

- 1** 例数設計が必要な理由
- 2** 例数設計の実際
- 3** 例数設計に対するオルトメディコの取り組み
- 4** 本日のまとめ

1

例数設計が必要な理由

2

例数設計の実際

3

例数設計に対するオルトメディコの取り組み

4

本日のまとめ

人 機能性表示食品とヒト臨床試験の症例数

■IV 資料作成に当たっての考え方, (V) 機能性に係る事項, 第2 最終製品を用いた臨床試験 (ヒト試験) の実施及び資料の届出, 1 最終製品を用いた臨床試験 (ヒト試験) の実施に当たっての留意事項, (3) 臨床試験 (ヒト試験) に係る提出資料, ①臨床試験 (ヒト試験) に関する査読付き論文

臨床試験 (ヒト試験) の結果について、その内容を誰もが適切に評価できるよう、国際的にコンセンサスの得られた指針 (本ガイドラインの施行時において、**ランダム化並行群間比較試験についてはCONSORT2010声明が該当する** (別紙3参照)。原則として、最新版の国際指針に基づく必要がある。) に準拠した形式で査読付き論文として公表された論文 (査読を経て採択された後、公表準備段階 (印刷中 (in press) 等) にある論文も含む。なお、公表後は速やかに公表論文を提出すること。) を提出する。 (...)

機能性表示食品の届出等に関するガイドライン (改正 令和6年4月1日 (消食表第217号)) より引用



CONSORTに則った
臨床試験が必要



CONSORTに則るためには
SPIRITに則った臨床試験が必要

CONSORT 2010 checklist (Schulz KF, et al 2010)

SPIRIT 2013 Checklist (Chan AW, et al 2013)

How sample size was determined
(どのように目標症例数が決められたか。)

When applicable, explanation of any interim analyses and stopping guidelines
(あてはまる場合には、中間解析と中止基準の説明。)

Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations
(研究目的を達成するために必要な推計参加者数とその決定方法。サンプルサイズのすべての算定のもととなる臨床的・統計学的仮定を含む。)

人 機能性表示食品における例数設計に対する指摘

■第1部 臨床試験WG報告書, 第3章 臨床試験方法論に関する特徴, 3-1 臨床試験のプロトコルの特徴

3-1-3 結果と考察

17) 有効性判定の有意水準の設定に関する記載

「 $p=0.01$ のみ」が0編 (0.0%)、「 $p=0.05$ のみ」が 27 編 (79.4%)、「 $p=0.01$ と $p=0.05$ の両方」が7編 (20.6%) であった。ほとんどの論文で有意水準が記載されていたが、単に有意水準を超えているかどうかだけで効果を判断することは、科学的に正しいとは言えない。検定では、サンプルサイズが大きいほど有意になりやすいため、適切なサンプルサイズの設定を行ったか否かを踏まえて解釈する必要がある。また、後述するが、多重検定している場合は5%有意水準の設定は適切ではないこともあるので、研究プロトコルにおいては注意が必要である。

3-1-4 課題

2) トクホと同様の機能性を標榜する際の留意点

トクホと同様の機能性を標榜する際には、同じ適格基準の参加者や介入期間、研究デザインで実施する必要がある。参加者に関しては、参加者数 (サンプルサイズ) をどのように設定したのか、ランダム化をどのように実施したのか、群間の背景因子の均衡性は保たれていたのかなどの情報や、割り付けから最終評価までを明確にするフローチャートの作成 (ドロップアウトがあった場合には理由を明記) が必要である。

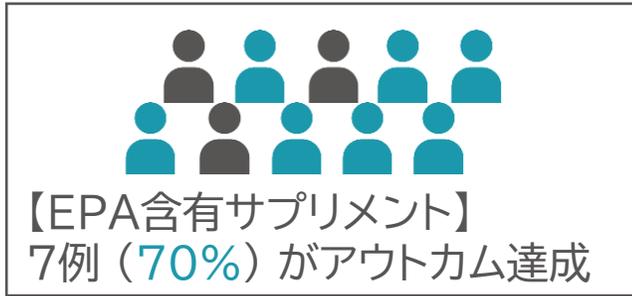
機能性表示食品制度における臨床試験及び安全性の評価内容の実態把握の検証・調査事業 報告書より引用

ガイドラインに明確な記述はないが、平成29年の臨床試験WGによる検証・調査事業の報告書では、例数設計について指摘がなされている。

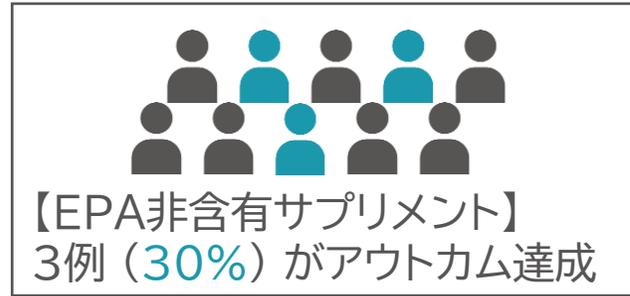
有意確率（P値）の問題点と症例数への影響

対象者	LDL-Cが境界域（120 mg/dL以上139 mg/dL以下）
介入	EPA含有サプリメント
対照	EPA非含有サプリメント
アウトカム	最終評価時点のLDL-Cが120 mg/dL未満の割合

Study1



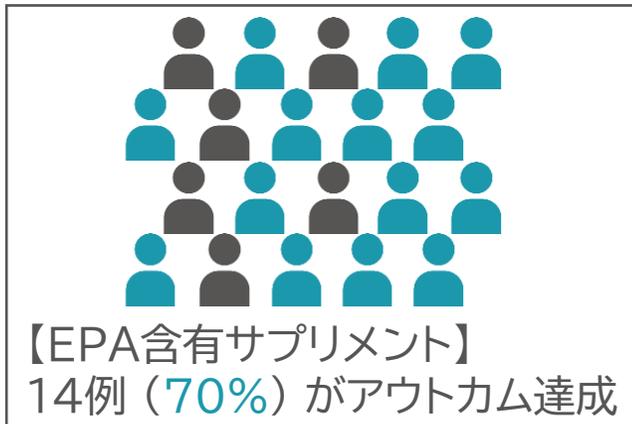
VS



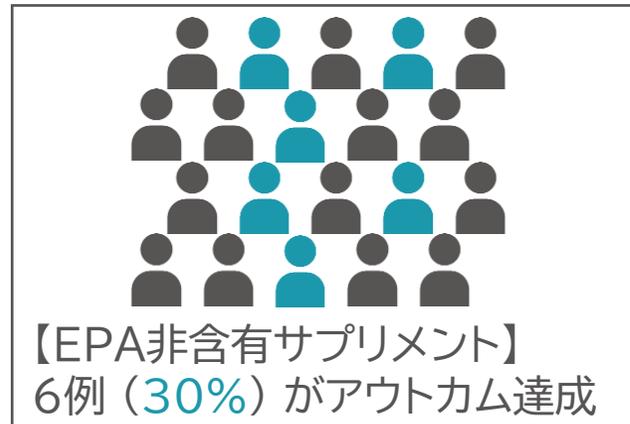
→ 群間差40%、有意差なし
(P=0.179)

臨床的に意味のある差を用いて症例数を計算することが重要

Study2



VS



→ 群間差40%、有意差あり
(P=0.026)

- ① 症例数が増えると解析の精度が上がる。
- ② 臨床的に意味がなくても有意差が検出される。

- 1 例数設計が必要な理由
- 2 例数設計の実際
- 3 例数設計に対するオルトメディコの取り組み
- 4 本日のまとめ

人 ICH-E9における例数設計の記述（1）

■III. 試験計画上で考慮すべきこと, 3.5 必要な被験者数

臨床試験の被験者数は、提示された問題に信頼のおける解答を与えられるよう常に十分多くすべきである。**試験に必要な被験者数は、通常試験の主要な目的により決められる。**被験者数がその他の理由から決定される場合には、その理由を明確にし正当化しておくべきである。例えば、安全性に関する問題若しくは要求に基づいた試験又は重要な副次目的に基づいて被験者数が決定される試験では、**主要な有効性の問題に基づいて被験者数が決定される試験よりも多くの被験者数を必要とするであろう**（例えば、ICH E1A参照）。



後半で説明

適切な被験者数を決定するために用いられる通常の方法を利用するためには、以下の項目を定めておくことが必要である。それらは、**主要変数、検定統計量、帰無仮説、選択された用量での対立（「作業」）仮説（その用量と選ばれた対象集団で検出すべき又は棄却すべき試験治療の差を考慮することも含めて）、誤って帰無仮説を棄却する確率（第一種の過誤）及び誤って帰無仮説を棄却できない確率（第二種の過誤）**であり、**更に試験治療を中止した被験者及び治験実施計画書違反を取り扱う方法も定める必要がある。**検出力の評価のために、イベント発生率が主要な関心事項となる場合には、試験に必要なイベント数から最終的な被験者数を外挿するための仮定も置くべきである。

ICH-E9 臨床試験のための統計的原則（平成10年11月30日付医薬審第1047号）より引用

人 ICH-E9における例数設計の記述（2）

■III. 試験計画上で考慮すべきこと, 3.5 必要な被験者数

被験者数を計算する方法は、計算に用いる見積値（分散、平均値、反応割合、イベント発生率、検出すべき差）とともに、治験実施計画書に定めておくべきである。また、これらの見積値の根拠も示すべきである。これらの仮定からの様々なずれに対して、必要な被験者数がどの程度変わり易いか調べることは重要であり、このためには実際に起こりうるずれの範囲に対応する被験者数の範囲を示すことで実施することが容易になるであろう。検証的試験では、通常これらの仮定は公表されたデータ又は先行する試験の結果に基づくべきである。検出すべき試験治療の差は、患者の治療管理を行う上で臨床的に意味をもつ最小限度の効果に関する判断又は新しい試験治療の予想される効果の方が大きい場合にはその効果に関する判断に基づいて決まるものであろう。**慣例的に、第一種の過誤は5%以下に設定され、多重性を考慮するために必要な調整がなされる場合はそれに従って設定される。**検証すべき仮説のもっともらしさ及び検定結果に望む影響力の強さにより、第一種の過誤の的確な選択に影響が及ぼされるであろう。**第二種の過誤は、慣例として10%～20%に設定される。**第二種の過誤を実施可能な範囲でできる限り小さくすることは、特に繰り返すことが困難又は不可能な試験の場合、治験依頼者の利益となる。慣例として用いている第一種の過誤の値及び第二種の過誤の値とは異なる値を用いることも許容される場合があり、むしろそれが好ましいこともあり得る。

被験者数の計算は、主要な解析で用いる解析対象集団に基づくべきである。解析対象集団が「最大の解析対象集団」である場合、効果の大きさに関する見積値は、治験実施計画書に適合した対象集団の場合に比べて小さくする必要があろう。これは、試験治療を中止した被験者又は服薬遵守状況の悪い被験者を解析に含めることにより、試験治療の効果が薄められることを考慮するためである。このときばらつきに関する仮定も再検討する必要があるであろう。

(...)

ICH-E9 臨床試験のための統計的原則（平成10年11月30日付医薬審第1047号）より引用

人 例数設計において考慮する事項

項目	説明
主要変数	目標症例数は、試験の主要な目的で決める
検定統計量、帰無仮説、対立仮説	分散、平均値、反応割合、イベント発生率や設定した仮説、その仮説を検証するための解析手法など
効果の大きさ	<ul style="list-style-type: none"> ● 先行研究などから見積もる ● 見積値の根拠も示す
第一種の過誤 (α)	<ul style="list-style-type: none"> ● 誤って帰無仮説を棄却する確率 ● 慣例的に5%以下（多重性問題がある場合は下げる）
第二種の過誤 (β) または 検出力 ($1-\beta$)	<ul style="list-style-type: none"> ● 誤って帰無仮説を棄却できない確率 ● 慣例的に10%~20%
例数比	同数であれば1:1
試験治療を中止した被験者及び治験実施計画書違反を取り扱う方法	検出力の低下が見込まれる場合など
主要な解析で用いる解析対象集団	Intent To Treat (ITT) やFull Analysis Set (FAS) など

上表の情報をプロトコルに記述する。

人 効果の大きさをどのように特定するか？

項目	説明
医学専門家の判断によるアプローチ	治療前と治療後の患者の健康状態を比較し、この変化を、より身近な結果を用いて改善または悪化を示した参加者と関連付けることで達成できる。
分布のばらつきに基づいて値を決定するアプローチ	内在する不正確さよりも大きく、したがって有意差の最小レベルを表すと思われる値を使用する。
経済評価の原則を用いたアプローチ	コストと健康アウトカムの両方を含み、意思決定者が支払ってもよいと考える健康効果単位のコストの閾値を定義し、治療の全体的な純便益を推定する。
医療従事者や患者などから得られた意見に基づいたアプローチ	専門家集団の結成、専門家団体や患者団体の会員への調査、個人への聞き取りなどのアプローチが考えられる。この意見聴取のプロセスは、臨床試験の文脈の中で明示的に構成することができる。
パイロット試験	期待値を導き、試験の適切な目標差を決定するための証拠や経験がほとんどない場合に実施されることがある。
エビデンスに基づくレビュー	研究課題に関する現在のエビデンスを用いて、目標とする差異を導き出すことができる。RCTのシステマティックレビューやメタアナリシスから得ることができる。ランダム化されたエビデンスがない場合は、観察研究からのエビデンスを同様の方法で使用することができる。もう一つの方法は、重要な差異が決定された研究のレビューを行うことである。
標準的な効果量	Cohenのdなどの効果量で設定されている小（0.2）、中（0.5）、大（0.8）といったカットオフ値を差の値として定義する。

人 例数設計の実例と書き方

<課題>

EPA含有サプリメント（EPA）のLDL-Cに対する有効性を証明する研究を行いたい。先行研究では、LDL-Cのベースラインからの変化率（%）に関して、EPA摂取群とプラセボ群の平均が、それぞれ-40%と-30%であり、共通の標準偏差が10%であった。この時、例数比が1:1、有意水準を両側5%とし、対応のない検定に対して検出力が80%となる症例数を計算したい。

$$n_T = \left(\frac{1+\varphi}{\varphi} \right) \frac{(Z_{\alpha/2} + Z_{\beta})^2}{\delta^2} = \frac{2(1.960 + 0.842)^2}{1^2} = 16$$

例数比=1:1

$Z_{\alpha/2} = 1.960$
 $Z_{\beta} = 0.842$

効果量 = (平均1-平均2)/標準偏差

<書き方>

目標症例数は、*****試験（UMINXXXXXXXXXX）の結果から算出した。*****試験の主要アウトカムである「LDL-Cの変化率」は、EPA群で-40%、プラセボ摂取群で-30%であった。群間差は、-10%であり、統合した標準偏差は10%であった。統計学的有意水準を5%、統計学的検出力を80%として症例数を計算すると、32名（各群16名）が目標症例数であった。加えて、試験期間中の脱落やプロトコルの遵守違反などを考慮し、最終的な必要症例数をXX名（各群XX名）とした。

- 1 例数設計が必要な理由
- 2 例数設計の実際
- 3 例数設計に対するオルトメディコの取り組み
- 4 本日のまとめ

人 安全性試験の症例数についての調査

機能性表示食品の届出等に関するガイドライン（2024年4月1日付け消食表第217号）

記載なし

ICH E1 致命的でない疾患に対し長期間の投与が想定される新医薬品の治験段階において安全性を評価するために必要な症例数と投与期間について

期間が6ヶ月以下の臨床試験	期間が12ヶ月の臨床試験
<ul style="list-style-type: none">● 短期間のうちに発現する有害事象の発現率（例えば、3カ月間の累積発現率が約1%）が明らかにされることが期待されている。● 妥当な頻度（一般的には0.5~5%程度）の遅発性の有害事象が観察できるとともに、より高頻度に発現した有害事象がその後の期間中に増加するのか、あるいは減少するのかを観察できるだけの十分な症例数が必要である。通常300~600例の対象症例数が適当である。	<ul style="list-style-type: none">● 100例の患者に対して最低1年間投与して得られた成績は、安全性データベースの一部として採用できると考えられる。そのようなデータを得るためには、治験薬を予定される臨床用量で少なくとも1年間投与するように適切に計画されたプロスペクティブな試験を実施すべきである。1年間の投与期間中に何ら重篤な有害事象が認められない場合には、そのような有害事象の1年間の累積発現率は3%未満と考えてよい。

要指導医薬品の製造販売後調査等の実施方法に関するガイドライン（2014年に廃止）

0.1%の発現頻度の副作用を95%以上の確率で少なくとも1件以上検出するために必要な症例数として**原則3000症例以上**を収集する。

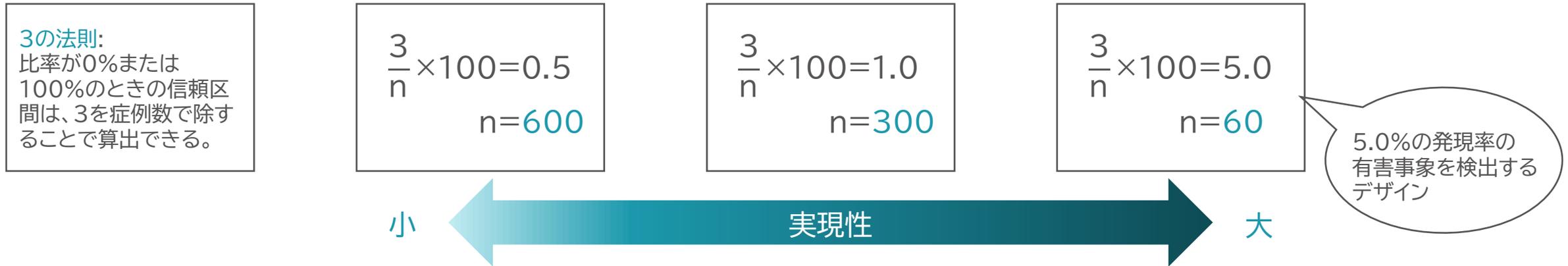
$$\rightarrow n = \frac{\log(1-p)}{\log(1-r)} = \frac{\log(1-0.95)}{\log(1-0.001)} \approx 2995 \approx 3000$$

人 有害事象発現率の妥当性と症例数の探索

安全性試験の症例数は、治験では安全性試験における症例数の明確な定義がなされているのに対して、機能性表示食品や特定保健用食品のガイドラインではあいまいな表現にとどまっている。そのため、発現率の根拠や実現可能性の高い症例数を提案する。

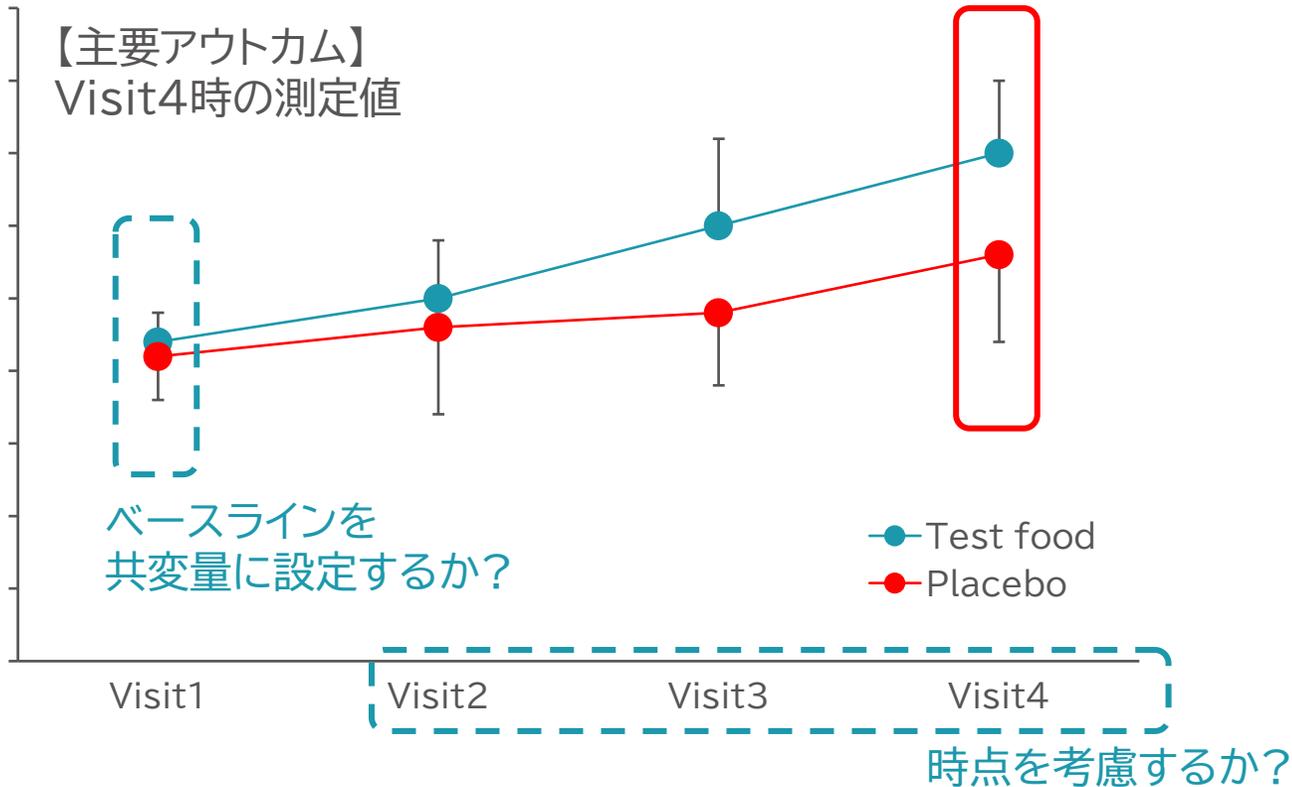


ICH E1では「例えば、3カ月間の累積発現率が約1%」や「一般的には0.5~5%程度」といった記載があるので、0.5%、1.0%、5.0%について3の法則を利用して考える。



食品を介入とした臨床試験を蓄積していき、より妥当性のある発現率の探索を行い情報の提供を目指す。

人 シミュレーションによる例数設計を実装し、複雑な解析モデルに対応する



層別ブロックランダム化を行っている場合は、
層別因子を解析モデルに投入するか？

単純なt検定であれば、話は簡単であるが...



食品の介入効果は医薬品よりも小さいので、
解析の精度を向上させる必要がある。

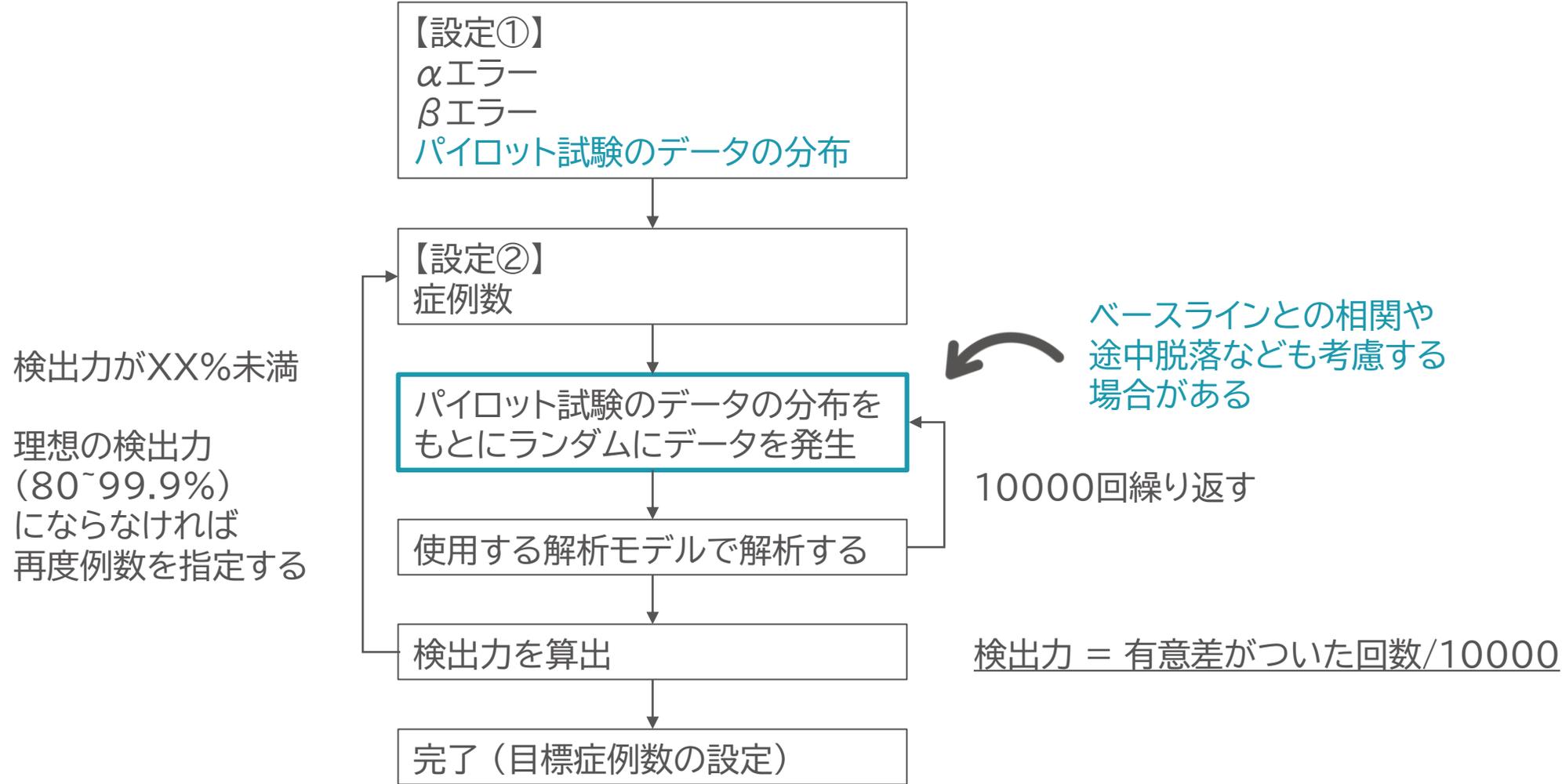


解析モデルが複雑になっていくが、公式が
存在しないため、例数設計が困難な場合がある。



モンテカルロシミュレーションの手法を
利用して例数設計を行うことで
現実に近い状況で例数設計を行うことができる

シミュレーションによる例数設計のアルゴリズム



様々な解析モデルに対応した例数設計を提供

- 1 例数設計が必要な理由
- 2 例数設計の実際
- 3 例数設計に対するオルトメディコの取り組み
- 4 本日のまとめ



例数設計は、試験の質を担保するためにも必要である。



機能性表示制度においても今後指摘される可能性がある。



安全性の例数設計は有効性試験と異なるアプローチが必要である。



シミュレーションによる例数設計で、様々な解析モデルに対応できる。

人 参考文献一覧

- 消費者庁. 機能性表示食品の届出等に関するガイドライン (2024年4月1日付け消食表第217号) (2024年4月1日アクセス可能: https://www.caa.go.jp/policies/policy/food_labeling/foods_with_function_claims/assets/notice_240401_01.pdf)
- Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *Trials*. 2010; 11: 32. (PMID: [20334632](https://pubmed.ncbi.nlm.nih.gov/20334632/))
- Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin JA, Doré CJ, Parulekar WR, Summerskill WS, Groves T, Schulz KF, Sox HC, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med*. 2013; 158 (3): 200-7. (PMID: [23295957](https://pubmed.ncbi.nlm.nih.gov/23295957/))
- 消費者庁. 機能性表示食品制度における臨床試験及び安全性の評価内容の実態把握の検証・調査事業 報告書 (平成29年3月) (2022年8月19日アクセス可能: https://www.caa.go.jp/policies/policy/food_labeling/foods_with_function_claims/pdf/foods_index_23_171025_0001.pdf)
- 厚生労働省. ICH E9 臨床試験のための統計的原則 (平成10年11月30日付医薬審第1047号) (2022年8月19日アクセス可能: <https://www.pmda.go.jp/files/000156112.pdf>)
- Cook JA, Hislop J, Altman DG, Fayers P, Briggs AH, Ramsay CR, Norrie JD, Harvey IM, Buckley B, Fergusson D, Ford I, Vale LD; DELTA group. Specifying the target difference in the primary outcome for a randomised controlled trial: guidance for researchers. *Trials*. 2015; 16: 12. (PMID: [25928502](https://pubmed.ncbi.nlm.nih.gov/25928502/))
- Hislop J, Adewuyi TE, Vale LD, Harrild K, Fraser C, Gurung T, Altman DG, Briggs AH, Fayers P, Ramsay CR, Norrie JD, Harvey IM, Buckley B, Cook JA; DELTA group. Methods for specifying the target difference in a randomised controlled trial: the Difference ELicitation in TriAls (DELTA) systematic review. *PLoS Med*. 2014; 11 (5): e1001645. (PMID: [24824338](https://pubmed.ncbi.nlm.nih.gov/24824338/))
- 厚生労働省. ICH E1 致命的でない疾患に対し長期間の投与が想定される新医薬品の治験段階において安全性を評価するために必要な症例数と投与期間について (平成7年5月24日付薬審第592号) (2022年8月19日アクセス可能: <https://www.pmda.go.jp/files/000156199.pdf>)
- 厚生労働省. 要指導医薬品の製造販売後調査等の実施方法に関するガイドライン (平成26年6月12日付薬食審査発0612第5号) (2022年8月19日アクセス可能: <https://www.mhlw.go.jp/file/06-Seisakujouhou-11120000-Iyakushokuhinkyoku/0000092788.pdf>)
- 岩崎学, 吉田清隆. 稀な事象の生起確率に関する統計的推測-Rule of Threeとその周辺-. *計量生物学*. 2005; 26: 53-63. (DOI: <https://doi.org/10.5691/jjb.26.53>)

問い合わせ・無料相談はコチラ

03-3818-0610
info@orthomedico.jp