

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

(案)

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. 解説	6
A.3.2. 中間事象に対応するためのストラテジー	8
A.3.3. Estimand の構成	11
A.3.3.1. 一般的留意事項	11
A.3.3.2. 治療背景及び試験の実験的側面に関する留意事項	12
A.4. 試験デザイン及び実施への影響	14
A.5. 試験の解析への影響	16
A.5.1. 主とする推定	16
A.5.2. 感度分析	18
A.5.2.1. 感度分析の役割	18
A.5.2.2. 感度分析の選択	18
A.5.3. 補足的解析	19
A.6. Estimand と感度分析の記載	20
A.7. 一般的な例	21
A.7.1. 中間事象が一つの例	22
A.7.2. 中間事象が二つの例	26
用語集	29

1 A.1. 目的と適用範囲

2
3 患者や処方医が治療を選択する際の選択肢を適切に提示するためには、薬剤の有効性に
4 ついて明確に説明すべきである。この説明は、個々の患者によって治療に対する反応が様々
5 であるため、複雑なものとなる。薬剤に忍容性があり、服薬計画を遵守できる被験者もい
6 れば、そうではない被験者もいる。また、併用薬の投与量の変更や、追加の薬剤の投与が
7 必要となる被験者もいれば、そうではない被験者もいる。忍容性や服薬計画の遵守状況、
8 追加薬剤が必要か否か等をどのように考慮するかによって、治療効果を定量化する複数の
9 方法を想定することができる。したがって、説明されている治療効果を正確に理解しなけ
10 れば、その大きさや意義を誤解する危険性がある。

11
12 多くの場合ランダム化比較試験として実施される検証的試験は、治療効果を定量化し、
13 規制上の意志決定の裏付けとなる有効性及び安全性の証拠を提示するために実施される。
14 ランダム化比較試験は、ベースラインにおける交絡は回避できるはずだが、臨床現場と同
15 様に、試験においても治療効果の説明と解釈を複雑にする特定の事象が発現することがあ
16 る。本補遺では、これらの事象を中間事象（用語集参照）と呼ぶ。中間事象には特に、代
17 替治療（レスキュー薬、治験実施計画書における併用禁止薬、又は後続ラインの治療）の
18 使用、治療の中止、治療の切り替え、状況によっては死亡といった終末事象が含まれる。

19
20 先に中間事象の発現に対応することなく、有効性及び安全性の変数、データ収集の基準、
21 及び統計解析の手法を選択し定義することは、推定される治療効果の曖昧さや、推定され
22 る治療効果と試験の目的との潜在的な不整合に繋がる。正しい順序はこれとは逆である。
23 試験の目的を明確にし、試験計画の段階で関心のある治療効果を説明する際に中間事象に
24 ついて明確に考慮することで、試験デザイン、データ収集及び統計解析に関する選択肢を
25 与えるべきである。

26
27 本補遺では、試験の目的とそれに適切に対応する試験デザイン及び推定や仮説検定の方
28 法をつなぐ、構造化されたフレームワークを提示する。このフレームワークは、試験の目
29 的を推定すべき治療効果の正確な定義に変換する **estimand**（用語集参照）の概念を導入し
30 ている（A.3 節）。これは、臨床試験の計画、実施、解析及び解釈に関わる専門分野間、及
31 び治験依頼者と規制当局との間での、臨床試験で取り組むべき関心のある治療効果に関す
32 る議論を促進することを目的とする。**Estimand** に整合した統計解析は仮定とデータの限界
33 に関連しており、その影響は感度分析（用語集参照）によって検討することができる。本
34 補遺では、感度分析の定義と役割を明確にする。なお、ICH E9 を参照する際は **x.y** を使用
35 し、本補遺内を参照する際は **A.x.y** を使用する。

37 本補遺は、ICH E9 をいくつかの点から明確にし、拡張するものである。

38

39 第一に、ICH E9 は、治療方針の効果、すなわち予定した試験治療のコースの遵守状況に
40 関わらない、ベースライン時に最初に割付けられた治療の効果と関連して、
41 **intention-to-treat (ITT)** の原則を導入し、ランダム化の維持が統計的検定の確固とした基
42 盤を与えることを示した。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
43 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、
44 治療方針の効果を理解することが常に規制上及び臨床上の意志決定に最も関連する治療効
45 果を目的としているものな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、
46 その他の治療効果、及び意志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
47 床試験のデザイン及び解析の留意点について、議論するための基盤を与える。

48

49 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）に含めて検討されてい
50 る問題について再考する。治療の中止や切り替え、又はレスキュー薬の使用といった中間
51 事象は、場合によっては、その発現後に変数の測定値を収集することが可能であっても、
52 その測定値を意味がない又は解釈が困難なものにしてしまうことがある。死亡の場合には、
53 被験者が死亡した後の測定値は存在しない。一方で、ICH E9 は、治療を中止した被験者が
54 追跡不能となった場合に **ITT** の原則を達成することが困難であることについて言及してい
55 る。本補遺では、割付けられた治療の不遵守又は割付けられた治療の中止と、試験の中止
56 の重要な違い、さらに、存在するが収集されなかった測定値と、存在しない又は存在し得
57 ない測定値との重要な違いを考慮する。**Estimand** を明確にすることにより、どのデータを
58 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対処すべき欠測デ
59 ータの問題があるかを計画するための根拠が得られる。その後、欠測データにより生じ
60 る問題に対処する方法が、選択された **estimand** と整合するように選択されることになる。

61

62 第三に、解析対象集団の概念を、提案するフレームワークにおいて考慮する。5.2 節では、
63 優越性試験においては、ランダム化された全ての被験者を可能な限り含むように定義され
64 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
65 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
66 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
67 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
68 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
69 という結果を招くこととなる。さらに、被験者が死亡した場合のように、結果変数の意味
70 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
71 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を決定し、さらに中
72 間事象の発現を考慮して解析に含めるべき個々の被験者の観測値を決めるというように、

73 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
74 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
75 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による単純な解析と比
76 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
77 る。

78

79 最後に、感度分析という項目の下で、安定性（ロバストネス）の概念について拡張した
80 議論を行う。特に、特定の解析における特定の仮定に対する推測の感度と、より広い意味
81 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
82 があり、統計解析がその **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
83 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
84 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。

85

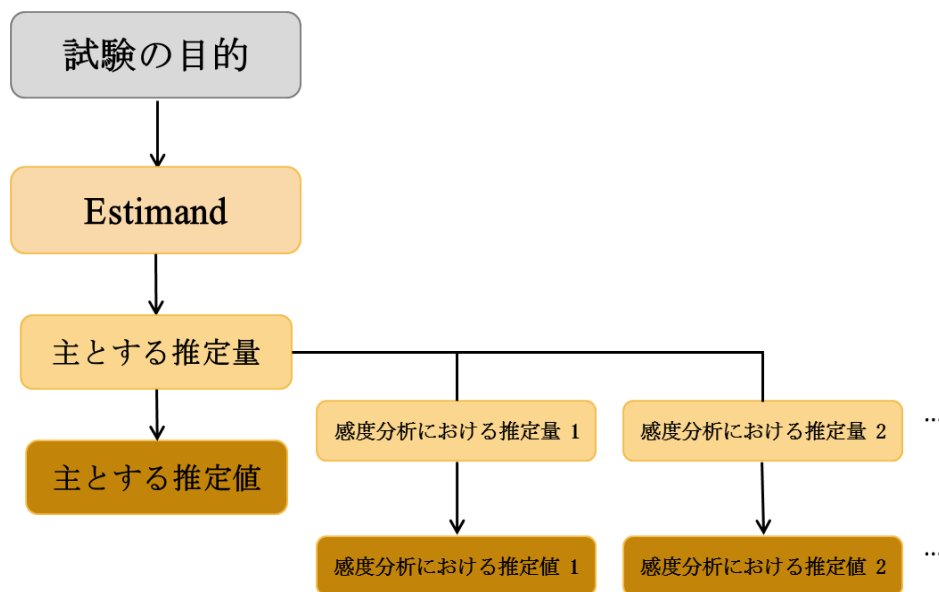
86

87 **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**

88

89 臨床試験の整合性と明確さを高めるために、試験計画は順に沿って進めるべきである（図
90 1）。明確な試験の目的を、適切な **estimand** を定義することにより、関心のある重要な科学
91 的疑問に変換すべきである。**Estimand** は、対象集団、変数、中間事象の取扱い及び集団レ
92 ベルの変数の要約を規定することにより、特定の試験の目的に対する推定の対象（すなわ
93 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
94 わち、主とする推定量と呼ばれる解析手法）を選択することができる。主とする推定量は
95 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の、その仮定からのずれに対
96 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
97 を実施すべきである（A.5 節）。

98



99

100 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

101

102 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
 103 法（推定量、つまり推定値を得る方法、用語集参照）及び感度分析を明確に区別した、適
 104 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
 105 ューに役立ち、また、医薬品の承認審査において試験デザインの適切性や試験結果の解釈
 106 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をより良いものとする。

107

108 通常は、estimand、つまり答えを得ようとしている科学的疑問を決めるために推定量を
 109 選択するのではなく、適切な順に沿って試験計画を進めることが重要である。

110

111 適切な estimand（A.3.3 節）の内容は、多くの場合、試験デザイン、実施（A.4 節）及び
 112 解析（A.5 節）の特徴に関する主要な決定要因となる。

113

114

115 A.3. Estimand

116

117 A.3.1 解説

118

119 医薬品開発と承認審査において中心となる課題は治療効果の定量化である。すなわち、
 120 試験治療による結果と、異なる治療条件下（例えば、試験治療を受けなかった場合や、異
 121 なる治療を受けた場合）で、同じ被験者に起きたかもしれない結果との比較である。関心
 122 のある変数の値と中間事象の発現はどちらも受けた治療に依存する可能性があるため、関

123 心のある変数における治療効果を記述する際は中間事象を考慮する必要がある。Estimand
124 によって規定される治療効果の定義においては、中間事象の発現後の変数の値が適切なも
125 のか、また、(受けた治療に関連している可能性がある) 中間事象の発現の有無をどのよう
126 に考慮するかを検討すべきである。

127

128 より正式には、estimand は関心のある科学的疑問に対する答えを得るために何を推定す
129 る必要があるのかを詳細に定義するものである。Estimand の記述には次の四つの要素が含
130 まれる。

- 131 A. 対象集団、すなわち、科学的疑問の対象となる患者
- 132 B. 科学的疑問の答えを得るために必要な、各患者について得るべき変数 (又は評価項目)
- 133 C. 関心のある科学的疑問を反映するために、中間事象をどのように考慮するかという説
134 明
- 135 D. 必要に応じて治療条件間の比較のための基礎となる、集団レベルでの変数の要約
136 これらの要素により、関心のある治療効果を定義する estimand を記述する。

137

138 多くの場合、対象集団は、治験実施計画書における選択/除外基準に基づいて組み入れ
139 られる患者に反映される。また、それらの患者のうち、起こり得る中間事象の観点から定
140 義される患者層、例えば治療計画を遵守するであろう患者層に関心がある場合もある。

141

142 一般的に、変数は、測定値 (例えば、血圧の測定値)、測定値の関数 (例えば、HbA1c
143 のベースラインから一年後までの変化量)、又は臨床結果に関連する量 (例えば、死亡まで
144 の期間、入院までの期間、再発回数) で構成される。また、変数が中間事象を包含する場
145 合もあるかもしれない。例えば、治療の中止といった中間事象に対して、変数に治療中止
146 の前までの測定値を用いる (例えば、中止までの HbA1c の曲線下面積や、治療中に血圧値
147 が維持されていた週数を用いる)、又は複合的な変数を用いる (例えば、無効又は治療中止
148 を治療失敗として定義する) 場合である。

149

150 関心のある科学的疑問を反映するように、起こり得る中間事象をどのように考慮するか
151 を規定しておく必要がある。複数の種類の中間事象が存在し、変数の解釈に影響を及ぼす
152 可能性がある。例えば、計画された血圧の測定時点よりも前に被験者が死亡した場合には、
153 その被験者の血圧値は観測されないことになる。被験者が、治療に加えてレスキュー薬を
154 使用した場合には、血圧値は観測され得るが、その値は治療とレスキュー薬を併用した効
155 果を反映することになる。被験者が毒性により治療を中止した場合には、血圧値は観測さ
156 れ得るが、その値は治療を受けていない時に治療効果がないことを反映することになる。
157 検討すべき中間事象の組は、特定の治療の設定や試験の目的に依存する。レスキュー薬の
158 使用を例とすると、治療の効果と中間事象 (この場合はレスキュー薬の使用) の複合的な

159 効果、及び仮想的であるかもしれないが、中間事象が発現しないという状況における治療
160 の効果、という二つの異なる説明が考えられる。A.3.2 節では、関心のある科学的疑問に最
161 も整合する estimand を構成する際の、中間事象に対応するための異なる複数のストラテジ
162 ーを説明する。

163

164 四つ目の要素は、集団レベルでの変数の要約指標である。例えば、HbA1c のベースライ
165 ンから一年後までの変化量の平均値や、反応に関する特定の基準を満たす被験者の割合が
166 あげられる。治療間の比較の場合、要約指標は、二つの異なる治療条件下における、HbA1c
167 のベースラインから一年後までの変化量の平均値の差、特定の基準を満たす被験者の割合
168 の差又は比といったものになる。

169

170

171 **A.3.2 中間事象に対応するためのストラテジー**

172

173 A.3.1 節で説明した estimand の要素 A から D は相互に関連しており、独立して考えるべ
174 きではない。Estimand の記述は、関心のある科学的疑問を起り得る中間事象にどのよう
175 に反映させるかを示さなければ、完全なものとはならない。少なくとも五つのストラテジ
176 ーが考えられるだろう。ストラテジーは、複数の異なる中間事象に対応するために、単独
177 で又は組み合わせて用いることができる。その他の estimand の要素とともに、どのよう
178 に中間事象に対応するかを選択することで、対象とする治療効果を記述する。A.7 節では、一
179 つ以上の中間事象を考慮して estimand を構成する際の、これら五つのストラテジーの使用
180 例を示す。

181

182 それぞれのストラテジーの適切性は、治療背景や試験の実験的側面に依存する。さらに、
183 それぞれの実験状況によって、これらのストラテジーを使って構成され、意志決定におい
184 て信頼できると考えられる特定の estimand の推定値を得ることができる場合とできない
185 場合がある。これらの留意事項は、A.3.3 節、A.3.4 節、A.4 節及び A.5 節で説明する。以
186 下に示すストラテジーの名称は参照を容易にするためのみに用いるものであり、実際に
187 estimand を構成する際には、選択したストラテジーについて十分な記述をする必要がある。

188

189 **治療方針ストラテジー**

190 中間事象の発現は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、
191 関心のある変数の値を用いる。

192

193 例えば、レスキュー薬の使用を中間事象としてどのように考慮するかを規定する際に、
194 中間事象の発現は無視し、関心のある変数の観測値を用いることとする。このストラテジ

195 ーが全ての種類の間事象に適用される場合には、これは ICH E9 の用語集
196 (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
197 を反映することになる。

198

199 通常、この戦略は、全ての被験者において中間事象の発現後も変数の値が存在
200 しなければ適用することができない。例えば、この戦略に基づく estimand は、死
201 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。

202

203 複合戦略

204 中間事象の発現を変数の構成要素とする、つまり中間事象を一つ以上の他の臨床結果の
205 指標と統合して関心のある変数とする。

206

207 この戦略では複数の異なる手法を考えることができる。レスキュー薬の使用を
208 必要とすることは、治療効果について意味のある情報を持つ可能性があるため、意味のある
209 治療効果を説明する適切な要約指標とともに変数に組み込まれるだろう。例えば、変数
210 を、レスキュー薬を使用しないことと良好な臨床結果を複合したものとして定義すること
211 ができる。あるいは、数値変数について、中間事象の発現は極端に好ましくない値である
212 と見なされ、適切な要約指標が選択されるだろう。また、別の手法として、計画された観
213 察期間を反映するが、中間事象の前までの変数の値に基づく、曲線下面積を用いることも
214 あるだろう。

215

216 中間事象と考えられる事象自体が、関心のある治療効果を定量化するために測定され得
217 る最も意味のある変数である場合もある。死亡が起きる場合はこれに相当し得る。被験者
218 が死亡したという事実は死亡前の観測値よりもはるかに意味がある可能性があり、また死
219 亡後の観測値は存在しないからである。例えば、心筋梗塞に注目した試験において、死亡
220 した被験者に心筋梗塞が起きていたか、又は起きたであろうかを調べることは必ずしも可
221 能ではないが、変数を死亡又は心筋梗塞を複合したものと定義すれば、この変数は完全に
222 特定できるかもしれない。

223

224 仮想戦略

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

227

228 例えば、倫理的な配慮からレスキュー薬の投与を許容しなければならないが、関心のある
229 治療効果としてはレスキュー薬を使用していない状況下での結果が重要であるかもしれ
230 ない。同様に、治療が失敗したためその治療を中止し、他の実薬治療を行う場合（対照群

231 に割付けられた被験者が試験治療に切り替える場合も含む) も、関心のある治療効果とし
232 てはその実薬治療を行わなかった場合の結果が重要であるかもしれない。これらの例にお
233 いて、レスキュー薬を使用していない状況と他の実薬治療がない状況は、異なる仮想的な
234 条件を反映している。

235

236 特定の試験の状況を踏まえて、関心のある科学的疑問を反映する仮想的な条件を正確に
237 記述するように注意する必要がある。例えば、**estimand** を正確に記述するために、中間事
238 象としてレスキュー薬の使用と治療遵守の両方に、仮想的な条件がうまく対応することが
239 ある。

240

241 主要層ストラテジー

242 対象集団を、中間事象が起こらないであろう主要層（用語集参照）とする。例えば、関
243 心のある対象集団は、治療計画の不遵守が起こらないであろう患者層であるとされるかも
244 しいない。言い換えると、主要層はより広い集団のうち中間事象の発現がないであろう部
245 分集団となる。関心のある科学的疑問は、その層のみでの治療効果に関連する。

246

247 主要層における効果は、試験のデータに基づいて定義されるあらゆる種類の部分集団や
248 試験実施計画書に適合した対象集団の解析によるものと明確に区別されるべきである。主
249 要層別（用語集参照）は、その患者に両方の治療において起こり得る中間事象によって定
250 義される。例えば、いずれの治療でも治療計画を遵守できるであろう患者といった定義が
251 なされる。一般に、中間事象の発現の有無を予測することはできないため、試験に先立っ
252 てこの被験者を特定することは不可能である。また、ランダム化比較試験においては、各
253 被験者は一つの治療のみを受ける状況で観察されることから、試験のデータに基づいてこ
254 の被験者を特定することも不可能である。したがって、主要層に該当する被験者かどうか
255 は、不完全となる場合が多いものの、共変量によって推論しなければならない。一方で、
256 割付けられた治療を受けている状況での中間事象の発現に基づいて主要層に該当する被験
257 者を特定すると、その集団に基づく解析による治療効果の推定には、交絡が生じる恐れが
258 ある。なぜなら、異なる被験者は、異なる治療で異なる中間事象を経験することになるか
259 らである。

260

261 治療下ストラテジー

262 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。変数が繰り返し測定さ
263 れる場合、中間事象の発現までの変数の値は、全ての被験者に対して同様に設定された時
264 点における値よりも、中間事象の発現を考慮するものであると考えられるかもしれない。

265

266 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

267 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
268 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、状況によっては治療
269 計画を遵守している期間の薬剤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
270 しいない。

271

272 この節では、五つの異なるストラテジーについて検討した。それぞれの中間事象を取り
273 扱うのに、より適切なストラテジーを記述する際には、正確であることが重要である。治
274 療計画の遵守について考えると、(i)「全ての被験者が治療計画を遵守したとしたら」という
275 仮想的な状況、(ii)「試験薬が投与される場合に治療計画を遵守できるであろう」患者層、
276 及び(iii)治療計画を遵守している間の効果、に基づく治療効果を区別することが最も重要で
277 ある。

278

279

280 **A.3.3. Estimand の構成**

281

282 **A.3.3.1. 一般的留意事項**

283

284 上述のとおり、関心のある治療効果を明確に説明し、被験者及び医師に説明される治療
285 効果をより妥当なものとするために、estimand の構成においては中間事象について明確に
286 検討する必要がある。Estimand の構成においては、臨床試験で発現する可能性があり試験
287 結果の解釈に影響を及ぼす、それぞれの中間事象に対応すべきである。試験の計画段階に
288 おける中間事象の記述は、理論的には、変数の観察のための特定の時間枠といった、治療
289 と追跡に関する非常に具体的な詳細を反映してもよい。このような特定の基準は、試験結
290 果の解釈に影響しないと考えられる。関連する全ての種類の中間事象を予見することは現
291 実的ではないかもしれない。したがって、試験の報告では、予見していなかった中間事象
292 を解析においてどのように取り扱ったかだけでなく、選択した解析により推定されるも
293 のに対する影響についても議論すべきである。Estimand の構成においては、異なる中間事
294 象に対応するために異なるストラテジー (A.3.2 節、A.7 節) を選択することもできる。

295

296 どのような臨床試験における estimand の構成も、臨床試験の計画及び実施に関与する臨
297 床家、統計家及びその他の専門分野の者も含む多分野にわたる作業となる。Estimand の構
298 成は、これから実施する臨床試験の目的やデザインに関する治験依頼者と規制当局との議
299 論の対象とすべきである。Estimand の構成は、試験の目的に基づいて行うべきあり、また
300 データ収集や解析手法に関する選択肢を与えるべきである。この手順を踏まなかったり、
301 過度に簡素化したりすると、試験の目的、試験デザイン、データ収集及び統計解析の間の
302 不整合を招く恐れがある。

303

304 Estimand の構成には、手順の繰り返しが必要となるかもしれない。Estimand の構成は、
305 試験の対象疾患や治療目標を含む個々の治療状況、及び個々の試験設定における臨床的な
306 妥当性を考慮して、正当化されるべきである (A.3.3.2 節)。さらに、推測のために信頼でき
307 る推定値が得られることを保証するために、試験デザインと統計手法の妥当性について検
308 討する必要がある。特に、臨床試験におけるランダム化の重要な利点を認識し、可能な限
309 り利用すべきである。いくつかの estimand、特に観測されたデータを使って推定する
310 estimand は、少ない仮定で安定した推定を行うことができるが、他の estimand は、正当
311 化することがより困難で、起こりそうな仮定の変化に対してより影響を受ける可能性があ
312 る特定の仮定を必要とする (A.5.1 節)。適切な試験デザインの策定や、特定の estimand
313 に対する信頼できる推定値の導出に重大な問題が存在する場合には、別の estimand、試験
314 デザイン、及び解析手法について検討する必要があるだろう。

315

316

317 A.3.3.2. 治療背景及び試験の実験的側面に関する留意事項

318

319 上述のとおり、疾患の状況と治療の目的の特徴は estimand の構成に影響を与える。治療
320 背景については、代替治療の選択肢があるかどうか、治療に対する個々の反応を観察でき
321 るかどうか、そして治療の目的が症状のコントロールか、疾患経過の修飾か、又は疾患の
322 予防なのかそれぞれ含まれる。例えば、複数の代替治療がある疾患領域においては、反
323 応に応じて個々の患者の治療の選択を変更できることを踏まえ、臨床的兆候や症状のコン
324 トロールを治療目標とすることもある。代替治療（レスキュー薬や、治験実施計画書で禁
325 止された薬剤による治療、又は後続ラインの治療）の使用は、中間事象と見なす必要があ
326 るだろう。関心のある科学的疑問を反映するためどのように中間事象を考慮するかについ
327 ての詳細は、それが代替治療を使用しなかった場合の治療効果の理解か、又は代替治療を
328 必要とせずに治療を続けられる被験者層における治療効果の理解かに基づくかもしれない。
329 場合によっては、これらの疑問に答えることは、例えば被験者が代替治療を使用したか否
330 かを区別しない治療方針の効果を定量化することよりも適切であるかもしれない。このよ
331 うな検討は、対照群に割付けられた被験者の試験治療への切り替えという中間事象に対し
332 てはさらに適切であるかもしれない。Estimand は、安定した推定値が得られることが合意
333 されているのであれば、これらのストラテジーのうち一つを使って構成することができる。
334 その他の状況では、実臨床においても存在する治療方針に関する治療効果を理解する必要
335 があるかもしれない。例えば、治療目的は不都合な臨床結果（死亡等）を予防する又は遅
336 延させることであるかもしれない。後続の治療ラインが確立している状況において、試験
337 治療が治療方針の一部として未治療患者に使用することが提案されている場合、治療方針
338 の効果がより重要になる可能性がある。治療方針ストラテジーに基づいて estimand を構成

339 する場合、推測は追加の **estimand** と中間事象自体に関する解析を定義することで補足する
340 ことができる。例えば、症状スコアに対する効果と、各治療条件の下で使用されたレスキ
341 ュー薬の量の両方を比較することが考えられる。

342

343 治療方針ストラテジーに基づく **estimand** は、規制上の意志決定を支持するのに、より一
344 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** に対
345 して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
346 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
347 合に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
348 味があり、信頼できる治療効果の推定値を得られる可能性がある。この状況では、臨床的
349 により重要と考えられる **estimand** も保持し、結果として得られた推定値をその特定の手法
350 に対する試験デザインや統計解析に関する限界についての議論とともに提示することが推
351 奨される。

352

353 複合ストラテジーについての一例は、連続変数を二値変数に置き換えることである。こ
354 の場合、中間事象がなければ、被験者はスコアの変化について事前に定義された閾値に基
355 づく反応例と非反応例と見なすことができる。このように連続量のスコアを二分化するこ
356 とは、結果として **estimand** を変えることになる。この **estimand** の臨床的な重要性と解釈
357 は、臨床的に解釈可能な反応例の基準と適切な集団レベルの要約（例えば、割合の差やオ
358 ッズ比）が利用できるかどうかにか依存する。

359

360 仮想ストラテジーを使用する場合、他の条件と比較して、規制上の意志決定において許
361 容されやすい条件もある。記述される仮想的な条件は、実臨床での薬剤の使用と関連する
362 解釈可能な治療効果の定量化のために正当化されなければならない。上述したとおり、レ
363 スキュー薬の使用も含めた特定の治療方針よりもむしろ、ある特定の条件における治療薬
364 の効果を対象とし、レスキュー薬を使用しなかったとしたら関心のある変数がどのような
365 値を取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
366 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条件の下で、関心
367 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
368 正当化されないかもしれない。全ての被験者が治療計画を遵守した場合の効果に基づく関
369 心のある科学的疑問は、全ての被験者が遵守すると仮定する仮想的な条件について綿密に
370 議論しなければ、明確に定義することはできない。さらに、試験及び実臨床において治療
371 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
372 得る。ストラテジーを選択する必要がある中間事象が、例えば治療計画への不遵守のみな
373 らず、不遵守の理由（例えば、毒性によるもの）にも依存して決まる場合には、臨床試験
374 においてこれらを定義し、正確に記録しなければならない。

375

376 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
377 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）
378 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
379 映されるだろう。特に、対照群の選択は、レスキュー薬又はその他の併用薬の投与が、試
380 験においてどのように許容されるかに影響を及ぼす可能性がある。

381

382 割付けられた治療以外の治療の使用は、一般的には中間事象と見なされる。レスキュー
383 薬や治験実施計画書で禁止されている薬剤による治療又は後続ラインの治療を含む代替治
384 療は多様となり得る。さらに、単にレスキュー薬といっても、効果不十分である長期治療
385 の代替や追加として用いられる薬剤、割付けられた治療に忍容性がない場合の代替薬、又
386 は疾患の症状における一時的な増悪を管理するための短期間の急性治療薬等、様々であろ
387 う。これらの例は、個々の試験の状況において固有の中間事象の取扱いを検討することの
388 重要性を示している。

389

390 非劣性又は同等性を示すことを目的とした試験における estimand の選択には、慎重な検
391 討が必要である。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が示され、
392 治験実施計画書の違反と逸脱、治療計画への不遵守及び中止を最小化することの重要性を
393 示している。5.2.1 節では、このような試験では一般的に FAS の結果が保守的ではないこと
394 と、その役割について慎重に考えるべきであることを説明している。治療方針ストラテジ
395 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
396 等性の試験に対して FAS に関する同様の問題が生じる。割付けられた治療の中止やその他
397 の治療の使用の後には、最初に割付けられた治療の類似性とは関係のない理由で、両治療
398 群間の反応がより類似して見えることがある。Estimand は治療群間の差の過小評価に繋がる
399 可能性がある中間事象（例えば、レスキュー薬の使用や対象集団からの違反）に直接的
400 に対応できるように構成することもあり得る。この状況では、estimand は、治療間で差が
401 あるのであればそれを高い感度で検出できるような治療効果の指標を対象とすることにな
402 るだろう。

403

404

405 A.4. 試験デザイン及び実施への影響

406

407 試験デザインは、試験の主要な目的を反映し、その目的が達成されたか否かを確認する
408 ための基盤を形成する一つ又は複数の estimand の選択と整合させる必要がある。特に、治
409 療効果の定量化の基盤となる estimand の明確な定義は、試験デザインに関する選択肢を与
410 えるべきである。例えば、特定の間事象が発現したかどうかに関わらない治療効果を理

411 解することに関心がある場合には、中間事象の発現の有無に関わらず全ての被験者の変数
412 を収集する試験デザインが適切である。一方で、規制上の意志決定を裏付けるために必要
413 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
414 めにこのようなデータを収集することの利点は、データ収集により複雑になるあらゆる点
415 及びその潜在的な欠点と比較して検討すべきである。

416

417 重要な追加の estimand も含め、関心のある estimand と整合した統計解析を裏付けるた
418 めに適切な、全てのデータを収集するように努めるべきである。治療計画の不遵守、治療
419 の中止、治療の切り替え、又はレスキュー薬の使用といった中間事象の発現は、死亡等の
420 終末事象とは異なり、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。
421 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
422 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
423 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
424 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の十分
425 な情報の取得を前向きに計画することは、関心のある中間事象とそれ以外の欠測データを
426 区別するのに役立つ、したがって、主要解析を改善させる可能性がある。また、より適切
427 な感度分析の選択に繋がる可能性もある。例えば、総称としての「追跡不能」は、より正
428 確に「有効性の欠如による治療の中止」と記録すべきだろう。それが関心のある中間事象
429 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
430 反映され、欠測データの問題としては扱われない。被験者を試験に留まらせる対策をとる
431 ことはできるが、実臨床に対する試験の外的妥当性が保持されるように注意すべきである。
432 例えば、毒性の影響を軽減するための、試験の対象集団の選択又は漸増法や併用薬の使用
433 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。

434

435 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
436 化治療中止デザイン又は漸増デザインといった標準的ではない試験デザインが必要となる、
437 又はそれらデザインの恩恵を受けるかもしれない。しかし、このような標準的ではないデ
438 ザインでは、その実施やその後の統計的推測に関して特別な注意が必要となるだろう。例
439 えば、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
440 の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
441 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
442 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
443 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。実際には、
444 これらの留意事項によって、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
445 使用は制限されることになるかもしれない。

446

447 中間事象を取り扱うストラテジーの記述による関心のある治療効果の詳細な説明は、被
448 験者数の計算に対して情報を与えるべきである。全ての被験者が解析に対する情報に寄与
449 し、中間事象とその取扱いの影響が対象とする効果の大きさと予期される分散に反映され
450 るのであれば、通常は算出された被験者数に想定される中止例の割合に応じた上乘せをす
451 る必要はない。

452

453 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
454 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
455 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統合したデータを用いること
456 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
457 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
458 アナリシスのデザインや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。
459 それぞれの試験のデータの表示や統計解析に対応している **estimand** について考慮、説明す
460 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
461 く可能性があり、また偏りの原因となる可能性がある。

462

463 臨床試験には複数の **estimand** に変換される複数の目的があることがより一般的である。
464 一つの **estimand** に対して適切な試験デザインは、他の重要である可能性がある **estimand**
465 に対しては適切ではないかもしれない。複数の目的や評価項目のある試験は、検定の多重
466 性の問題を生じさせる可能性があるが、原則としてこの問題は複数の **estimand** がある場合
467 にも同様に適用される。特に、一つ以上のものを主要とし、他のものを副次とする場合に
468 おけるこれらの問題に対処するために用いられる手法は、**estimand** に対しても同様に適応
469 することができる。

470

471

472 **A.5. 試験の解析への影響**

473

474 **A.5.1. 主とする推定**

475

476 対照と比較した治療効果のための **estimand** は、治療効果を比較の基盤となる被験者の群
477 間でのあらゆる差異から分離できるようにするため、治療群の被験者の結果と治療群に類
478 似した対照群の被験者の結果の比較を反映すべきである。規定した **estimand** に整合した解
479 析手法又は推定量を用いるべきである。それにより、信頼できる解釈の基盤となり得る推
480 定値を提示することができる。安定した推定値が得られるかどうかに関する重要な判断は、
481 必要となる仮定の範囲でのものである。主とする推定量及び感度分析における推定量につ
482 いて、必要となる仮定を明確に説明すべきである。その仮定は正当化できるべきであり、

483 ありそうにない仮定は避けるべきである。仮定に対する結果の安定性は、同じ estimand に
484 対する感度分析によって評価されるべきである (A.5.2 節)。

485

486 特に、中間事象の発現の有無に関わらず被験者を完全に追跡した場合には、治療方針ス
487 トラテジーに基づく estimand は最小限の仮定によって推定することができる。このストラ
488 テジーを採用した estimand に対する推定は、中間事象発現後の測定値を収集しなかった場
489 合には、より強い確認できない仮定を必要とする。複合ストラテジーを用いると、本来の
490 変数が完全に確認されない場合でも、中間事象発現後の反応の補完やモデル化の必要なく、
491 また関連する仮定を必要とせずに解析を行うことができるかもしれない。対照的に、中間
492 事象に対応するために仮想的な状況を必要とするストラテジーを用いて構成される
493 estimand に対する推定では、仮想的な条件の慎重な説明が必要であり、確認できず、そし
494 て感度分析によって検討する必要があるモデル化の仮定に必然的に依存することになる。
495 ランダム化試験では、対象集団の主要層における治療効果の推定には、ランダム化前にそ
496 の層の被験者が特定できない限り交絡が生じる。特定できない場合、推定は、特に全ての
497 関連する交絡因子は測定され考慮されているという仮定に依存することになる。例えば、
498 治療の計画を遵守することができるであろう被験者の層のために、観測された治療の遵守
499 例と対照治療の遵守例を単純に比較することは不適切である。これらは系統的に異なる被
500 験者集団であり、治療の効果の推定に交絡が生じる。この場合、モデルの誤特定に対抗す
501 るような保証を与えることは難しいが、少数の事前に規定した共変量よりもむしろ、全て
502 の重要な交絡因子を考慮することが不可欠である。治療下ストラテジーについては、中間
503 事象の発現とその時期が治療に関連する場合、治療効果の推定にはより強い仮定が必要と
504 なるだろう。

505

506 中間事象に対応する estimand を適切に定義し、推定に必要なデータの収集に努めた場合
507 にも (A.4 節)、いくつかのデータは欠測となるかもしれない。この欠測データは、推定に
508 必要な情報の収集の系統的な失敗や収集をやめることとは区別する。例えば、治療方針ス
509 トラテジーに基づいて estimand が構成された場合、割付けられた治療を中止した後であっ
510 ても、被験者を試験に留まらせ、評価の計画を遵守させるようにあらゆる努力をすべきで
511 ある。その努力が功を奏しなかった場合には、観測されたデータのみに基づいて、個々の
512 観測値の予測や補完、又は統計手法の正当化のために、欠測した観測値に関する仮定をお
513 くことが必要になる。欠測データの取扱いはもっともらしい仮定に基づくべきであり、ま
514 た可能であれば、estimand の記述において用いられているストラテジーによって導かれるべ
515 きである。ある被験者に対する予測は、その被験者の観測されたデータ (共変量及びベー
516 スライン後の値) 及び他の類似した被験者の観測されたデータに基づくかもしれない。類
517 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
518 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予測モデルでは、治療を続

519 けた被験者のデータよりも、むしろ治療を中止したが、データの収集は継続した被験者の
520 データを用いるかもしれない)。これらの手法の仮定からの起こり得るずれは、感度分析の
521 重要な要素となる。

522

523

524 **A.5.2. 感度分析**

525

526 **A.5.2.1. 感度分析の役割**

527

528 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
529 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
530 て評価される。

531

532 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
533 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を確認する目的
534 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。感度分析とは
535 異なり、試験のデータをより十分に評価し、理解するために計画、提示又は要求されるそ
536 の他の各解析は補足的解析(用語集参照)と呼ぶ。それぞれの補足的解析は異なる estimand、
537 又は同じ estimand に対する異なる推定量に関連させることができるだろう。関心のある主
538 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
539 定されていれば、一般的には補足的解析は感度分析よりも優先順位が低いものとすべきで
540 ある。

541

542

543 **A.5.2.2. 感度分析の選択**

544

545 感度分析を計画し、実施する際には、主とする解析の多くの要素を同時に変更しないこ
546 とが推奨される。多くの要素を同時に変更した場合には、結果の違いの原因になる仮定が
547 あったとしても、見られる可能性があるどの結果の違いについてどの仮定が原因となっ
548 ているかを特定することが困難になるからである。より透明性があり有用な方法は、一度に
549 一つの仮定だけを変え、その影響を評価することである。さらに、確認できる仮定と確認
550 できない仮定を区別することは、異なる解析の解釈と関連性について評価する際に役立つ
551 だろう。

552

553 欠測データは感度分析において特に注意が必要である。なぜなら、どのような方法にお
554 ける仮定も完全に正当化することが困難であり、また確認することができない可能性があ

555 るからである。欠測データは個々の estimand に対して定義し、考慮しなければならない。
556 例えば、治療方針ストラテジーに基づく estimand のために必要な情報を得るため、試験治
557 療中止後も収集が計画されていたデータは、収集されなかった場合には欠測となる。しか
558 し、同じデータは他のストラテジーでは重要ではなく、その副次的な estimand の目的に対
559 しては収集されなかったとしても欠測とはならないかもしれない。幸い、多くの場合、重
560 要な仮定からのずれは簡単に特徴付けることができる。例えば、連続量の結果に対する平
561 均の解析において、それぞれの治療群における欠測データと非欠測データがその平均にお
562 いて異なる限り、また特にこれらの差異自体が治療群によって異なる場合、元の解析には
563 偏りが生じている可能性がある。これらの違いについて想定される値のもっともらしい範
564 囲を調べ、結論の安定性を評価すべきである。例えば、有意性検定において、治療効果が
565 事前に規定されていた水準で統計学的に有意又は有意ではない場合のこれらの違いの値を、
566 tipping point analysis によって図示することができる。治療効果が一定の臨床的な意義を
567 保つ又は保てない場合のこれらの違いの値を確かめるためにも同様の手法の利用が考えら
568 れる。他のデータ構造に対しても同様の方法を適用することができる。例えば、治療成功
569 例の割合又は事象までの時間のデータに対するハザード比も、欠測データと非欠測データ
570 の間で差があり、その程度が治療群間で異なると仮定することができる。

571

572

573 A.5.3. 補足的解析

574

575 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が確認される
576 のであれば、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
577 る。

578

579 異なる estimand を対象とした補足的解析は、追加の考察を与えることができるが、試験
580 結果の解釈においては副次的な役割を担うものである。例えば、反応例の割合に基づく解
581 析は、連続量の平均変化量の群間差によって定量化される治療効果の解釈に対して有用で
582 あるかもしれない。あるいは、異なる定義の反応例に基づく解析が、その定義に対して結
583 果が安定しているかを評価するために実施されるかもしれない。補足的解析の必要性と有
584 用性は、それぞれの試験において判断すべきである。

585

586 5.2.3 節では、FAS の解析と 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
587 よう、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。FAS と PPS に基づく解
588 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
589 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
590 フレームワークにおいては、特定の治療群に割付けられた上で試験実施計画書を遵守した

591 被験者の部分集団に基づく解析を実施することはできるが、それ自体は関心のある治療効
592 果を明確に定義するものではない。上述したとおり、治験実施計画書を遵守した被験者集
593 団のデータに基づく解析では、治療を遵守する被験者集団における効果を推定するという
594 目標を達成しない。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
595 にならないからである。したがって、このような解析の役割は、治験実施計画書の違反や
596 逸脱の程度が試験結果の信用度を損なわせていないかどうかを評価することに限定される。
597 治験実施計画書の違反や逸脱には、中間事象として扱われるものもある。中間事象の多く
598 が estimand の構成によって取り扱われる場合には、残された治験実施計画書の違反や逸脱
599 の数は少なく、PPS の解析は追加の考察を加えることにならないかもしれない。

600

601

602 A.6. Estimand と感度分析の記載

603

604 Estimand は治験実施計画書において定義し、明確に特定すべきである。治験実施計画書
605 では、試験の結果の解釈に影響を及ぼすであろう予見可能な中間事象の種類を特定し、試
606 験の主要な目的に対応する主要な estimand を事前に規定すべきである。さらに、治験実施
607 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
608 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ともに事前に規定す
609 べきである。規制上の意志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
610 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
611 切な感度分析とともに、適切に記載すべきである。探索的な目的のために、追加の estimand
612 に繋がる追加的な試験目的が検討されるかもしれない。

613

614 個々の探索的な疑問に対する estimand の詳細を記載すること、特に、それら estimand
615 が主要又は副次的な estimand の中間事象の取扱いに関する細かなバリエーションである
616 場合、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、これ
617 らの estimand の詳細を記載することは規制上の要件ではない。しかし、異なる関心のある
618 科学的疑問のために大きく異なる estimand を必要とする場合には、十分に文書に記載する
619 ことが推奨される。

620

621 主要な estimand の選択は、多くの場合、試験デザインや実施の特徴の主要な決定要因と
622 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
623 る。追加的な estimand に重要な関心がある場合には、試験デザインや実施の特徴の検討は、
624 必要に応じてそれらの estimand も支持するように拡張されるかもしれないし、同様に記載
625 すべきである。これらの特徴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
626 はこれまでと同様である。例えば、一つ以上の estimand があり、検定の多重性の問題が生

627 じる可能性がある場合には、第一種の過誤確率の制御に関する通常の留意点が適用され、
628 適切に記載されるべきである（A.4 節）。

629

630 主とする解析、感度分析及び補足的な解析の結果は、それぞれの解析が事前に規定され
631 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
632 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。試験の計
633 画段階で予見していなかった、又は試験実施中に特定された中間事象に対応した場合は、
634 中間事象の解析における取扱い方法だけでなく、選択された解析が推定したもの及び試験
635 結果の解釈に対する影響についても議論すべきである。

636

637

638 A.7. 一般的な例

639

640 以下では、連続変数の一般的な事例を使って、本補遺で提案したフレームワークを説明
641 する。この例は規制上の推奨事項と捉えられるべきではなく、それぞれの試験の設定（特
642 に、又はこれらに限らないが、二値変数や事象までの時間の変数を用いる場合）での要求
643 に合わせて適応させるべきである。

644

645 特定の慢性的かつ致死性ではない疾患の患者に対する新しい試験治療（薬剤 X）が検討さ
646 れる。治療に対する反応は連続値を用いて毎月観測される。薬剤 X の最大効果は治療開始
647 から 4~6 ヶ月後に見られると期待される。主とする科学的疑問は、6 ヶ月時点における薬
648 剤 X とプラセボとの比較であり、これはランダム化比較試験によって最もよく評価できる。
649 この臨床試験におけるプラセボの使用は倫理的と考えられており、有効性が欠如している
650 場合には割付けられた治療を中止し、レスキュー薬に切り替えられるように規定されてい
651 る。レスキュー薬への切り替えは中間事象であるが、その後の変数の測定値は収集可能で
652 ある。この点は有害事象による治療中止といった他の中間事象の後では同様だが、死亡（こ
653 の例の状況では非常に稀と考えられる）のような中間事象には当てはまらない。

654

655 中間事象が全く発生しないと予想されるという非現実的な状況においては、適切な
656 estimand の定義を以下の四つの要素とすることに議論の余地はない。

657 A. 対象集団：承認のための対象患者集団を反映する適切な選択／除外基準によって定義
658 される

659 B. 変数：計画された測定値のベースラインから 6 ヶ月後までの変化量

660 C. 中間事象：考慮すべき中間事象はない

661 D. 集団レベルでの要約：治療条件間での変数の平均値の差

662

663 このとき、estimand は、対象患者集団における計画された測定値のベースラインから 6
664 ヶ月後までの変化量の治療条件間での平均値の差である。

665

666 この estimand を対象とした試験デザインは、試験を通して全ての測定値を収集するラン
667 ダム化並行群間比較デザインである。測定値を収集できなかった場合には、欠測データと
668 なる。全ての測定値が収集されている限り、治療群を因子とした分散分析モデルによる解
669 析が、この estimand に対する統計解析の一例となる。測定値が欠測している場合には、欠
670 測データによる不確実性を考慮したもっともらしい仮定に基づいてデータを予測する必要
671 がある。例えば、欠測データは試験を継続した類似する被験者に基づいて補完されるかも
672 しれない。類似性は、ベースラインの共変量が同じであること、割付けられた治療群が同
673 じであること、測定値の推移や中間事象に関する情報が同じであることによって定められ
674 るだろう。例えば、補完方法に用いる仮定を評価するための感度分析を、治験実施計画書
675 に事前に規定すべきである。推測は、レスキュー薬への切り替えの割合や時期を治療群間
676 で比較するといった、異なる estimand を対象とする可能性がある補足的解析を追加するこ
677 とによって補うことができる。

678

679 要素 C は略して「中間事象」と呼ぶが、関心のある科学的疑問を反映するために、起こ
680 り得る中間事象をどのように考慮するかの説明を示すものである。

681

682

683 A.7.1. 中間事象が一つの例

684

685 実際には、複数の中間事象が発現すると考えられる。説明を簡単にするため、まずは「有
686 効性欠如によるレスキュー薬への切り替え」という中間事象のみが発現すると考えられる
687 状況について考える。以下では、異なる科学的疑問に対応する様々な estimand について、
688 試験デザイン、実施、及び解析に関するハイレベルな留意事項とともに説明する。

689

690 治療方針ストラテジー

691 A. 対象集団：承認のための対象患者集団を反映する適切な選択／除外基準によって定義
692 される

693 B. 変数：計画された測定値のベースラインから 6 ヶ月後までの変化量

694 C. 中間事象：レスキュー薬への切り替えの有無に関わらない

695 D. 集団レベルでの要約：治療条件間での変数の平均値の差

696

697 この例において、治療方針ストラテジーにより記述される estimand は、「薬剤 X+必要
698 に応じたレスキュー薬の投与」による効果と「プラセボ+必要に応じたレスキュー薬の投

699 与」による効果の、変数の測定値による比較となる。つまり、この **estimand** は両治療群間
700 のレスキュー薬への切り替え例の割合に依存して、治療とレスキュー薬の効果を併せた効
701 果を捉えることになる。また、この **estimand** は、レスキュー薬への切り替えが「有効性の
702 欠如」という好ましくない事象により生じていることは捉えていない。

703

704 この場合、**estimand** は、レスキュー薬への切り替えの有無に関わらない、対象患者集団
705 における計画された測定値のベースラインから 6 ヶ月後までの変化量の治療条件間での平
706 均値の差となる。

707

708 以下のそれぞれの例においても同様の説明文を構成することが可能であり、また、中間
709 事象の取扱い方法の説明をまとめることもできる。

710

711 この **estimand** を対象とした試験デザインは、レスキュー薬への切り替えの有無に関わら
712 ず試験を通して全ての測定値を収集する、ランダム化並行群間比較デザインである。

713

714 全ての測定値が収集されている限り、治療群を因子とした分散分析モデルによる解析が、
715 この **estimand** に対する統計解析の一例となる。測定値が欠測している場合には、欠測デー
716 タによる不確実性を考慮したもっともらしい仮定に基づいてデータを予測する必要がある。
717 例えば、欠測データは試験を継続した類似する被験者に基づいて補完されるかもしれない。
718 類似性は、ベースラインの共変量が同じであること、割付けられた治療群が同じであるこ
719 と、測定値の推移や中間事象に関する情報が同様であることによって定められるだろう。
720 例えば、補完方法に用いる仮定を評価するための感度分析を、治験実施計画書に事前に規
721 定すべきである。推測は、レスキュー薬への切り替えの割合や時期を治療群間で比較する
722 といった、異なる **estimand** を対象とする可能性がある補足的解析を追加することによって
723 補うことができる。レスキュー薬の使用についての科学的疑問に対応するために、別の
724 **estimand** が構成される可能性がある。

725

726 複合ストラテジー

727 A. 対象集団：承認のための対象患者集団を反映する適切な選択／除外基準によって定義
728 される

729 B. 変数：計画された測定値のベースラインから 6 ヶ月後までの変化量が事前に規定した
730 閾値を超え、かつレスキュー薬への切り替えがなかった場合を有効として扱う二値変
731 数

732 C. 中間事象：中間事象は変数の定義により捉えられる

733 D. 集団レベルでの要約：治療条件間での有効例の割合の差

734

735 複合ストラテジーにより記述される estimand は、6 ヶ月時点の変数の測定値のみに関する
736 治療効果を評価するものではない。治療効果は、臨床的な意味の有無で二分化された変
737 数の測定値の変化量と、「レスキュー薬への切り替え」という中間事象を組み合わせた、複
738 合的な変数に基づいて定められる。レスキュー薬への切り替えは有効性の欠如によるもの
739 であり、この estimand はレスキュー薬の投与が好ましくない結果であるという情報を含む。

740

741 この estimand を対象とした試験デザインは、ランダム化並行群間比較デザインである。
742 レスキュー薬への切り替え後の測定値は、そのようなデータを必要とする試験の他の目的
743 (例えば、中間事象の後でも安全性情報を収集すること)に関心がない限りは、収集する
744 必要はない。この例では、レスキュー薬投与後に収集し得るデータは対象とした estimand
745 の推定においては関心がないことから、欠測とは見なされない。

746

747 有効例か否かを判定するための測定値が全て収集されている限り、ロジスティック回帰
748 分析がこの estimand に対する統計解析の一例となる。欠測データがある場合、すなわち、
749 中間事象の発現がなく、評価時点より前に欠測となった場合には、欠測データによる不確
750 実性を考慮したもっともらしい仮定に基づいて有効例か否かを補完する必要がある。例え
751 ば、欠測データは試験を継続した類似する被験者に基づいて補完されるかもしれない。類
752 似性は、ベースラインの共変量が同じであること、割付けられた治療群が同じであること、
753 測定値の推移が同じ同様であることによって定められるだろう。例えば、補完方法に用い
754 る仮定を評価するための感度分析を、治験実施計画書に事前に規定すべきである。推測は、
755 変数の定義のための閾値を変えるとといった、この複合ストラテジーによる estimand の個々
756 の構成要素を対象とした、結果として異なる estimand に基づく補足的解析を追加すること
757 によって補うことができる。

758

759 仮想ストラテジー

760 A. 対象集団：承認後の対象患者集団を反映する適切な選択／除外基準によって定義され
761 る

762 B. 変数：計画された測定値のベースラインから6ヶ月後までの変化量

763 C. 中間事象：6ヶ月時点まで被験者はレスキュー薬を使用しなかったとする

764 D. 集団レベルでの要約：治療条件間での変数の平均値の差

765

766 仮想ストラテジーによって記述される estimand は、レスキュー薬が被験者に使用しなかつたという別の仮想的な状況での治療効果に対応する。この科学的疑問を直接的に対象とした臨床試験の実施は、倫理的に正当化されないだろう。

769

770 仮想ストラテジーを用いた estimand を対象とした試験デザインは、ランダム化並行群間

771 比較デザインである。レスキュー薬への切り替え後の測定値は、そのようなデータを必要
772 とする試験の他の目的（例えば、中間事象の後でも安全性情報を収集すること）に関心が
773 ない限りは、収集する必要はない。この例では、レスキュー薬投与後に収集し得るデータ
774 は対象とした **estimand** の推定においては関心がないことから、欠測とは見なされない。

775

776 この **estimand** に対する統計解析は、被験者がレスキュー薬を使用しなかったという仮想
777 的な状況において観測されたであろう測定値に関する仮定に基づくこととなる。一般的に、
778 このような予測において必要となる仮定は観測されたデータに基づいて確認することがで
779 きないため、結論の安定性を評価するために感度分析が必要となる。仮定のもっともらし
780 さに関する議論が、これらの仮定に対して十分な信頼性を与えるため、また結果として治
781 療効果の推定に対して十分な信頼性を与えるために必要となる。推測は、レスキュー薬へ
782 の切り替えの割合や時期を治療条件間で比較するといった、異なる **estimand** を対象とする
783 可能性がある補足的解析を追加することによって補うことができる。

784

785 主要層ストラテジー

- 786 A. 対象集団：選択／除外基準によって定義される対象集団のうち、割付けられた治療に
787 関わらず、6 ヶ月間にわたりレスキュー薬の投与を必要としないであろう被験者とし
788 て定義される
- 789 B. 変数：計画された測定値のベースラインから6 ヶ月後までの変化量
- 790 C. 中間事象：中間事象は対象集団の定義により捉えられる
- 791 D. 集団レベルでの要約：治療条件間での変数の平均値の差

792

793 主要層ストラテジーによって記述される **estimand** は、どちらの治療に割付けられたかに
794 関わらず、6 ヶ月間にわたりレスキュー薬を必要としないであろう対象集団の層における、
795 最初に割付けられた治療の効果を評価する。

796

797 実際にはこの集団を事前に特定することが困難であることが、この **estimand** の複雑さの
798 一つである。したがって、実際には、6 ヶ月間にわたりレスキュー薬を必要としないであろ
799 う被験者を対象とするために、導入期を用いたデザイン、エンリッチメントデザインやラ
800 ンダム化治療中止デザインといった標準的ではない試験デザインを採用しなければならな
801 いかもしれない。

802

803 この **estimand** に対する統計解析は、6 ヶ月間にわたりレスキュー薬を必要としないであ
804 る被験者のみがランダム化され、またそれらの被験者が試験期間全体を通して追跡され
805 る限り、単純なものとなる。しかし、上述したとおり、一般的にはこの集団に含まれる被
806 験者を事前に特定することは困難である。対象集団を特定することができない場合、レス

807 キュー薬への切り替えをしなかった被験者のみを対象とした解析では、適切な解析を実施
808 することはできない。なぜなら、異なる割付けられた治療により系統的に異なる被験者を
809 除外する可能性があり、被験者がレスキュー薬へ切り替える傾向に影響する被験者の特徴
810 と治療効果が交絡する可能性があるからである。適切な解析を行うには、この交絡を考慮
811 する必要がある。さらに、設定された仮定に対する結論の安定性を適切な感度分析を用い
812 て評価する必要がある。推測は、治療条件間でのレスキュー薬への切り替えの割合や時期
813 を治療条件間で比較するといった、異なる estimand を対象とする可能性がある補足的解析
814 を追加することによって補うことができる。

815

816 治療下ストラテジー

- 817 A. 対象集団：承認のための対象患者集団を反映する適切な選択／除外基準によって定義
818 される
- 819 B. 変数：割付けられた治療を受けている間の計画された測定値の平均
- 820 C. 中間事象：中間事象は変数の定義によって捉えられる
- 821 D. 集団レベルでの要約：治療条件間での変数の平均値の差

822

823 この estimand は変数の測定値により平均的な治療効果を評価する。ここで選択された変
824 数は、治療を受けている間、すなわちレスキュー薬への切り替えの前までの結果を平均し
825 たものである。

826

827 この estimand を対象とした試験デザインは、ランダム化並行群間比較デザインである。
828 レスキュー薬への切り替え後の測定値は、そのようなデータを必要とする試験の他の目的
829 （例えば、それらのデータを必要とする別の estimand や、中間事象の後でも安全性情報を
830 収集すること）に関心がない限りは、収集する必要はない。この例では、レスキュー薬投
831 与後に収集し得るデータは対象とした estimand の推定においては関心がないことから、欠
832 測とは見なされない。

833

834 割付けられた治療を受けている間の全ての測定値が収集されている限り、治療群を因子
835 とした分散分析モデルによる解析が、この estimand に対する適切な統計解析となる。測定
836 値が間欠的に欠測している場合には、欠測データによる不確実性を考慮したもっともらし
837 い仮定に基づいてデータを補間する必要がある。例えば、補間方法に用いる仮定を評価す
838 るための感度分析を、治験実施計画書に事前に規定すべきである。推測は、治療を受けて
839 いる間の最終測定値に焦点を当てることにより変数の定義に対する他の選択を考慮すると
840 いった、異なる estimand を対象とする可能性がある補足的解析を追加することによって補
841 うことができる。

842

843

844 A.7.2. 中間事象が二つの例

845

846 一般的な例を、「レスキュー薬への切り替え」及び「有害事象による治療の中止」の二種
847 類の中間事象が起こるであろう状況に拡張する。臨床的に意味のある estimand の定義は、
848 estimand に中間事象を含めることなく対象としている治療効果の説明を完全には理解する
849 ことができないのであれば、その臨床試験の設定において起こりそうで臨床的に重要な、
850 全ての中間事象を網羅する必要がある。試験デザイン、実施及び統計解析に関する選択に
851 ついても、同様のことが言える。上で議論した五つのストラテジーを考えると、二種類の
852 中間事象に対するストラテジーの可能性のある全ての組み合わせを考えることができるが、
853 全ての組み合わせが臨床的に重要というわけではない。

854 説明を簡単にするため、以下では、estimand に対する二つの異なるストラテジーのみにつ
855 いて、試験デザイン、実施及び解析に関するハイレベルな留意事項とともに説明する。

856

857 両方の中間事象を治療方針ストラテジーにより考慮する場合

858 A. 対象集団：承認のための対象患者集団を反映する適切な選択／除外基準によって定義
859 される

860 B. 変数：計画された測定値のベースラインから 6 ヶ月後までの変化量

861 C. 中間事象：レスキュー薬への切り替え及び有害事象による治療の中止の有無に関わら
862 ない

863 D. 集団レベルでの要約：治療条件間での変数の平均値の差

864

865 この estimand は、変数の測定値による治療を開始するという方針の効果を対象としてい
866 る。この estimand は、レスキュー薬の投与開始も有害事象による治療中止も考慮しない。
867 特に、レスキュー薬への切り替え及び有害事象が好ましくない結果であることを捉えてい
868 ない。

869

870 この estimand を対象とした試験デザインは、レスキュー薬への切り替えや有害事象によ
871 る治療中止に関わらず、試験を通して全ての測定値を収集するランダム化並行群間比較デ
872 ザインである。

873

874 全ての測定値が収集されている限り、治療群を因子とした分散分析モデルによる解析が、
875 この estimand に対する適切な統計解析となる。測定値が欠測している場合には、欠測デー
876 タによる不確実性を考慮したもっともらしい仮定に基づいてデータを予測する必要がある。
877 例えば、欠測データは試験を継続した類似する被験者に基づいて補完されるかもしれない。
878 類似性は、ベースラインの共変量が同じであること、割付けられた治療群が同じであるこ

879 と、測定値の推移や中間事象に関する情報が同じであることによって定められるだろう。
880 例えば、補完方法に用いる仮定を評価するための感度分析を、治験実施計画書に事前に規
881 定すべきである。推測は、レスキュー薬への切り替え及び有害事象による治療中止の割合
882 や時期を治療群間で比較するといった、異なる estimand を対象とする可能性がある補足的
883 解析を追加することによって補うことができる。

884

885 二種類の間件事象を仮想ストラテジーと治療方針ストラテジーの組み合わせにより考慮す 886 る場合

- 887 A. 対象集団：承認のための対象患者集団を反映する適切な選択／除外基準によって定義
888 される
889 B. 変数：計画された測定値のベースラインから 6 ヶ月後までの変化量
890 C. 中間事象：6 ヶ月時点まで被験者はレスキュー薬を使用しなかったとし、有害事象に
891 による治療中止に関わらない
892 D. 集団レベルでの要約：治療条件間での変数の平均値の差

893

894 この estimand は二種類の間件事象を考慮するために、二つの異なるストラテジーを組み
895 合わせて利用している。レスキュー薬への切り替えに対応するために仮想ストラテジーを
896 用い、有害事象による治療中止に対応するために治療方針ストラテジーを用いている。こ
897 のような estimand は、薬理学的な効果を対象としているがレスキュー薬の使用を控えるこ
898 とは倫理的ではなく、有害事象によって治療を中止した後被験者が治療されないままと
899 なる場合には、関心のあるものとなり、また解釈が容易であるかもしれない。

900

901 この estimand を対象とした試験デザインは、有害事象による治療中止に関わらず試験を
902 通して全ての測定値を収集する、ランダム化並行群間比較デザインである。レスキュー薬
903 への切り替え後の測定値は、そのようなデータを必要とする試験の他の目的に関心がない
904 限りは収集する必要はない。この例では、レスキュー薬投与後に収集し得たであろうデー
905 タは欠測とは見なされない。

906

907 この estimand に対する統計解析は、両方の中間事象を考慮する必要がある。

- 908 ● レスキュー薬への切り替え：関心があるのは、被験者が 6 ヶ月時点までレスキュー薬
909 を使用しなかった場合の効果である。この状況での測定値を直接観察することはでき
910 ないため、この仮想的な状況において観測されたであろう測定値に関する仮定を用い
911 る必要がある。
- 912 ● 有害事象による試験治療の中止：関心があるのは、この中間事象に関わらない効果で
913 ある。したがって、この中間事象に関わらず全ての測定値を解析に含める必要がある。
914 測定値が欠測している場合には、欠測データによる不確実性を考慮したもってもらし

915 い仮定に基づいてデータを予測する必要がある。例えば、欠測データは試験を継続し
916 た類似する被験者に基づいて補完されるかもしれない。類似性は、ベースラインの共
917 変量が同じであること、割付けられた治療群が同じであること、測定値の推移や、発
918 現時期といった中間事象に関する情報が同じであることによって定められるだろう。

919

920 個々の予測が、観測された中間事象と関心のある **estimand** に合致するようになされるの
921 であれば、例えば、ランダム化された全ての被験者に基づく分散分析を用いた統計解析は
922 適切である。測定値が欠測している場合には、欠測データによる不確実性を考慮したもっ
923 ともらしい仮定に基づいてデータを予測する必要がある。例えば、欠測データは試験を継
924 続した類似する被験者に基づいて補完されるかもしれない。類似性は、ベースラインの共
925 変量が同じであること、割付けられた治療群が同じであること、測定値の推移や中間事象
926 に関する情報が同じであることによって定められるだろう。例えば、補完方法に用いる仮
927 定を評価するための感度分析を、治験実施計画書に事前に規定すべきである。推測は、レ
928 スキュー薬への切り替え及び有害事象による治療中止の割合や時期を治療群間で比較する
929 といった、異なる **estimand** を対象とする可能性がある補足的解析を追加することによって
930 補うことができる。

931

932

933 用語集

934

935 **Estimand**

936 試験の目的により提示される関心のある科学的疑問に対応する推定の対象。**Estimand** の要
937 素には、対象集団、関心のある変数（又は評価項目）、関心のある科学的疑問を反映するた
938 めに中間事象をどのように考慮するかという説明、及び集団レベルでの変数の要約が含ま
939 れる。

940

941 **推定値 Estimate**

942 観測された臨床試験データに基づいて、推定量によって算出される数値。

943

944 **推定量 Estimator**

945 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。

946

947 **中間事象 Intercurrent Event**

948 治療開始後に発現し、変数を観測できなくする、又は変数の解釈に影響を与える事象。

949

950 **欠測データ Missing Data**

951 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。
952 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される
953 べきである。

954

955 **主要層別 Principal Stratification**

956 全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集団を区分すること。治療
957 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
958 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
959 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
960 るであろう被験者集団の、四つの主要層が考えられる。

961

962 **主要層 Principal Stratum**

963 本補遺では、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）を意味す
964 る。

965

966 **感度分析 Sensitivity Analysis**

967 主とする推定量の、モデル化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
968 性を調べるために実施される、同じ estimand を対象として、異なる仮定を用いた一連の解
969 析。

970

971 **補足的解析 Supplementary Analysis**

972 主とする解析及び感度分析に加えて、治療効果の理解に追加の考察を与えるための解析の
973 総称。この用語は感度分析よりも広い種類の解析を指す。