

医薬薬審発 0620 第 1 号
令和 6 年 6 月 20 日

各都道府県衛生主管部（局）長 殿

厚生労働省医薬局医薬品審査管理課長
（ 公 印 省 略 ）

「臨床試験のための統計的原則」の補遺について

近年、優れた新医薬品の研究開発を地球規模で促進し、患者へ迅速に提供するため、承認審査資料の国際的な調和の推進を図ることの必要性が指摘されています。このような要請に応えるため、医薬品規制調和国際会議（以下「ICH」という。）が組織され、品質、安全性及び有効性の各分野で、承認審査資料の国際的な調和の推進を図るための活動が行われています。

臨床試験における統計的原則については、「臨床試験のための統計的原則」について（平成 10 年 11 月 30 日付け医薬審第 1047 号厚生省医薬安全局審査管理課長通知）により通知したところです。

今般、ICH において、臨床試験の計画、デザイン、実施、解析及び解釈をよりよいものとすることを目的として、「臨床試験のための統計的原則」に関する補遺が別添のとおり合意されました。本補遺では、臨床試験の計画、デザイン、実施、解析及び解釈を整合させる構造化されたフレームワークを明示し、試験の目的に基づく臨床的疑問を反映する治療効果の詳細な説明である estimand の考え方を導入するとともに、試験結果の安定性を検討するための感度分析の役割を明確にしています。つきましては、貴管下関係業者等に対して周知方御配慮願います。

なお、本通知の写しについて、別記の関係団体宛てに発出しますので、念のため申し添えます。

(別記)

日 本 製 薬 団 体 連 合 会
日 本 製 薬 工 業 協 会
米 国 研 究 製 薬 工 業 協 会 在 日 執 行 委 員 会
一 般 社 団 法 人 欧 州 製 薬 団 体 連 合 会
独 立 行 政 法 人 医 薬 品 医 療 機 器 総 合 機 構

ICH E9(R1)

臨床試験のための統計的原則 補遺

臨床試験における estimand と感度分析

目次

A.1. 目的と適用範囲	3
A.2. 計画、デザイン、実施、解析及び解釈を整合させるフレームワーク	5
A.3. Estimand	6
A.3.1. 関心のある臨床的疑問に反映させるべき中間事象	7
A.3.2. 関心のある臨床的疑問の定義における中間事象に対応するためのストラテジー	8
A.3.3. Estimand の要素	11
A.3.4. Estimand を構成する際の留意事項	12
A.4. 試験デザイン及び実施への影響	15
A.5. 試験の解析への影響	17
A.5.1 主とする推定	17
A.5.2. 感度分析	19
A.5.2.1. 感度分析の役割	19
A.5.2.2. 感度分析の選択	20
A.5.3. 補足的解析	20
A.6. Estimand と感度分析の記載	21
用語集	22

A.1. 目的と適用範囲

製薬会社、規制当局、患者、医師及びその他のステークホルダーが意思決定する際の情報を適切に与えるためには、病状に対する治療（薬剤）のベネフィットとリスクを明確に説明すべきである。報告された「治療効果」は、そのような明確な説明がない場合には、誤って理解されてしまう懸念がある。本補遺は、臨床試験の目的の設定、デザイン、実施、解析及び解釈に関わる専門分野間、及び治験依頼者と規制当局との間での、臨床試験で検討すべき関心のある治療効果に関する議論をより良くするための、構造化されたフレームワークを提示する。

関心のある臨床的疑問に対応する「estimand」（用語集、A.3 節）を構成することにより、関心のある治療効果の説明を明確にすることが容易になる。この説明を明確にするためには、割り付けられた治療の中止、追加又は他の治療の使用、及び死亡などの終末事象といった中間事象について慎重に想定することが必要である。Estimand の説明は、これらの中間事象について関心のある臨床的疑問を反映すべきであり、本補遺では、提起される可能性のある様々な関心のある疑問を反映するためのストラテジーを導入する。ストラテジーの選択は、例えば関心のある治療、集団、又は変数（評価項目）といった従来の試験の要素が、臨床的疑問を説明する際にどのように反映されるのかに影響を与える可能性がある。

臨床試験のデータの統計解析は estimand に整合させるべきである。本補遺では、主要な統計解析に基づく結論の安定性を検討する「感度分析」（用語集参照）の役割を明確にする。

項番号の表記については本補遺を通して、ICH E9 を参照する際は x.y を使用し、本補遺内を参照する際は A.x.y を使用する。

本補遺は、ICH E9 の以下の話題について明確にし、拡張するものである。第一に、ICH E9 は、ランダム化の維持が統計学的検定の確固とした基盤を与えることを示し、ランダム化比較試験において、計画された治療のコースの遵守状況に関わらず被験者を追跡、評価及び解析をする、治療方針の効果と関連した Intention-To-Treat (ITT) の原則を導入した。ITT の原則から生じるいくつかの結論は以下の通りである。一つ目は、リサーチクエストョンに関連したすべての被験者は試験の解析に含めるべきであることである。二つ目は、被験者は割付に従って解析に含めるべきであることである。ITT の原則（ICH E9 用語集）の定義から直接的に読み取ると、三つ目の結論は、計画された治療のコースの遵守状況に関わらず、被験者を追跡及び評価し、その評価を解析に用いるべきであることである。ランダム化は比較試験の基礎であり、解析はランダム化の利点を最大限に活用することを目指すべきであることについては、議論の余地がない。しかし、ITT の原則に従った効果を推定するこ

とが常に規制上及び臨床上の意思決定に最も関連する治療効果を示すことになるのか、という疑問が残る。本補遺で概説するフレームワークは、異なる治療効果を記述するための基盤、及び意思決定のために信頼できる治療効果の推定値を与えるための臨床試験のデザイン及び解析における留意点を与える。

第二に、一般的にデータの取扱い及び「欠測データ」（用語集参照）と関連して検討されている問題について再考する。2つの重要な区別をする。まず、本補遺では、割り付けられた治療の中止と試験の中止を区別する。前者は、**estimand** による試験の目的の詳細な説明において対応すべき中間事象を示す。後者は、統計解析において対応すべき欠測データを生じさせるものである。例えば、がんの臨床試験において、治療をスイッチした被験者や、試験自体が完了したために結果が観察されなかった被験者について考える。前者は中間事象であり、この事象に関連する関心のある臨床的疑問を明確にしなければならない。後者は試験管理上の打ち切りであり、統計解析における欠測データの問題として対応する必要がある。**Estimand** を明確にすることにより、どのデータを収集する必要があるか、つまり、どのデータが収集されなかった場合に統計解析において対処すべき欠測データの問題が生じるかを計画するための根拠が得られる。そして、欠測データにより生じる問題に対処する方法が **estimand** と整合するように選択される。次に、本補遺は、異なる中間事象による異なる結果について強調する。治療の中止、治療のスイッチ、又は追加の薬剤の使用のような事象は、それ以降の変数の測定値が収集できる場合でも、その測定値を意味のない又は解釈が困難なものにしてしまうかもしれない。一方、被験者の死亡後には測定値自体が存在しない。

第三に、解析対象集団の概念に関連した論点をフレームワークの中で考慮する。5.2 節では、優越性試験においては、ランダム化されたすべての被験者を可能な限り含むように定義された最大の解析対象集団に基づいて解析を行うことが強く推奨されている。一方で、臨床試験では多くの場合、同一の被験者に対して繰り返し測定を行っている。何例かの被験者におけるいくつかの計画された測定値について、おそらく意味がない又は解釈が困難であると考えられるためにその測定値を解析から除外することは、最大の解析対象集団から被験者を完全に除外する場合と同様の結果、つまり最初のランダム化が完全には維持されないという結果を招くこととなる。この結果、ランダム化が治療効果に関する仮説検定に与える理論的な利点と、ベースラインでの交絡因子の均衡を保つという実用的な利点が損なわれる可能性がある。さらに、被験者が死亡した場合のように、結果変数の意味のある値が存在しないこともある。5.2 節では、これらの問題について直接的には言及していない。関心のある治療効果は、その推定に含めるべき被験者の集団と解析に含めるべき個々の被験者の観測値の両方を中間事象の発現を考慮して決定するという方法で、注意深く定義することにより明確になる。本補遺では、治験実施計画書に適合した対象集団の解析の意味と役割についても再考する。特に、治験実施計画書に対する違反と逸脱の影響を調べる必要がある際

に、治験実施計画書に適合した対象集団の単純な解析と比較して、より偏りが少なく、より解釈可能な方法で対応することができるかについて考慮する。

最後に、感度分析という項目の下で、安定性（ロバストネス、1.2 節）の概念について拡張した議論を行う。選択された解析手法における仮定に対する推測の感度と、より広い意味での解析手法の選択に対する感度の区別をする。合意された **estimand** について詳細な説明があり、解析手法がその **estimand** に整合し、さらに第三者にも正確に再現可能なほど詳細に事前に定められているのであれば、規制上の関心としては特定の解析に関する仮定からのずれ及びデータの限界に対する感度に焦点を絞ることができる。

本補遺で述べる原則は、有効性及び安全性のいずれについても、治療効果の推定や治療効果に関連した仮説の検定の際に常に関係するものである。これらの原則の主な焦点はランダム化比較試験に置かれるが、単群試験や観察研究にも適用可能である。このフレームワークは、経時測定データ、初回イベント発現までの時間のデータ、再発イベントのデータを含む、あらゆるタイプのデータに対して適用される。検証的試験や、検証的な結論を導くために用いる複数の試験を統合したデータに対しては、この原則の適用についての規制当局の関心はより大きなものとなる。

A.2. 計画、デザイン、実施、解析及び解釈を整合させるフレームワーク

試験計画は順を追って進めるべきである（図 1）。明確な試験の目的を、適切な **estimand** を定義することにより、関心のある重要な臨床的疑問に変換すべきである。**Estimand** は、特定の試験の目的に対する推定の対象（すなわち、「推定されるべきもの」）を定義する（A.3 節）。それにより、適切な推定の方法（すなわち、主とする「推定量」と呼ばれる解析手法、用語集参照）を選択することができる（A.5.1 節）。主とする推定量は特定の仮定に裏付けられている。主とする推定量による推測について、その仮定からのずれに対する安定性を調べるために、同じ **estimand** を対象とした一つ以上の解析として、感度分析を実施すべきである（A.5.2 節）。

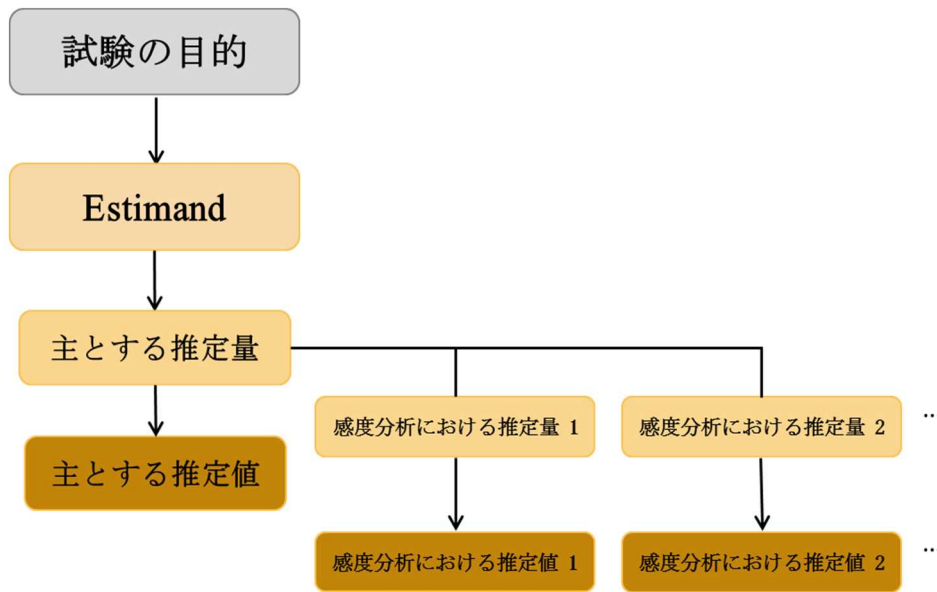


図 1：定められた試験の目的に対して、推定の対象、推定の方法及び感度分析を整合させる

このフレームワークに従うことにより、推定の対象（試験の目的、estimand）、推定の方法（推定量）、その数値的な結果（「推定値」、用語集参照）及び感度分析を明確に区別した、適切な試験の計画を行うことができる。これは、治験依頼者の試験計画及び規制当局のレビューに役立ち、また、試験デザインの適切性について議論する際の治験依頼者と規制当局との間の意思疎通や、試験結果の解釈をより良いものとする。

適切な estimand（A.3 節）の説明は、通常、試験デザイン、実施（A.4 節）及び解析（A.5 節）の特徴に関する主要な決定要因となる。

A.3. Estimand

医薬品開発と薬事承認において中心となる課題は治療効果の存在を確認し、その大きさを推定することである。すなわち、試験治療による結果と、他の治療（すなわち、試験治療を受けなかった場合や、異なる治療を受けた場合）の下で同じ被験者に起きたであろう結果との比較である。Estimand は、定められた臨床試験の目的によって提起される臨床的疑問を反映する治療効果の詳細な説明である。それは、比較されている異なる治療状況下で同じ患者の結果がどのようになるかを集団レベルで要約するものである。推定の対象は臨床試験の前に定義される。それが定義されれば、対象とする治療効果の信頼できる推定を可能にするための試験を計画することができる。

Estimand の説明には特定の要素の詳細な規定が必要である。それらの要素は、臨床的な

留意事項だけでなく、中間事象をどのように臨床的疑問に反映させるのかにも基づいて決めるべきである。A.3.1 節では中間事象について述べる。A.3.2 節では、中間事象について関心のある疑問を説明するためのストラテジーを示す。A.3.3 節では estimand の要素について説明し、A.3.4 節ではその構成における留意事項を述べる。各ストラテジーの違いを理解し、estimand の構成においてどのストラテジーを使用するのかを明確にすることが非常に重要である。

A.3.1. 関心のある臨床的疑問に反映させるべき中間事象

中間事象とは、治療の開始後に発現し、関心のある臨床的疑問に関連する測定値の解釈や測定値の有無に影響を及ぼす事象である。推定すべき治療効果を正確に定義するためには、関心のある臨床的疑問を説明する際に、中間事象に対応する必要がある。

変数の測定値は中間事象の影響を受ける可能性があり、また、中間事象の発現は治療に依存する可能性があるため、治療効果の説明において中間事象を考慮する必要がある。例えば、二人の患者が最初に同じ治療を受け、結果の同じ測定値が得られることがあるかもしれないが、一人の患者は追加の薬剤の投与を受けていたとすると、その二つの測定値が治療について与える情報はこの二人の患者の間で異なる。さらに、患者が追加の薬剤を必要とするかどうか、及び患者が治療を継続できるかどうかは、患者が受ける治療に依存する可能性がある。欠測データとは異なり、中間事象は臨床試験において避けるべき欠点と考えるべきではない。規定された治療の中止、追加の薬剤の使用、及びその他のそのような事象は、臨床試験と同様に実臨床でも起こり得るものであり、これらの事象の発現は関心のある臨床的疑問を定義する際に明確に考慮する必要がある。

測定値の解釈に影響を及ぼす可能性のある中間事象の例として、割り付けられた治療の中止や追加又は他の治療の使用が挙げられる。追加又は他の治療の使用は、背景治療又は併用治療の変更や関心のある治療間のスイッチなど、様々な種類である可能性がある。測定値の有無に影響を及ぼす中間事象の例としては、これら事象が変数自身の一部でない場合の、死亡や下肢の切断（糖尿病性足部潰瘍の症状を評価する場合）のような終末事象が挙げられる。ある特定の臨床的な事象も、その事象が発現すること又はしないことが関心のある主要層（A.3.2 節）を定義する場合には、中間事象になり得る。例として、がん領域における客観的奏効の持続時間に対する治療効果を評価する場合の、客観的奏効を定義する腫瘍の縮小や、当初は感染していなかった被験者のワクチン接種後に発症する感染症の重症度に対する治療効果を評価する場合の、感染症の発症が挙げられる。

中間事象は、治療の中止のように事象自体のみにより定義される場合もあれば、より細か

く定義される場合もある。例えば、毒性による治療中止や有効性の欠如による治療中止のように事象の理由が特定される場合、規定された期間や用量を超える追加の薬剤の使用のように一定の量や程度を必要とする場合、又は変数の評価時点との近さに関連して事象の時期が特定される場合がある。治療の中止のように結果の測定値の解釈に永続的に影響を与える事象もあれば、追加治療の短期的な使用のように一時的にのみ影響を与える事象もある。実際のところ、追加治療や他の治療は多様となり得る。被験者が割り付けられた治療に忍容性がない場合の他の治療として、又は疾患の症状の一次的な増悪を制御するための短期の急性期治療として、被験者が十分な利益を得られない治療を置き換えるか、追加の治療を行うことがある。臨床試験では多くの場合、追加治療や他の治療は、例えば、背景治療、レスキュー薬、併用禁止薬と特定され、それらの異なる役割を区別し、それらを別々に考慮することができる。異なるストラテジーを用いる場合には、異なる中間事象として識別するためにさらなる詳細さが必要である。ストラテジーを選択する必要がある中間事象が、例えば治療継続の失敗だけでなく、その失敗の理由、程度、又は時期にも依存する場合には、臨床試験においてこの追加の情報を正確に定義し記録すべきである。中間事象の記述は、理論的には、長期的な投与のうち一回だけ服用しない、または一日のうちで誤った時間に服用するといった、治療と追跡に関する非常に具体的な詳細を反映することもできる。変数の解釈に影響しないと考えられる場合には、このような特定の基準に中間事象として対応する必要はないだろう。

上述のとおり、**estimand** を構成する際には中間事象を考慮する必要がある。**Estimand** は試験デザインに先立って定義すべきであるため、試験の中止や欠測データの理由（例えば、生存を評価する試験における試験管理上の打ち切り）はそれ自体が中間事象ではない。試験を中止した被験者では、中止の前に中間事象が発現した可能性がある。

A.3.2. 関心のある臨床的疑問の定義における中間事象に対応するためのストラテジー

個別の中間事象について異なる関心のある臨床的疑問をそれぞれ反映するいくつかのストラテジーを以下で説明する。本補遺で用いるようなストラテジーの名称を使うかどうかに関わらず、**estimand** の構成においてストラテジーの選択を一義的に明確にする必要がある。すべての中間事象に同じストラテジーを用いて対応する必要はない。実際には、多くの場合、関心のある臨床的疑問に対応するために、異なる中間事象に対して異なるストラテジーが用いられるだろう。A.3.4 節では **estimand** を構成するためのストラテジーの選択における留意事項を述べる。

治療方針ストラテジー

中間事象の発現は関心のある治療効果の定義において問題としないものとする、つまり、

中間事象の発現の有無に関わらず、関心のある変数の値を用いる。例えば、追加の薬剤の使用という中間事象に対応する方法を規定する場合には、関心のある変数の値はその患者が追加の薬剤を投与されたかどうかに関わらず用いることになる。

患者が治療を継続するかどうか、及び患者が他の治療（例えば、背景治療又は併用治療）を変更するかどうかに関して適用される場合には、その中間事象は比較している治療の一部とみなされる。そのような場合には、これは ICH E9 の用語集（Intention-To-Treat の原則の項）で治療方針により得られる効果とされているものの比較を反映することになる。

通常、治療方針ストラテジーは、終末事象である中間事象に対してはその中間事象の発現後の変数の値が存在しないため適用することができない。例えば、このストラテジーに基づく estimand は、死亡により測定できない変数に関しては構成することができない。

仮想ストラテジー

中間事象が発現しなかった状況を想定する。関心のある臨床的疑問を反映する変数の値は、定義された仮想的な状況において得られたであろう値であるとする。

多種多様な仮想的な状況を想定することができるが、いくつかの仮想的な状況に対しては、他の状況よりも臨床的に又は規制の観点からより関心が高い可能性がある。例えば、実施可能な臨床試験とは異なる状況下での治療の効果を検討することが、臨床的に又は規制の観点から重要であることがある。特に、倫理的な配慮から追加の薬剤を使用可能としないが、関心のある治療効果としてはその追加の薬剤が使用可能でなかった状況下での結果が重要であるかもしれない。中間事象が起こらない又は異なる中間事象が起こるということを仮定する、実際とは大きく異なる仮想的な状況について考える可能性がある。例えば、有害事象を経験し治療を中止する被験者に対して、その被験者が有害事象を経験しないか、又は有害事象があっても治療を継続できる場合について検討する可能性がある。このような仮想的な状況に対する临床上及び規制上の関心は限定的であり、通常、中間事象又はその結果が実臨床と臨床試験とで異なることが予想される理由と程度が、明確に理解できるかどうか依存する。

仮想ストラテジーが提案される場合、どのような仮想的な状況を想定しているのかを明確にする必要がある。例えば、「患者が追加の薬剤を使用していない場合」といった記述は、追加の薬剤が使用可能でないため患者が追加の薬剤を使用していないという仮想的な状況なのか、又は特定の患者は追加の薬剤の使用を必要としないと考えられるためその患者が追加の薬剤を使用しないという仮想的な状況なのか、混乱を招く可能性がある。

複合変数ストラテジー

このストラテジーは関心のある変数に関連している (A.3.3 節)。中間事象自体が患者の結果に関する情報を持つと考え、その中間事象を変数の定義に組み込む。例えば、毒性のために治療を中止した患者について、治療が成功しなかったと考える場合がある。結果変数が成功又は失敗であった場合、毒性による治療の中止は単純に失敗のもう一つの状態とみなされるだろう。しかし、複合変数ストラテジーの使用は結果が二値の場合に限定する必要はない。例えば、身体機能を測定する臨床試験では、変数は連続尺度の結果を用いて構成し、死亡した被験者については身体機能の欠如を反映した値とするかもしれない。複合変数ストラテジーは、患者が死亡した場合など、変数の本来の測定値は存在しない又は意味がないが、中間事象自体が患者の結果について意味のある説明をする場合、*intention-to-treat* の原則に則っていると見なすことができる。

死亡などの終末事象は、複合変数ストラテジーの必要性が最も顕著な例であろう。治療によって生存する場合、生存する患者の様々な測定値におけるその治療の効果に関心があるかもしれないが、関心のある測定値の要約が、生存する患者における連続量の測定値の平均値のみであるというのは適切ではない。連続量の測定値と共に、生存も関心のある結果である。例えば、がん領域の臨床試験における無増悪生存期間は、腫瘍の増殖と生存期間の組み合わせに対する治療効果を測定している。

治療下ストラテジー

このストラテジーでは、中間事象の発現前までの治療に対する反応に関心の対象とする。例えば死亡を中間事象とする場合には「生存下」とするように、このストラテジーの表現は関心のある中間事象によって決まるだろう。

変数が繰り返し測定される場合、中間事象の発現までの変数の値は、すべての被験者に対して設定された同じ時点における値よりも、臨床的疑問に関連していると考えられるかもしれない。中間事象が起こる時点までの関心のある二値の結果の発現についても同様である。例えば、純粋な対症療法について、末期の疾患のある被験者で死亡により対症療法が中止となっても、その治療の成功については死亡前の症状に対する効果に基づいて評価することができる。あるいは、被験者は治療を中止するかもしれないが、状況によってはその患者が治療を受けている間の副作用のリスクを評価することが関心の対象となるかもしれない。

したがって、この場合、関心のある観察期間を中間事象の前までの期間に限定することで、治療下ストラテジーは複合変数ストラテジーと同様に、変数の定義に影響を与えると考えることができる。比較している治療間で中間事象の発現状況が異なる場合には、特別な注意

が必要である (A.3.3 節)。

主要層ストラテジー

このストラテジーは関心のある集団に関連している (A.3.3 節)。対象集団を中間事象が発現するであろう「主要層」(用語集参照) とする場合がある。あるいは、対象集団を中間事象が発現しないであろう主要層とする場合がある。関心のある臨床的疑問はその主要層のみにおける治療効果に関連している。例えば、ワクチンの接種後に感染する患者という主要層における感染の重症度に対する治療効果を評価することが望ましい場合がある。あるいは、毒性により試験治療を継続できない患者がいる可能性があるが、試験治療に忍容性のある患者における治療効果を評価することが望ましい場合がある。

潜在的な中間事象に基づく「主要層別」(用語集参照。例えば、試験薬に割り付けられた場合に治療を中止するであろう被験者集団) と、実際の中間事象に基づいて部分集団を構成すること(割り付けられた治療を中止する被験者集団) を区別することは重要である。多くの場合、試験治療を受けて中間事象を経験する被験者の部分集団は、対照治療を受けて同じ中間事象を経験する被験者の部分集団とは異なるだろう。これらの部分集団の結果を比較することにより定義される治療効果は、異なる治療による効果と異なる被験者の特性による結果の違いを交絡させる可能性がある。

A.3.3. Estimand の要素

以下の要素により estimand を構成し、関心のある治療効果を定義する。

関心のある**治療**の状況と、必要に応じて、比較を行う他の治療の状況(本補遺では以下「治療」と呼ぶ)。これは、個別の介入、同時に行われる介入の組み合わせ(例えば、標準治療への上乗せ)、又は複雑な一連の介入を含む総合的な治療計画で構成される場合がある(A.3.2 節の治療方針及び仮想ストラテジーを参照)。

臨床的疑問に対応する**対象集団**。これは、試験の対象集団全体、ベースラインで測定される特性によって定義される部分集団、又は特定の間事象の発現(又は、状況によっては非発現)によって定義される主要層(A.3.2 節の主要層ストラテジー参照)で説明される。

臨床的疑問に対応するために必要な、各患者について収集する**変数**(又は評価項目)。変数の詳細には、患者が中間事象を経験するかどうかが含まれる場合がある(A.3.2 節の複合変数及び治療下ストラテジー参照)。

治療、対象集団及び変数の詳細な説明は、治験依頼者と規制当局が関心のある臨床的疑問について議論する際に考慮される多くの中間事象に対応しているだろう。**その他の中間事象**に関する関心のある臨床的疑問は、通常、治療方針、仮想、又は治療下ストラテジーを用いて反映される。

最後に、変数の**集団レベルでの要約**を規定し、治療状況間の比較の基盤を与える。

関心のある治療効果を定義する際には、その定義により、観察期間や患者の特性の違いといった潜在的な交絡因子による効果ではなく、治療による効果を特定することを保証することが重要である。

A.3.4. Estimand を構成する際の留意事項

関心のある臨床的疑問とそれに対応する **estimand** は、臨床試験計画の初期段階で特定すべきである。多くの臨床試験では、試験の目的の詳細な説明に、治療の中止や追加又は他の治療の使用を反映させる必要がある。死亡といった終末事象に対応すべき状況もある。試験の目的が、例えば反応があった被験者におけるその反応が得られた期間のように、臨床的な事象に関連してのみ説明できる場合もある。

Estimand の構成においては、個々の治療環境における個々の治療に対する臨床的妥当性を考慮すべきである。対象疾患、臨床的背景（例えば、他の治療が使用可能であるか）、投与方法（例えば、1 回のみ投与、短期投与又は長期投与）及び治療の目標（例えば、予防、疾患修飾、症状のコントロール）についても考慮すべきである。また、治療効果の推定値が意思決定に対して信頼できるものとなるかどうかも重要である。例えば、臨床結果が得られるまでに他のどの治療が行われるかに関わらない臨床結果に対する治療効果に関する臨床的疑問は、追加の薬剤が使用可能でなかった場合の治療効果に関する臨床的疑問とは異なる。状況によっては、どちらも関心のある臨床的疑問を説明するものであることもある。しかし、いずれの場合でも、これらの治療効果を推定するために計画された臨床試験では、医学的に必要とされる場合には追加の薬剤を使用する可能性が含まれることがよくある。前者の臨床的疑問に対しては、追加治療の後の値は意味がある。後者の臨床的疑問に対しては、追加治療の後の値は、その追加の薬剤の影響も反映しているため、直接的には意味がない。**Estimand** の選択を最終決定する前に、信頼できる推定が可能であることに合意すべきである。後者の臨床的疑問では、解析に用いられない観測値を置き換える方法も含めて合意すべきである。

Estimand を構成する際には、関心のある臨床的疑問に対応する治療を明確にする必要が

ある (A.3.3 節)。関心のある治療についての明確な説明は、複数の関連する中間事象をすでに反映していることもある。具体的には、背景治療の変更、併用薬の使用、追加治療又は後続ラインの治療の使用、治療のスイッチ、及び前処置レジメンについて、治療が関心のある臨床的疑問をすでに反映していることもある。例えば、治療を、必要に応じて投与される背景治療 B に追加される介入 A として規定することができる。この場合、背景治療 B の用量の変更は中間事象として考慮する必要はないだろう。しかし、追加治療は中間事象として考慮する必要があるだろう。例えば、治療方針ストラテジーを用いていずれの追加の薬剤の使用も反映させる場合には、治療は、必要に応じた追加の薬剤と共に、必要に応じて投与される背景治療 B に追加される介入 A と規定されることもある。あるいは、治療が介入 A として規定される場合、背景治療の変更及び追加治療の両方が中間事象として扱われるだろう。

中間事象について関心のある臨床的疑問を反映するのに、対象集団や変数の規定を用いるべきかについても検討すべきである。その後、その他の中間事象に対して用いるストラテジーについて検討することができる。通常、意思決定に対して臨床的に適切であり、信頼できる推定を行うことができる **estimand** を構成するには、繰り返しの作業が必要である。特に、臨床的疑問に対して意味のある測定値が得られる **estimand** では、ほとんど仮定を必要とせずに安定性をもって推定することができることもある。他の **estimand** では、正当化するのがより困難である可能性があり、起こり得るずれに対して結果が変わりやすい可能性がある、より特定の仮定を前提とした解析手法が必要となることがある (A.5.1 節)。適切な試験デザインを設計する又は特定の **estimand** に対して十分に信頼できる推定値を導くために、重大な問題が存在する場合には、代替りの **estimand**、試験デザイン及び解析手法を検討する必要がある。

Estimand に関する議論と構成の作業をしなかつたり、単純化しすぎたりすると、臨床試験の目的、試験デザイン、データ収集及び解析手法の間の不整合のリスクが生じる。信頼できる推定値を得られないことから、特定のストラテジーの選択が困難になる可能性はあるが、試験の目的及び関心のある臨床的疑問の理解から順に進めることが重要であり、**estimand** を決定するためにデータ収集及び解析手法を選択するのではない。

試験の実験的な側面についても考慮すべきである。治験実施計画書に基づく被験者の管理 (例えば、忍容性がないことによる用量調整、不十分な反応に対するレスキュー治療、臨床試験の評価の負担) が、実臨床で予期されるものと異なることが正当化される場合、これは **estimand** の構成に反映されるだろう。

Estimand を構成したら、推定の対象を明確かつ一義的に定義すべきである。治療中止という中間事象について考える際には、試験治療を受けた場合に治療を継続することができ

るであろう患者の主要層と、治療を継続している間の効果の、それぞれに基づく関心のある治療効果を区別することが最も重要である。さらに、これらのいずれも、すべての患者が治療を継続できる場合の効果を表すと捉えるべきではない。

前述のとおり、仮想ストラテジーを使用する場合、想定する他の状況と比較して、規制上の意思決定において許容されやすい状況もある。説明される仮想的な状況は、規制当局の意思決定と関連する解釈可能な治療効果の定量化及び実臨床での薬剤の使用を踏まえて正当化すべきである。レスキュー薬が使用可能でなかったとしたら関心のある変数がどのような値をとっていたかという疑問が重要な場合があるかもしれない。対照的に、薬剤の副作用により治療を中止した被験者が治療を継続していたらという仮想的な状況の下で関心のある変数がどのような値を取っていたかという疑問は、科学的及び規制上の関心に対して正当化できないかもしれない。すべての被験者が治療を継続できた場合の効果に基づく関心のある臨床的疑問は、すべての被験者が継続するであろうと仮定する仮想的な状況について綿密に議論をしなければ明確に定義することはできない。治療に対して忍容性がないことは、それ自体が好ましい結果を達成できないことの根拠となり得る。

治療方針ストラテジーに基づく **estimand** を用いて有益な効果の特徴づけることは、規制上の意思決定を支持するためにより一般的に受け入れられやすいかもしれない。特に、他のストラテジーに基づく **estimand** に、より臨床的な関心があると考えられるものの、信頼できる推定値又は安定性のある推測を支持すると合意できるような主とする推定量及び感度分析の推定量が特定できない場合にも受け入れられるだろう。治療方針ストラテジーに基づく **estimand** は、それでもなお意味のある治療効果の信頼できる推定値を得られる可能性がある。この状況では、臨床的により重要と考えられる **estimand** も含め、結果として得られた推定値を、その特定の手法に関する試験デザインや統計解析の観点からの限界についての議論と共に、提示することが推奨される。治療方針ストラテジーに基づいて **estimand** を構成する場合、推測はそのストラテジーを使用するそれぞれの中間事象に関する追加の **estimand** や解析を定義することで補足することができる。例えば、症状スコアに対する効果と、各治療の下で追加の薬剤を使用した被験者の割合の両方を比較することが考えられる。同様に、治療下ストラテジーを使用した **estimand** を構成する場合には、通常、中間事象の発現までの時間の分布に関する追加の情報を併せて提示すべきであり、主要層ストラテジーに基づく **estimand** では、可能であればその層に含まれる患者の割合に関する情報を併せて提示することが有益であろう。

非劣性又は同等性の評価を目的とした規制上の意思決定を支持するための **estimand** の構成においては、優越性に対する **estimand** の選択とは異なる注意が必要となる場合がある。ICH E9 で説明されているとおり、規制当局が意思決定の際に直面する問題は、非劣性

又は同等性の試験に基づく場合と優越性試験に基づく場合とでは異なる。3.3.2 節では、このような試験は本質的に保守的ではないことについて述べられ、治験実施計画書の違反と逸脱、治療計画への不遵守及び試験の中止を最小化することの重要性を示している。5.2.1 節以降では、このような試験では一般的に最大の解析対象集団（FAS）の結果が保守的ではないことと、その役割について慎重に考えるべきであることを述べている。治療方針ストラテジーを用いて一つ以上の中間事象を考慮するように構成された estimand では、非劣性及び同等性の試験に対して、ITT の原則に基づく FAS の解析に関連した問題と同様の問題が生じる。割り付けられた治療の中止やその他の薬剤の使用の後は、最初に割り付けられた治療の類似性とは関係のない理由で、両治療群間の反応がより類似して見える可能性がある。Estimand は治療群間の差を小さくする可能性がある中間事象（例えば、治療の中止や追加の薬剤の使用）に直接的に対応できるように構成することもあり得る。ストラテジーを選択する際には、同一又は類似の有効成分を含む治療間に差が存在するかどうかを検出するように計画された試験（例えば、バイオ後続品と対照バイオ医薬品との比較）と、有効性の証拠を確立し定量化するために非劣性又は同等性の仮説が用いられる試験とを区別することが重要だろう。Estimand は、規制上の意思決定に対して適切ならば、治療間の差を検出する感度を優先する治療効果を対象として構成することができる。

A.4. 試験デザイン及び実施への影響

試験デザインは、試験の目的を反映する estimand と整合させる必要がある。一つの estimand に対して適切な試験デザインは、他の潜在的に重要な estimand に対しては適切ではないかもしれない。治療効果の定量化の基盤となる estimand の明確な定義は、試験デザインに関連する選択のための情報を与えるべきである。これには、対象集団を規定する選択基準と除外基準、治験実施計画書で許可される薬剤や禁止される薬剤を含む治療、患者管理とデータ収集に関する他の側面に関する規定が含まれる。例えば、特定の中間事象が発現したかどうかに関わらない治療効果を理解することに関心がある場合には、すべての被験者の変数を収集する試験デザインが適切である。一方で、規制上の意思決定を裏付けるために必要な estimand において中間事象の後の変数の収集が必要でないならば、他の estimand のためにこのようなデータを収集することの利点は、データ収集することにより複雑になるあらゆる点及びその潜在的な欠点と比較して検討すべきである。

中間事象の特徴、発現状況及び発現時期の情報に関するデータも含め、推定の裏付けに関連するすべてのデータを収集するように努めるべきである。データは常に収集できるとは限らない。被験者を被験者自身の意志に反して臨床試験に留めることはできない。また、生存を評価する試験における管理上の打ち切りのように、一部の被験者のデータの欠測が試験デザインにより回避できない場合もある。反対に、治療の中止、治療のスイッチ、又は追

加の薬剤の使用といった中間事象の発現は、測定値は適切ではない可能性があるものの、中間事象発現後に変数を測定できないことを意味するものではない。死亡のような終末事象では、変数を中間事象の後に測定することはできないが、これらのデータは一般的に欠測と見なすべきではない。

Estimand を評価するのに必要ないずれかのデータが収集されなかった場合には、それに続く統計的推測において欠測データの問題が生じることとなる。統計解析の妥当性は確認できない仮定に基づいていることがあり、欠測データの割合によっては、結果の安定性が損なわれる可能性がある (A.5 節)。収集しようとしていたデータが欠測となった理由に関する有益な情報の取得を前向きに計画することは、中間事象の発現と欠測データを区別するのに役立つ可能性がある。これは同様に解析を改善し、より適切な感度分析の選択に繋がる可能性もある。例えば、「追跡不能」は、より正確に「有効性の欠如による治療の中止」と記録されるかもしれない。それが中間事象として定義されている場合には、それを考慮するために選択されたストラテジーによって反映され、欠測データの問題としては扱われない。欠測データを減らすために、被験者を臨床試験に留まらせるような方策をとることもできる。しかしながら、実臨床でも通常は発現するであろう中間事象を減らす又は回避する方策をとることは、試験の外部妥当性を低下させるリスクがある。例えば、毒性の影響を軽減するために、試験の対象集団を選択する又は漸増法や併用薬を使用することは、それが実臨床においても行われるものでなければ適切ではないだろう。

ランダム化と盲検化が比較試験の基盤であることに変わりはない。2.3 節では偏りを回避するための計画上の技法について述べている。特定の **estimand** では、導入期を用いたデザインやエンリッチメントデザイン、ランダム化治療中止デザイン又は漸増デザインといった試験デザインの利用が必要となる、又はそれらデザインの恩恵を受けるかもしれない。試験治療群又は対照群への割付を行う前に、導入期を用いて治療に忍容性がある被験者の主要層を特定しようとするに関心があるかもしれない。規制当局と治験依頼者の間の協議では、提案された導入期が対象集団を特定するのに適切であるかどうか、またそれに続く試験デザインの選択 (例えば、ウォッシュアウト期間、ランダム化) が対象とする治療効果の推定とそれに関連する推測の裏付けとなるかどうかを検討する必要がある。これらの検討により、このような試験デザインの使用や特定のストラテジーの使用が制限されることになるかもしれない。

関心のある治療効果の詳細な説明は、被験者数の計算に対して情報を与えるべきである。暗黙的に又は明示的に異なる **estimand** に基づいて推定された治療効果又はばらつきが報告された可能性のある過去の臨床試験を参照する場合には、特に注意すべきである。すべての被験者が解析に対する情報に寄与し、中間事象を反映するストラテジーの影響が対象と

する効果の大きさと予想される分散に含まれるのであれば、通常は、算出された被験者数に、想定される試験の中止例の割合に応じた上乘せをする必要はない。

7.2 項は複数の臨床試験のデータの要約に関する問題について述べている。関心のある変数について一貫した定義を持つことの必要性が強調されており、これは **estimand** の構成に拡張することができる。したがって、臨床試験プログラムを通じたエビデンスの統合を行うことが計画段階で想定される状況では、適切な **estimand** を構成し、治験実施計画書に含め、統合に寄与する試験のデザインに関する選択に反映させるべきである。同様の留意点はメタアナリシスのデザイン、非劣性の限界値の決定のための完了した試験から推定された効果の大きさの利用、又は単群試験の解釈のための外部対照群の使用についても適用される。それぞれの試験のデータの表示や統計解析で扱われている **estimand** を考慮、特定することなく、データ間の単純な比較や複数の試験のデータの統合を行うことは、誤解を招く可能性がある。

臨床試験には、複数の **estimand** に変換され、それぞれが統計的検定や推定に関連する、複数の目的があることがより一般的である。これにより生じる多重性の問題には対応する必要がある。

A.5. 試験の解析への影響

A.5.1 主とする推定

対照と比較した治療効果のための **estimand** は、試験治療群の被験者の結果と試験治療群に類似した対照群の被験者の結果の比較により推定される。規定した **estimand** に対し、信頼できる解釈の基盤となり得る推定値を提示することができる、整合した解析手法又は推定量を用いるべきである。解析手法は信頼区間の算出及び統計的有意性の検定の裏付けともなる。解釈可能な推定値が得られるかどうかに関して考慮すべき重要な点は、解析において必要となる仮定の程度である。**Estimand** とそれに対応する主とする推定量及び感度分析における推定量と共に、必要となる重要な仮定を明確に示すべきである。その仮定は正当化できるべきであり、ありそうにない仮定は避けるべきである。仮定からのずれの可能性に対する結果の安定性は、同じ **estimand** に整合した感度分析によって評価されるべきである (A.5.2 節)。多くの仮定又は強い仮定を必要とする推定には、より広範な感度分析が必要である。仮定からのずれの影響を感度分析によって包括的に確認することができない場合には、その特定の **estimand** と解析手法の組み合わせが意思決定において受け入れられない可能性がある。

すべての解析手法には仮定が必要であり、異なる解析手法は、同じ **estimand** に整合していても、異なる仮定を必要とするかもしれない。それでも、ある種の仮定は、概説した異なるストラテジーそれぞれを用いた **estimand** に整合したすべての解析手法に固有のものである。例えば、仮想的な状況で観測されたであろう結果を予測するための方法論や、主要層ストラテジーにおける適切な対象集団を特定するための方法論に対する仮定がある。中間事象の発現を反映するために用いる様々なストラテジーに関連したいくつかの例を以下に示す。ここで強調する問題は、**estimand**、主とする解析及び感度分析について合意するにあたって治験依頼者と規制当局との間で議論する際の重要な要素となる。

ある中間事象に治療方針ストラテジーで対応する解析では、試験デザイン及び実施状況によって、より強い仮定又は弱い仮定が必要となる場合がある。多くの被験者がそれぞれの中間事象（例えば、治療の中止）の発現後でも追跡された場合には、欠測データに関する残された問題は比較的軽微なものだろう。対照的に、中間事象の後に観察が終了している場合は、このストラテジーに対して望ましくないことは明白であり、中止した被験者の（未観測の）結果が治療を継続した被験者の（観測された）結果と類似しているという仮定は、多くの場合、ありそうにないものである。欠測データを取り扱うための別の方法を正当化する必要があり、感度分析が必要となる可能性がある。

仮想ストラテジーに整合した解析では、実際に観測されたものとは異なる結果、例えば、実際にはレスキュー薬が投与されたのに、投与されなかった場合の結果などが含まれる。レスキュー薬を投与する前の観測値とレスキュー薬を必要としなかった被験者における観測値は、強い仮定の下でのみ有益である可能性がある。

複合変数ストラテジーは、中間事象の発現を結果の構成要素と考えることにより、中間事象の発現後のデータに関する統計的仮定を避けることができる。推定における仮定よりも、推定された治療効果の解釈に関して潜在的な懸念がある。**Estimand** を解釈可能なものとするため、中間事象の発現による失敗に評点が付けられるのであれば、患者にとっての有益性の欠如を意味のある形で反映させるべきである（例えば、死亡は有害事象による治療の中止とは異なる形で反映されるだろう）。

治療下ストラテジーに基づいて構成される **estimand** は、中間事象の発現時点までの結果が収集されていれば推定が可能である。これに対しても、重要な仮定は解釈に関するものである。例として治療の中止について考える。治療下での結果では改善しているかもしれないが、治療によって中止が誘発されたり遅延されたりすることにより治療期間が短縮又は延長する可能性があり、治療下での結果と治療期間の効果の両方を臨床的有益性の解釈と評価において考慮すべきである。

主要層ストラテジーに整合した解析では、通常、強い仮定が必要である。例えば、被験者のベースライン特性から推論する主要層別の方法があるが、この推論の正しさを評価することは困難だろう。しかし、単純な方法ではこの問題は回避できない。例えば、治療と中間事象は無関係であると仮定し、試験治療において中間事象が認められなかった被験者と対照治療において中間事象が認められなかった被験者とを単純に比較することの正当性を説明することは非常に困難である。

中間事象に適切に対応する **estimand** を定義し、推定に必要なデータの収集に努めた場合でも (A.4 節)、例えば生存を評価する試験での試験管理上の打ち切りのように、いくつかのデータは欠測となるかもしれない。適切なデータの収集に失敗することと、中間事象の発現により意味がなくなったデータを収集しない又は収集しても解析に用いないという選択をすることを混同してはならない。例えば、治療方針ストラテジーに基づく **estimand** に対応するために収集しようとしていた治療中止の後のデータは、収集されなければ欠測となる。しかし、別のストラテジーにおいては同じ時点のデータが意味のないものとなる可能性があり、このような **estimand** の目的に対しては収集されなかったとしても欠測とはならない。データの収集に努めたが収集できなかった場合には、統計解析において欠測データを取り扱うために仮定をおくことが必要になる。欠測データの取扱いは臨床的にもっともらしい仮定に基づくべきであり、また可能であれば、**estimand** の記述の中で用いられているストラテジーによって導かれるべきである。用いられる方法は、個々の被験者及び他の類似した被験者の観測された共変量及びベースライン後のデータに基づくかもしれない。類似した被験者を特定する基準は、中間事象の発現の有無を含むかもしれない。例えば、治療を中止し、その後のデータを収集しなかった被験者に対するモデルでは、治療を中止したがデータの収集は継続した他の被験者のデータを用いるかもしれない。

A.5.2. 感度分析

A.5.2.1. 感度分析の役割

特定の **estimand** に基づく推測は、データの限界と、主とする推定量に対する統計モデルで用いる仮定からのずれに対して安定しているべきである。この安定性は感度分析によって評価される。感度分析は、規制当局の意思決定及び医薬品の情報の表示において重要となるすべての **estimand** の主とする推定量について計画すべきである。これは治験依頼者と規制当局とで議論し、合意する話題となり得る。

主とする推定量を裏付ける統計的な仮定は明記すべきである。そして、主とする推定量に

よって算出された推定値がその仮定からのずれに対して安定しているかどうかを確認する目的で、同じ **estimand** を対象とした一つ以上の解析を事前に規定すべきである。これは、統計学的又は臨床的有意性の観点から、結果の解釈を変えるような仮定からのずれの程度により特徴付けられることがある（例えば、**tipping point analysis**）。

仮定からのずれに対する安定性を調べる目的で評価を行う感度分析とは異なり、試験のデータをより十分に評価し、理解するために実施される他の解析は「補足的解析」（用語集、A.5.3 節）と呼ばれる。関心のある主要な **estimand** が治験依頼者と規制当局の間で合意され、主とする推定量が事前に明確に規定され、感度分析により主とする推定量から得られた推定値が解釈において信頼できることが確認されるのであれば、一般的には補足的解析は評価において優先順位がより低いものとすべきである。

A.5.2.2. 感度分析の選択

感度分析を計画、実施する際に、主とする解析の多くの要素を同時に変更することは、結果の違いの原因になる仮定があったとしても、それがどの仮定であるかを特定することを困難にする可能性がある。したがって、異なる複数の仮定に基づく異なる解析の結果を単に比較するのではなく、異なる解析の前提となる仮定の変更を特定した体系的な方法で行うことが望ましい。複数の仮定を同時に変えた解析の必要性は状況に応じて検討する必要がある。確認できる仮定と確認できない仮定を区別することは、異なる解析の解釈と適切性について評価する際に役立つだろう。

欠測データに対する感度分析の必要性は確立されており、このフレームワークにおいてもその重要性は変わらない。欠測データは個別の **estimand** に対して定義し、検討すべきである（A.4 節）。特定の **estimand** に対して欠測となっているデータと、特定の **estimand** に対して直接的には意味のないデータの区別により、感度分析で確認すべき仮定は異なる。

A.5.3. 補足的解析

試験結果の解釈は、対応する推定値が感度分析によって安定していることを確認した上で、合意された個々の **estimand** に対する主とする推定量に焦点を当てるべきである。ある **estimand** に対する補足的解析は、治療効果の理解にさらなる考察を与えるために、主とする解析及び感度分析に加えて実施することができる。通常それらが試験結果の解釈において担う役割はより小さい。補足的解析の必要性和有用性は、それぞれの試験において考慮すべきである。

5.2.3 節では、FAS の解析と治験実施計画書に適合した対象集団 (PPS) の解析との相違を明示的な議論と解釈の対象にできるように、通常両方の解析を計画することが適切であると述べている。また、FAS と PPS に基づく解析結果の一貫性により試験結果の信用度は高くなると述べている。5.2.2 節では、PPS に基づく結果には重大な偏りが生じている可能性があることも説明している。本補遺で示したフレームワークにおいては、PPS の解析に整合するような適切な estimand を構成できない可能性がある。上述したとおり、PPS の解析では、例えば試験治療に忍容性があり治療を続けることが可能な被験者集団といった、いずれの主要層における効果を推定するという目標も達成しない。なぜなら、異なる治療を受けている同様の被験者集団を比較することにならない可能性があるからである。

必ずしも中間事象が発現しているわけではなく、例えば規定した期間外の来院など、治験実施計画書の違反及び逸脱がある場合に、その被験者を PPS から除外する場合がある。同様に、治験実施計画書からの逸脱はなく、死亡といった中間事象が発現する可能性もある。治験実施計画書の違反及び逸脱が中間事象に関わらず、測定値の解釈や存在に影響を及ぼす可能性のある事象については、estimand の記述において考慮する。PPS の解析に関連した目的に対して、より適切に対応するために、整合した解析手法と共に estimand を構成できるかもしれない。その場合には、PPS の解析はさらなる考察を与えるものではない場合がある。

A.6. Estimand と感度分析の記載

治験実施計画書では、試験の主要な目的に対応する主要な estimand を定義し、明確に規定すべきである。治験実施計画書と統計解析計画書では、主要な estimand と整合し、主要解析を導く主とする推定量を、仮定からのずれに対する安定性を調べるための適切な感度分析と共に事前に規定すべきである。規制上の意思決定を支持する可能性がある副次的な試験目的 (例えば、副次変数に関するもの) に対する estimand についても、それぞれ対応する主とする推定量と適切な感度分析と共に定義し、明確に規定すべきである。探索的な目的のために追加の試験目的が検討され、それが estimand の追加に繋がるかもしれない。

主要な estimand の選択は、通常、試験デザイン、実施及び解析の特徴の主要な決定要因となる。通常行われているように、これらの特徴は治験実施計画書に十分に記載すべきである。副次的な estimand に重要な関心がある場合には、必要に応じてこれらの留意事項を拡張し、同様に記載すべきである。これらの側面以外の試験デザイン、実施及び解析に関する従来の留意事項はこれまでと同様である。

何を推定しようとしているのかを明確にすることは治験依頼者には役立つが、個々の探索的な目的に対する **estimand** を記載することは規制上の要件ではない。

主とする解析、感度分析及び補足的解析の結果は、それぞれの解析が事前に規定されていたか、試験の盲検性が維持されている間に導入されたのか、又は事後的に実施されたのかを明記した上で、治験総括報告書において体系的に報告されるべきである。それぞれの治療群における中間事象の発現数及び発現時期に関する要約は報告する必要がある。

試験実施中に **estimand** を変更すると問題が生じ、試験の信頼性が低下する可能性がある。試験の計画段階で予見されず、試験実施中に明らかになった中間事象に対応する際には、解析手法の選択だけでなく、**estimand** に対する影響、すなわち推定される治療効果の説明に対する影響、及び試験結果の解釈についても考察すべきである。通常、**estimand** の変更は治験実施計画書の改定に反映すべきである。

用語集

Estimand

試験の目的によって提起される臨床的疑問を反映する治療効果の詳細な説明。比較されている異なる治療状況下において同じ患者の結果がどのようになるかを集団レベルで要約するものである。

推定値 Estimate

観測された臨床試験データに基づいて、推定量により算出される数値。

推定量 Estimator

観測された臨床試験データから推定値を算出するための解析手法。

中間事象 Intercurrent Event

治療開始後に発現し、関心のある臨床的疑問に関連する測定値の解釈や測定値の有無に影響を及ぼす事象。推定すべき治療効果を正確に定義するためには、関心のある臨床的疑問を説明する際に、中間事象に対応する必要がある。

欠測データ Missing Data

規定した **estimand** の解析に対して意味があると考えられるが、収集されなかったデータ。存在しないデータや中間事象の発現により意味があると思えないデータとは区別す

べきである。

主要層別 Principal Stratification

すべての治療において中間事象が発現する可能性に応じて被験者を分類すること。治療群が二つの場合には、ある中間事象に対して、いずれの治療でも中間事象が発現しないであろう被験者集団、治療 A では発現するが治療 B ではしないであろう被験者集団、治療 B では発現するが治療 A では発現しないであろう被験者集団、及びいずれの治療でも発現するであろう被験者集団の、四つの主要層が考えられる。本補遺では、**主要層**は主要層別によって定義されるいずれかの層（又は層の組み合わせ）を指す。

感度分析 Sensitivity Analysis

主とする推定量に対する統計モデルにおける仮定からのずれとデータの限界に対する推測の安定性を評価することを目的として実施される一連の解析。

補足的解析 Supplementary Analysis

主とする解析及び感度分析に加えて、治療効果の理解にさらなる考察を与えることを目的として実施される解析の総称。