床データを受け入れる際に考慮すべき民族的要因についての指針」に関するQ&Aについて（その2）平成18年10月5日厚生労働省医薬食品局審査管理課事務連絡）の質問11において明確に示されているので参照されたい。

国際共同治験を計画・実施する際の基本的な考え方については、以下にQ&A形式で示す

こととするが、これらは一般的な原則を示すものであり、最適な開発戦略は個々の医薬品で異なることも想定される。個々のケースにおいては、開発戦略、試験デザイン等、本文書に掲げられた事項を予め検討すべきであり、可能な限り早期に総合機構との対面助言を活用すべきである。

なお、この文書において「国際共同治験」とは、新規の医薬品の世界的規模での開発及び承認を目指して企画される治験であって、一つの治験に複数の国や地域の医療機関が参加し、共通の治験実施計画書に基づき、同時並行的に進行するものを言う。

### 1. 国際共同治験を実施する上での基本的な要件は何か？

以下のすべての条件を満たすことが必要である。

- 参加するすべての国、医療機関等で ICH-GCP に準拠した臨床試験が実施可能であること。
- 参加するすべての国、医療機関等で日本からの GCP 実地調査を受入れ可能であること。
- 治験薬の有効性及び安全性に影響を及ぼしうる要因（人種、地域、患者背景等）を予め検討するとともに、当該要因に関する部分集団解析が実施可能であり、適切な考察が可能であること。
- 慣習などの社会的相違や試験の管理・運営等各施設における治験実施状況を適切に把握でき、認められた差異が試験結果に影響を及ぼしうるものであるかどうかについて、適切に考察可能な状況であること。

### 2. 日本はいつからグローバル開発に参加すべきか？

世界的に進行している臨床開発について、できるだけ早期に参加することが望ましい。このため、遅くとも用量反応性を探索的に検討する段階の試験から参加できるよう予め検討しておくことが重要である。

### 3. 患者を対象とした国際共同治験を実施する場合に、その試験開始前に日本人での第 I 相試験や日本人での薬物動態情報は必須か？

国際共同治験で用いる用法・用量が日本人においても安全性上特段の問題がないかについて、予め確認しておく必要がある。そのためには、国際共同治験を開始する前に、少なくとも日本人の健康な志願者又は患者を対象とした治験薬の単回投与試験による安全性や薬物動態等を検討し、外国人における結果と比較して、日本人におけるリスクが外国人におけるリスクと遜色ないことを確認しておくことが求められる。

ただし、海外で実施された第 I 相試験の結果から日本人に対する安全性を判断することが可能な場合や類薬での状況等から日本人と外国人における推奨用量が同様と判断できる場合等においては、必ずしも国際共同治験開始前に第 I 相試験を実施する必要はない。なお、この場合においても薬物動態と臨床効果との関連等を日本人と外国人で比較検討しておくことが日本人に適切な用量を設定する上で有用であると考えられること、また、国際共同治験の結果の解釈に際しても重要な情報となりうると考えられることなどから、必要に応じ、国際共同治験の実施と並行して適切な臨床薬理試験を実施したり、国際共同治験の中で薬物動態と臨床効果との関連を検討するなどして、その結果を承認申請資料に含めることが望まれる。

4. 海外臨床試験成績に基づき用量設定を行い、国内での用量反応試験を実施せずに、検証的な第Ⅲ相試験から日本人を組み入れるという開発計画は受入れ可能か？

これまでの承認事例及び ICH-E5 ガイドラインに基づく承認審査の経験等を踏まえると、日本人と外国人との間で薬物の体内動態等が異なることもあり、また、外国人での臨床試験結果に基づき設定された推奨用量が日本人での推奨用量であると結論付けることは困難である場合もみられることから、質問のような開発計画を基本的考え方とすることは有効かつ安全な医薬品を日本人患者の元へ届けるといふ本来の目的からして適切ではない。

したがって、開発を円滑に進め、日本における承認時期を諸外国と同時期とするためには、用量反応試験に日本人の患者等を含め、民族間での用量反応性の差異を臨床開発の早期に同定し、その後の検証的試験を計画することが望まれる。また、仮に日本人と外国人とで推奨用量が異なっている場合、各地域ごとに設定した用量について、有効性及び安全性の検証が同等に扱えることを適切に説明できるのであれば、その後の第Ⅲ相国際共同治験（検証的試験）において各地域での結果を統合し主要な解析集団として取り扱うことも可能である。

なお、PK (pharmacokinetics) に類似性があり、PK と臨床効果との関連が明らかとなっている PD (pharmacodynamics) との間で相関性が示されているような場合等には、臨床効果を指標とした日本人での用量反応試験は必ずしも必要ないと考えられる。

(注) 希少疾病用医薬品又は生命に関わるような疾患で他の治療法が確立していないような場合、そもそも国内での用量反応試験を行うことが困難な場合があり、このような場合には、医師の厳重な管理下で第Ⅲ相試験を行うなどの工夫を検討すべきである。

5. 国際共同治験を計画する場合の基本的な留意事項は何か？

基本的には以下の事項について留意すべきである。なお、詳細については、ICH-E5ガイドラインのQ&Aの質問11を参照されたい。

- 国際共同治験を実施する場合には、それぞれの地域における民族的要因が治験薬の有効性及び安全性に及ぼす影響について評価し、また、日本人における治験薬の有効性及び安全性について評価できるよう計画することが必要である。
- 実施する国際共同治験のデザイン及び解析方法が我が国にとって受入れ可能なものであることが必要である。
- 主要評価項目は、各地域に許容されているものであるべきであり、主要評価項目が地域により異なる場合には、すべての地域においてすべての主要評価項目に関するデータを収集し、地域間での差異を検討できるようにすべきである。
- 安全性評価を適切に実施するため、全地域での有害事象の収集方法及び評価方法をできる限り統一すべきである。

6. 用量反応試験等の探索的な試験あるいは検証的な試験を国際共同治験として実施する場合に、症例数の設定及び日本人の症例の割合の決定はどのようにすることが適切であるか？

- 国際共同治験では、全集団での結果を前提とした症例数の設定も可能であり、日本人の部分集団において統計的な有意差を検出するだけの検出力を必ずしも確保する必要はない。しかし、検証的試験での主要な解析対象を全集団として規定する場合には、各地域での集団ではなく、全地域での集団を一つの集団としてみなすことができることを根拠を説明する必要がある。
- 仮に日本人の部分集団での結果が、全集団での結果と著しく乖離している場合には

その理由を十分検討すべきであり、必要に応じさらなる臨床試験の実施も考慮すべきであることから、このような場合には、総合機構における対面助言を活用することが推奨される。

- 国際共同治験は、全集団での結果と日本人集団での結果に一貫性が得られるよう計画すべきであり、各地域での一貫性が担保されることで、全集団での結果を適切に各地域に外挿することが可能となると考えられる。
- したがって、症例数については、一般的に推奨できる方法は現時点で確立されておらず、実施地域の数、試験規模、対象疾患、全体での症例数と日本人症例数との割合等を考慮して決定することが必要である。例えば、国際共同治験を計画する際に全集団と日本人集団において一貫した結果が得られる可能性を考慮する方法として、定量的な評価変数を用いたプラセボ対照試験を例にとると以下のような方法があると考えられる。

(1) 方法 1: プラセボ群と治験薬群での群間差を  $D$ 、その場合の全集団での群間差を  $D_{all}$ 、日本人集団における群間差を  $D_{Japan}$  とすると、 $D_{Japan}/D_{all} > \pi$  が成立するような確率が 80 %以上となるように日本人症例数を設定する。 $\pi$  については、適切な値を設定する必要があるが、一般的には 0.5 以上の値をとることが推奨される。この方法では、日本人症例数を最小にしようとする、全体での症例数が増加し、全体での症例数を最小にしようとする日本人症例数が増加するという関係が認められる。

(2) 方法 2: 全集団におけるプラセボ群と治験薬群での群間差を  $D_{all}$ 、例えば 3 地域が試験に参加し、各地域でのプラセボ群と治験薬群での群間差をそれぞれ  $D_1$ 、 $D_2$ 、 $D_3$  とすると、 $D_1$ 、 $D_2$ 、 $D_3$  が全て同様の傾向にあることを示す。例えば  $D_{all}$  が正の値をとるとすると、 $D_1$ 、 $D_2$ 、 $D_3$  のいずれの値も 0 を上回る確率が 80 %以上となるように症例数を設定する。この方法では、各地域から均等に症例数を集積した場合に、確率が高くなるという傾向があり、全体の症例数を変更することなく日本人症例数を検討することが可能であるが、日本人の構成比率が小さく、症例数が少ない場合に、地域間比較が十分に行えない場合があることに留意すべきである。

(参考) 例えば、数百例程度を対象に、プラセボを対照とした 2 群での並行群間比較試験を実施する場合、各地域での有効性が同様であると仮定すると、試験結果に基づき科学的に適切な評価を行うためには、方法 1 においては、全体の症例数の増加を適度に抑制しながら日本人症例数を最小にしようとする場合、およそ 20 %の症例を確保することが必要になる。また、方法 2 の場合には、全集団での検出力を 90 %として 3 地域で実施するのであれば、日本人症例数としておよそ 15 %以上を確保することが必要になる。

(注) 個々のケースにおける具体的な目標症例数の設定については、総合機構と相談することが可能である。

7. 国際共同治験においては、諸外国では確立されているが、我が国ではまだ確立されていないような指標であっても、主要評価項目とせざるを得ない場合もあるが、このような場合でもその指標は受け入れ可能か？

そのような場合が想定されるのであれば、できる限り早期に国内でパイロット試験等を実施し、海外臨床試験結果と同様の反応が得られるかどうか確認しておく必要があると考えられる。また、国際共同治験実施前には、予め統一的な評価方法に関する研修プログラムを作成し、実施するなど、評価者間、施設間、各地域間での差を最小限にする工夫が必要である。



質問のような場合、何ら国内での検討がない状況で国際共同治験に参加することは、日本での成績が適切に得られないばかりか、試験全体に悪影響を及ぼす可能性があることに留意すべきである。

8. 日本を含まない諸外国で既に実施されている国際共同治験と同一のプロトコルによる小規模の国内治験を別途実施し、それらの結果をもって国内外での有効性及び安全性が同様と結論付けることは可能か？

既に海外で実施されている臨床試験があつて、国内試験結果とは別に解析が実施されるような場合には、一般的に別試験として考えるべきである。このような場合、設定されている主要評価項目が適切であるとの前提で、国内で統計的な検討も実施できるよう症例数を確保したブリッジング試験をプロスペクティブに計画・実施し、ICH-E5 ガイドラインに基づき、海外臨床試験成績との差異等について検討することが適切であると考えられる。

9. 第Ⅲ相の検証的な国際共同治験での対照群について

- (1) 国際共同治験でプラセボのみが対照群と設定されている場合があるが、そのような場合においても日本では別途実薬を対照群として設定することが必要であるか？
- (2) 対照群として用いられる実薬が国際的な標準薬であっても、国内では未承認の場合があるが、このような場合に国内での対照薬をどのように設定すべきか？

- (1) 原則としてその必要はなく、治験薬のプラセボに対する優越性が検証され、全集団と日本人集団との間で一貫した結果が得られるよう臨床試験を計画することが適切である。なお、別途実薬を対照とした比較試験成績があれば、治験薬の臨床的位置づけが明確になる場合もあると考えられるが、臨床的位置付けについては他の方法によって説明可能な場合もあり、一律にそのような試験の実施を求めるものではない。
- (2) 用いようとしている実薬が国際的に標準薬であることが諸外国等のガイドライン等の記載から客観的に説明できるのであれば、国内で未承認の医薬品であっても、治験における対照薬として試験を実施することは可能であるが、当該未承認薬が日本人にどのような影響を及ぼすか、特に当該未承認薬の安全性について予め検討しておく必要がある。なお、非劣性検証を目的とした試験における結果の解釈については慎重に判断する必要があるので、当該未承認の対照薬の有効性・安全性に関するデータ、特に日本で既承認である医薬品と当該未承認薬の相違等について、可能な限り情報を収集し、得られる結果の日本人患者への外挿性を予め検討しておくことが望ましい。

10. 国際共同治験で設定される併用薬、併用療法等を完全に国内外で同一に設定することは困難であるが、どのように設定することが適切であるか？

併用薬、併用療法といっても多種多様であり、一概に回答することは困難であるので、2つの例を挙げることにする。

なお、いずれの場合についても、併用薬の用法用量の妥当性について、個々のケースに応じた十分な検討が必要である。その前提の上で、実施しようとする国際共同治験の計画の根

拠が諸外国におけるエビデンスに大きく依存しているような場合には、基本的には、諸外国と同一の併用薬、併用療法等を用いて国内での臨床試験を実施する方が理論的であり、かつ試験の成功率を高めるものと考えられる。

(1) 抗がん剤の併用療法のように、治療域が狭く毒性の強い薬剤同士を併用するような場、国内外で厳密に同一の用法・用量を設定することが望ましい。

(2) ある疾患に対して用いられる標準的治療法として通常併用される薬剤又は治療法については、各地域の用法・用量を同一に規定することが困難であることも想定される。このような場合、必要最小限の差異については許容可能であるが、これらの差異が治験薬の有効性及び安全性に及ぼす影響を最小限にとどめるため、併用時の用法・用量等を前観察期も含めて治験期間中は変更せず、患者ごとに一定にするなどの配慮が必要である。

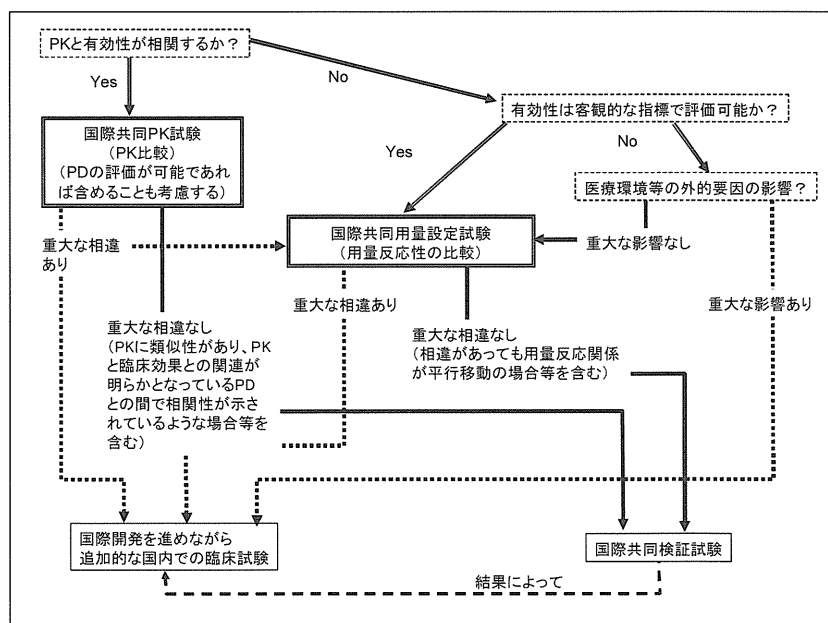
11. 国際共同治験を実施することが望ましい領域はあるか？

どのような領域であっても国際共同治験を実施することは可能であるが、希少疾病等の国内で大規模な検証試験を実施することが困難と考えられる疾患であれば、より積極的に国際共同治験の実施を検討すべきである。これまでは国内で少数例の使用経験的な試験が実施されている場合もあったが、このような開発戦略よりも、国際共同治験の中に可能な限り多くの日本人の症例を組み入れ、より適切な臨床試験デザインに基づくエビデンスを構築することが望ましい。また、希少疾病でなくても、生存率等の真の臨床的エンドポイントを用いた臨床試験で症例数が数千例規模に上る場合等、世界的にも症例の集積に時間を要すると予想される時には、日本も積極的に参加してエビデンスの構築に貢献することも考えられる。このような方策をとることで、最終的には各地域での承認申請時期を同様とすることが可能になると考えられる。

12. 国際共同治験を実施することの適否を判断する上で、参考となるフローチャートはあるか？

数多くの例外があると想定され、万能なアプローチは存在せず、個別に判断することが適切であると考えられるが、現時点での一般的な考え方としては、下図のようなフローチャートが参考になるとと思われる。

なお、図中[ ]で示した事項は、同一疾患に対する先行試験の結果等からも確認可能な場合があると考えられる事項であり、また、重大な相違及び影響の有無については、実施された臨床試験のデザインや各集団での比較等も踏まえて判断する必要がある。



### Ⅲ. 研究成果の刊行に関する一覧表

## 研究成果の刊行に関する一覧表

### 雑 誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
手良向 聡	臨床試験デザイン	京都府立医科大学 雑誌	123 (11)	769-777	2014

## **IV. 研究成果の刊行物・別刷**

---

## 総 説

---

### 臨床試験デザイン

手良向 聡\*

京都府立医科大学大学院医学研究科生物統計学

### Clinical Trial Design

Satoshi Teramukai

*Department of Biostatistics,*

*Kyoto Prefectural University of Medicine Graduate School of Medical Science*

### 抄 録

臨床試験はデザイン・計画の段階から始まり、試験実施・データ管理・モニタリングを経て、データ解析・報告書作成に至る。この各ステップが統計的方法を必要としている。再現性によって結果を保証することが可能な基礎実験と異なり、同じデザインで繰り返すことが困難な臨床試験においては、デザインと手続きの妥当性から結果を保証するしかない。計画段階から生物統計学の専門家が参画していれば、質の高い臨床試験を実施できる可能性は高い。近年、臨床試験にベイズ流の方法が有用であるという報告は着実に増えている。効率的かつ倫理的な試験デザインの開発は、資源を有効に活用するという観点から今後ますます重要になるであろう。本稿では、ランダム化対照試験の標準的方法、および探索的臨床試験のデザインとして有用であるベイズ流の方法について述べる。

キーワード：技術評価，統計的仮説，ランダム化，統計的考察，ベイズ流統計学。

### Abstract

Every clinical trial starts from the design and planning stage, moves to trial conduct, data management and monitoring, and finally to the data analysis and conclusions. Each step along the way calls for statistical methods. While basic research which can be guaranteed a result by reproducibility, clinical trials which cannot be repeated with the same design must be guaranteed by the validity of design and procedure. If a biostatistician take part in the planning stage, we can do the high quality of clinical trials. In recent years, reports insisting on the usefulness of Bayesian statistics in clinical studies have steadily increased. The development of efficient and ethical design will become important in the future, from the viewpoint of the best use of resources. In this paper, I will describe the standard methods for randomized clinical trials, and Bayesian methods which could be useful for the design of exploratory clinical trials.

**Key Words:** Technology assessment, Statistical hypothesis, Randomization, Statistical consideration, Bayesian statistics.

---

平成26年 9月16日 受付

\*連絡先 手良向聡 〒602-8566 京都市上京区河原町通広小路 上ル梶井町465

steramu@koto.kpu-m.ac.jp

## はじめに

臨床試験はデザイン・計画の段階から始まり、試験実施・データ管理・モニタリングを経て、データ解析・報告書作成に至る。この各ステップが統計的方法を必要としている。臨床試験は、良いデザインおよび正しい遂行がなければ、台無し（例えば、結論の出ないまたは誤った知見を導く）になり、さらには大惨事（例えば、不必要に多数の患者が毒性または死を被ることを引き起こす）になる可能性さえある。臨床試験は効率的かつ倫理的であるべきで、資源を節約し、より多くの患者に恩恵を与え、より迅速に正しい結論を引き出し、結果として不必要な毒性をより少なくすべきである<sup>1)</sup>。本稿では、ランダム化対照試験の標準的方法、および探索的臨床試験のデザインとして有用であるベイズ流 (Bayesian) の方法について述べる。

### 医学・医療と技術評価

医学は普遍性のある真実を追求する科学の一分野である一方、医療は多様性のある個人に対して最適な技術を選択して適用することが要求される場である。技術評価は、主に統計学に基づく科学的方法を駆使して医療技術を相対的に評価し、医学から医療への橋渡しを行う (図1)。

統計学を医学・医療の領域に導入するには、2つの大きなギャップを認識しておく必要がある。1つは、「決定論」と「非決定論 (確率論)」のギャップである。1800年代半ばにクロード・ベルナールが「統計学に立脚している限り、医学は永久に推測科学に止まるであろう」と決定論的な考え方を主張して以来、医学の世界では決定論的な思想が支配的である。もう1つのギャップは、意思決定の主体に関わる問題であり、「対集団の確率」と「対個人の確率」とのギャップである。たとえば、ある医薬品を承認すべきかどうかという判断は、その国の人々という集団に対するベネフィットとリスクのバランスで決定される。その決定は「対集団の確率」に基づく一方、医療の場で診断や治療を行う際には、個人に対するベネフィットとリスクを評価しなければならない。たとえば、胎児診断を行って、医師が「胎児に異常がある確率は80%」と言ったとき、その80%は集団での頻度に基づいたものである。しかしながら、それを聞いた母親の「子供には異常があるか (100%)、ないか (0%) のどちらか」という感覚で、この確率を解釈することはそれほど容易ではない。これは「確率とは何か」という哲学的課題につながっている。

臨床試験は、20世紀を代表する英国の統計学

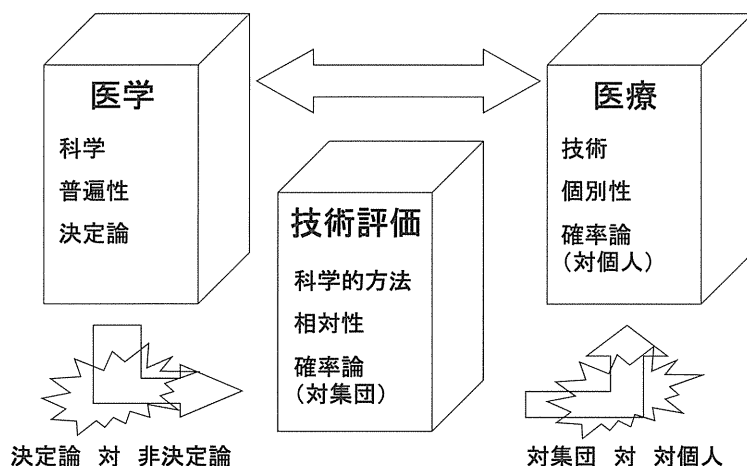


図1 医学・医療と技術評価

者 R.A. フィッシャー (1890~1962) が創始した統計的実験 (技術的実験とも呼ばれる) の方法論を基礎としている。科学的実験は、人工的に作り出された純粋な条件のもとでの因果関係を確定しようとするのに対して、統計的実験は以下の特徴を有する<sup>2)</sup>。

- ・実験の場は、現実の応用の場に近い状況に設定される
- ・結果の分析には誤差の存在を前提にしなければならない
- ・いくつかの因子を同時に変化させて結果を見る必要があることがある
- ・目的は、何らかの基準によって現実の場において最も良い結果が得られるような条件を求めることである

様々な種類の誤差を伴うデータを扱うためには統計的方法が不可欠である。また、再現性によって結果を保証することが可能な基礎研究と異なり、同じデザインで繰り返すことが困難な臨床試験においては、デザインと手続きの妥当性から結果を保証するしかない。

### 探索的試験と検証的試験

臨床試験の性格は検証的試験と探索的試験に大別される。検証的試験は有効性または安全性の確固たる証拠を提示するための試験と位置付けられる。ただし、個々のいかなる試験も検証的側面と探索的側面の両方を持つ。試験実施計画書 (以下、プロトコル) には、各試験について検証的な証明として用いられる側面と、探索的解析のためにデータを提供する側面とを、明確に区別しておくべきである<sup>3)</sup>。探索的解析から得られた結果は仮説に過ぎず、その仮説は検証的試験によって確認しなければならない。

### 統計的仮説

統計的仮説の代表的なものは、優越性仮説と非劣性仮説である。優越性仮説を証明しようとする試験 (優越性試験) とは、試験治療の効果が対照治療 (活性対照またはプラセボ対照) よりも「臨床的に優れること」を示すことが目的の試験である。一方、非劣性仮説を証明しよう

とする試験 (非劣性試験) とは、試験治療の効果が対照治療よりも「臨床的に劣らないこと」を示すことが目的の試験である。多くの場合は優越性仮説が設定されるが、対照治療 (通常、活性対照) に比べて安全性あるいは経済性に優れていることが見込まれる場合に、このような非劣性仮説が許容されることがある。

非劣性試験を計画する際には、非劣性マージン (臨床的に意味のある最小の差:  $\Delta$ ) の決定、データの質などについて十分な注意が必要である。ハザード比を治療効果の尺度とした臨床試験の場合、優越性試験では、ハザード比の 95% 信頼区間の上限が 1 より小さければ、有意水準 5% で試験治療が優れると判断される。一方、非劣性試験では、その 95% 信頼区間の上限が  $1+\Delta$  よりも小さければ、有意水準 5% で試験治療は対照治療に  $\Delta$  以上は劣らない、と判断される (図 2)。

### ランダム化

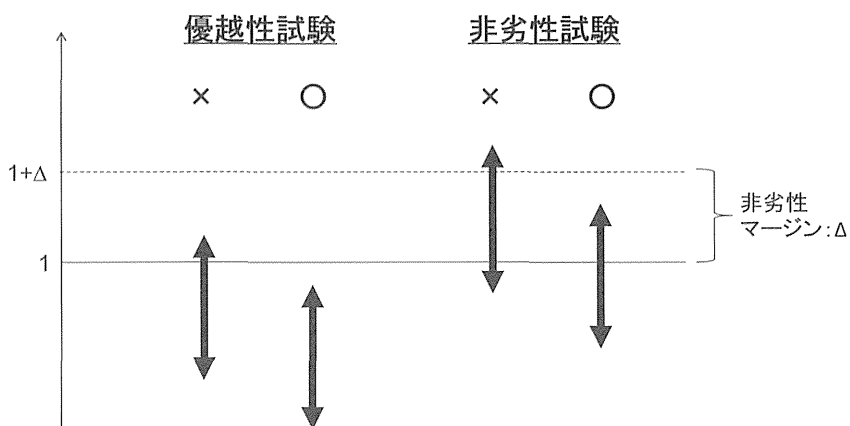
ランダム化対照試験では、登録の際にランダム化という操作が必要になる (図 3)。ランダム化を行うことにより、試験群と対照群の比較可能性 (内的妥当性とも呼ばれる) が保証される。ランダム化は「実験計画法」を確立したフィッシャーの偉大な発明の一つである。実験に伴う誤差には以下の 2 種類がある。ちなみに、フィッシャーは臨床試験ではなく、農事実験に従事していた。

- ・偶然誤差…測定誤差のようにある確率分布に従うと想定できる誤差であり、繰り返し測定を行えばその大きさについて推定可能である
- ・系統誤差またはバイアス…圃場の肥沃度や日当たりの不均一性のように確率変動と見なせない誤差であり、繰り返しには関係なく結果を歪める原因となる

ランダム化の目的は、一言で言うと「系統誤差を偶然誤差に転化すること」である。臨床試験におけるその意義は、

- ・予後因子が既知か未知かにかかわらず、予後因子の分布が類似したグループを作る





矢印は、ハザード比(試験薬のハザード率/対照薬のハザード率)の95%信頼区間

図2 優越性試験と非劣性試験の判断規準

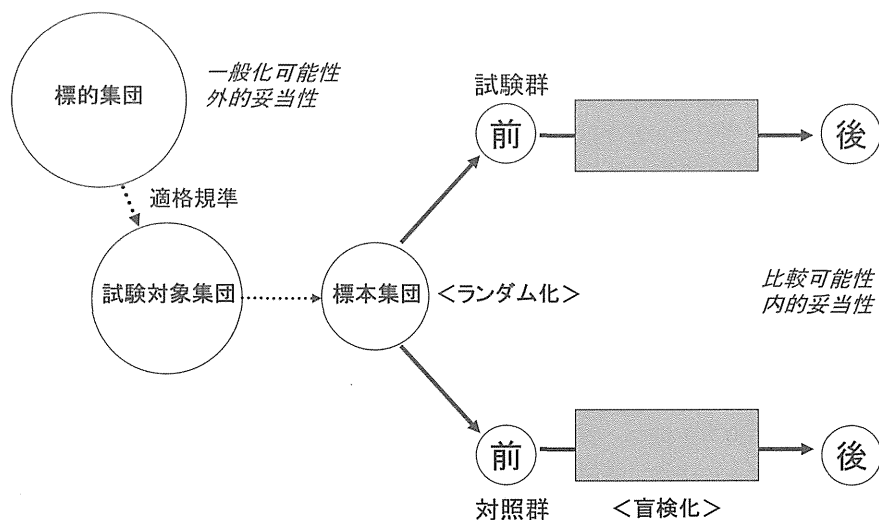


図3 ランダム化対照試験

・データ解析の際に治療効果の定量的な評価のための正しい統計的根拠を与えることである<sup>3)</sup>。

臨床試験で利用されるランダム化の方法は、固定的割付(単純ランダム化、並べ替えブロックランダム化など)と適応的割付(偏コインデザイン、最小化法など)とに分類される。重要な予後因子の分布を治療群間で均等にするためには、それらの因子で層を作り、その層ごとに

ランダム化(層別ランダム化)を行う必要がある。ランダム化の方法には多くの選択肢があり、試験統計家がそれぞれの利点と欠点を考慮して決定すべきである。

### 統計的考察

試験統計家がプロトコルに記載すべき統計的事項として、「標本サイズの設定根拠」、「解析対象集団」、「解析項目・方法」、「中間モニタリン

グ」がある。

### 1. 標本サイズの設定根拠

臨床試験に参加していただく対象数（標本サイズ）は科学性、倫理性、実施可能性のバランスを考慮して決める必要があり、多すぎても少なすぎてもいけない。標本サイズは主要評価項目に関する情報に基づいて計算される。その時点における情報を最大限利用するものの、一時的な仮定に基づく概算であることに注意が必要である。例えば、ある癌の補助化学療法の臨床試験において、標準治療を受ける患者の5年生存率を予測する場合、利用できる情報にはかなり大きなばらつきがある。これらの前提条件を慎重に検討した上で、通常は仮説検定に基づく決定方式に従って、帰無仮説、対立仮説、検定統計量、有意水準、第I種の過誤、第II種の過誤または検出力などが試験の性格（探索的または検証的）を考慮して決定される。

### 2. 解析対象集団

解析対象集団は、ITT (intention-to-treat) の原則に従って定義すべきである。これは、「被験者が実際に受けた治療ではなく、被験者を治療しようとした意図 (intention to treat) に基づいて評価する」という原則である<sup>3)</sup>。従って、登録されたすべての被験者を解析対象とすることが原則であるが、登録後に判明した不適格例、試験治療を全く受けなかった例を対象から除くことは一般的に許容される。いずれにしても、有効性および安全性に関する主要な解析対象集団の定義をプロトコルに明記し、報告時には解析対象から除外した対象数とその理由を明記する必要がある。

全体集団のある一部のグループを対象とした解析をサブグループ解析（サブセット解析、部分集団解析）と呼ぶ。通常、患者特性やベースライン情報に基づいてサブグループ化を行う。「試験治療を完遂した集団」のように介入後の情報に基づいて対象を選択する場合には、比較可能性など別の問題が生じるため、サブグループ解析とは区別しておく必要がある。ランダム化の際に層別に用いた因子によるサブグループ解析は、ランダム化に基づく比較可能性の条件を

満たしている。そうでない場合は比較可能性の条件を満たさない。ただ、その条件を無視したとしても、多数の検定の実施により第1種の過誤確率が上昇すること、一方ではサブグループ内の標本サイズ不足により検出力が低下することが問題となる。対応策としては、1) 関心のある少数のサブグループを事前にプロトコルに記載する、および2) 交互作用の検定が有意な場合にのみサブグループでの検定を行うことが推奨されている。交互作用には、量的交互作用と質的交互作用の2種類があり、交互作用が検出されたときは有意差の有無だけでなく、医学的な意義と解釈について十分な吟味が必要である。いずれにしても、サブグループ解析は、探索的解析の代表的なものであり、その目的は仮説の生成である。

### 3. 解析項目・方法

臨床試験で利用される標準的な統計解析手法について表1に示す。ランダム化対照試験では、ランダム化によって比較可能性が保証されているので、通常は観察研究のように複雑な回帰モデルを用いて交絡因子を調整する必要はない。

### 4. 中間モニタリング

中間モニタリング（中間解析、中間評価とも呼ばれる）の目的は、

- ・ 試験治療の優越性が疑いなく立証された場合
- ・ 適切な試験治療の差を示す見込みがないことが判明した場合
- ・ 許容できない有害事象が明らかになった場合

に試験を早期に中止することである<sup>3)</sup>。検定の多重性を考慮した中止規則の設定には多くの方法（グループ逐次法など）が開発されてきているが、実際に臨床試験を中止すべきかどうかという判断は純粋に統計的な問題ではなく、臨床的のみならず社会的な影響も考慮する必要がある。そのような判断を公正に行う場として、中間モニタリングを実施する際には、当該臨床試験に関与しない第三者からなるデータモニタリング委員会（効果安全性評価委員会とも呼ばれ

表1 変数の型別の標準的な統計解析手法

目的	連続変数	分類変数	時間-イベント変数
分布の記述	ヒストグラム、箱ヒゲ図、散布図	ヒストグラム、分割表	生存曲線 (Kaplan-Meier法)
要約統計量	平均、分散、中央値、パーセント点、相関係数	頻度、一致度、相関係数	x年生存率、中央生存期間
検定(単純)	t検定、分散分析、Wilcoxon検定	$\chi^2$ 検定、Fisher正確検定	ログランク検定
検定(層別)	共分散分析	Mantel-Haenszel検定	層別ログランク検定
回帰モデル	重回帰分析	ロジスティック回帰分析	Cox回帰分析

る)を設置しなければならない。

### 探索的臨床試験のデザイン —ベイズ流の方法

1950年頃に臨床試験の方法論がほぼ確立して以来、統計的評価の方法として、フィッシャーあるいはネイマン・ピアソンによる頻度流(frequentist)の仮説検定・推定が主に用いられてきた。データ解析へのベイズ流統計学の適用は、物理学をはじめとする多くの自然科学分野および社会科学分野で広く行われており、医学・生物学分野においても、ベイズ流の統計モデルを適用したデータ解析の事例は多く存在する。しかしながら、臨床試験のデザインにベイズ流の方法を適用した事例としては、抗がん剤の第I相試験(最大耐用量を決定するための試験)でのCRM(continual reassessment method)、ランダム化試験の中間解析でのベイズ流予測確率の利用などがあるが、未だそれほど多くない。

大学等の研究機関が主体となって実施するトランスレーショナルリサーチおよび臨床試験の対象疾患は、難治性かつ重篤であり、そのうえ患者数が限られているという特徴がある。このような状況では、基礎研究で認められたコンセプトを実証するためのPOC(proof of concept)試験と呼ばれる探索的試験の実施が主であり、

疾患の重篤性を考えると同時対照を設定すること自体困難な場合が多い。また、被験者のリスクを最小にするために臨床試験の途中で結果をモニタリングしながら意思決定を行うというような柔軟な対応も必要である。さらに、被験者数を最小にするために、過去に得られた証拠や情報(事前情報)を十分に生かすことも重要となる。これらを鑑みると、予期しない事態が発生して試験途中でデザイン(標本サイズや中間モニタリングの時期・方法など)の変更を行う場合などに、頻度流接近法に基づく方法は柔軟性の観点から不十分であり、新しい方法の開発が必要となる。近年、ベイズ流接近法は柔軟性と効率性の面から有望と考えられている。臨床試験におけるベイズ流接近法の主な特長は以下の通りである。

- ① 解釈が容易な「確率」だけを用いて整合性のある推測と意思決定を行うことができる
- ② 標本サイズに関わらず事前分布を事後分布に更新して推測ができる
- ③ 予測分布を用いて試験結果を予測することができる

また、ベイズ流デザインの動作特性が頻度流に評価できることも1つの利点である。

### ベイズ流デザインの例

すべての被験者に同一の試験治療を行う単群臨床試験は、探索的な臨床試験の大部分を占めている。その主目的は、治療効果に対する確定的な証拠を得ることではなく、さらに研究を継続すべき有望な治療をスクリーニングすることである。単群臨床試験デザインの多くは、効果が認められない場合は試験を早期中止することが望ましい致死的な疾患の領域で開発されてきた。抗がん剤の第Ⅱ相単群臨床試験については、1960年代から頻度流の方法が開発され、1990年代以降ベイズ流接近法を用いたデザインがいくつか提案されている。その中には、効用/損失関数を明示的に用いるベイズ流決定理論に基づく手法も含まれる。

ここで、被験者20名にある試験治療を行い、14名に「効果あり(成功)」、6名に「効果なし(失敗)」という結果が得られた、二値(成功または失敗)評価項目の単群臨床試験という単純な事例を用いてベイズ流の方法を概説する。まず、事前情報が存在しないと仮定し、事前分布を一様分布(Beta(1,1))と表現されるベータ分布と定める。次に、ベイズの定理を用いて、

事前分布と観察データ(実際には尤度[ゆうど]と呼ばれる形に変換されたもの)を結合し、事後分布Beta(15,7)(=Beta(1+14, 1+6))を得る(図4)。このように更新された事後分布から、この治療の成功確率の平均は0.68(=15/(15+7))、成功確率が0.5以下の確率は分布下面積から0.039と得られる。この結果から、引き続いて5名の被験者に同じ治療を行ったときに何名の被験者に成功が観察されるかという予測分布を得ることもできる(図5)。

### おわりに

フィッシャーは1938年に次のように述べている:「同じだけの時間と労力をかけたとしてもデータ収集の過程、または実験計画を厳密に検討しているか否かによって、得られる収穫は10倍から12倍にもなる。実験終了後に統計学者に相談を持ちかけるのは、統計学者に、単に死後診察を行って下さいと頼むようなものである。統計学者はおそらく何が原因で実験が失敗したかという実験の死因について意見を述べてくれるだけであろう」<sup>6)</sup>。

臨床試験に統計的方法は必須であり、計画段階から生物統計学の専門家が参画していれば、

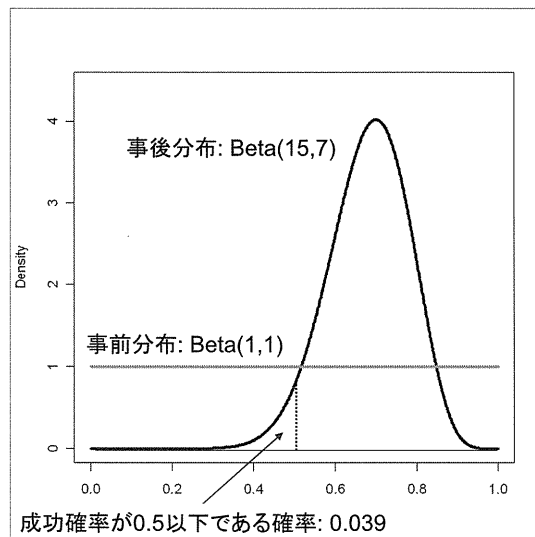
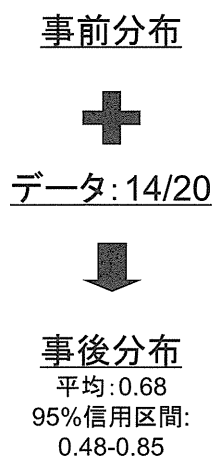


図4 事前分布から事後分布へ