臨床試験デザイン
- プロトコルの統計学的考慮点 -

手 良 向　聡
京都府立医科大学
生物統計学／附属病院臨床研究推進センター（CTREC）

金沢大学「臨床研究実施のための講習会」、2023.12.5

内容

- 試験実施計画書
- 評価項目
- ランダム化
- 統計的な研究仮説
- 統計学的考察
  - 標本サイズ、解析対象集団、適応的デザイン、サブグループ解析

臨床試験に必要なもの

- 研究仮説
- 試験実施計画書（試験プロトコル）
  - 診断・評価の基準・手順
  - 統計解析計画
- 実施体制
  - 実施主体（Sponsor）
  - 実施医療機関（Investigator）
  - データの質管理・質保証システム
    - データマネジメント、統計解析、モニタリング、監査
    - 独立データモニタリング委員会

試験実施計画書の標準様式

0. 概要
1. 目的
2. 背景と根拠
3. 本試験で用いる基準・定義
4. 試験薬／試験機器情報
5. 適格基準
6. 登録、割付
7. 倫理
8. 試験治療計画
9. 有害事象の評価・報告
10. 臨床検査、観察、調査項目・スケジュール
11. データの収集
12. 目標症例数と試験予定期間
13. 評価項目
14. 統計学的事項
15. 試験実施計画書の遵守又は変更並びに改訂
16. 試験の終了又は中止
17. 補償
18. 利益相反と資金源
19. 金銭の支払いに関する取り決め
20. データの質管理・質保証
21. 記録の保存
22. 研究内容の発表
23. 文献
24. 付録（実施体制など）

試験実施計画書の科学的妥当性

- 背景と根拠　Background and rationale
- 目的　Objective
- 適格基準　Eligibility criteria
- 治療計画　Treatment plan
- 評価項目　Endpoint

- 試験デザイン　Study design

- 解析方法　Statistical method

臨床家主導

統計家主導

試験デザイン

- 対照の選択　Selection of controls
- ランダム化の有無・方法　Randomization
- 盲検化の有無・方法　Blinding/Masking
- 統計的な研究仮説　Statistical hypothesis
- 比較形式（並行群間比較　Parallel、クロスオーバー　Crossover、用量漸増　Dose escalation など）
- 中間解析の計画・方法　Interim analysis plan/method

- 主要評価項目に基づいて
- 主要な解析方法を決めて
- 統計的仮説（帰無仮説、対立仮説）、
  許容できるエラー（有意水準、検出力）を設定して

標本サイズ（目標症例数）設定　Sample size determination

評価項目

- 試験の目的に関連する仮説を検証するうえで臨床的に意味があり、客観的に評価できる観察・検査項目またはそれらの合成指標

  - 主要（Primary）と副次（Secondary）に分ける
  - 各患者について定義する
    - 例えば、「〇〇週後の改善の有無」は個人に対して定義されるので妥当な表現であるが、「〇〇週後の改善割合」は集団のデータから統計的に推定されるものであり、対象集団や推定方法により変わり得るため、評価項目として妥当な表現ではない

臨床・代替評価項目

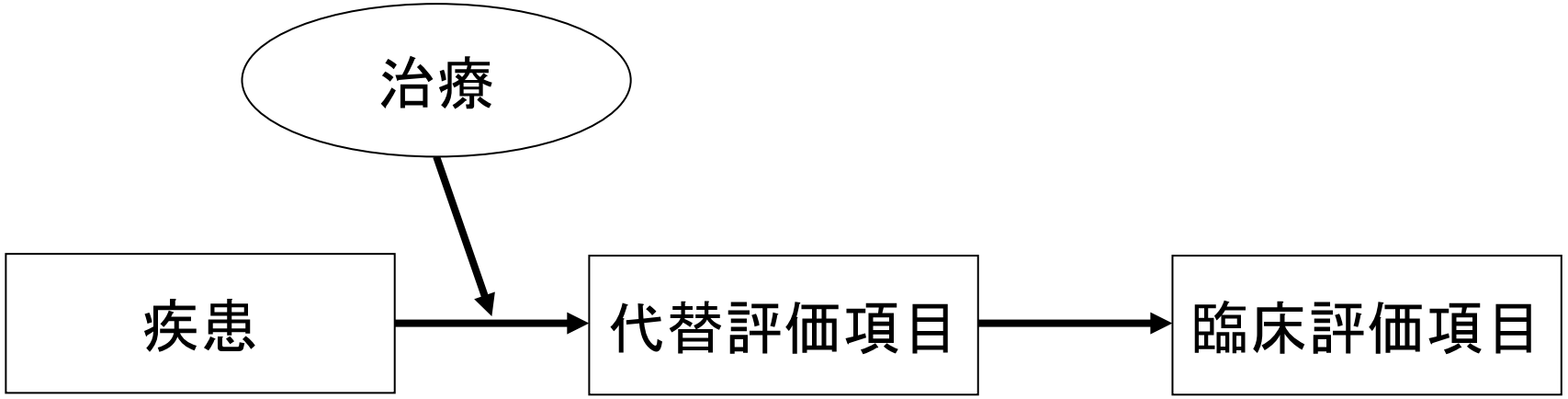
- 臨床評価項目 Clinical endpoint
  - 患者がどのように感じ, あるいは機能し, どのくらい生存しているかを反映する特性あるいは変数
- バイオマーカー Biomarker
  - 正常な生物学的プロセス, 病態形成プロセスあるいは治療的介入に対する薬理学的反応の指標として客観的に測定および評価されるある特性
- 代替評価項目 Surrogate endpoint
  - バイオマーカーのうち、臨床評価項目の代わりになることが意図されたもので, 疫学, 治療学, 病態生理学または他の科学的根拠に基づき, 臨床上の便益・害の有無を予測することが期待されるもの

Biomarkers Definitions Working Group. Clin Pharmacol Ther 2001;69:89-95

代替評価項目の候補

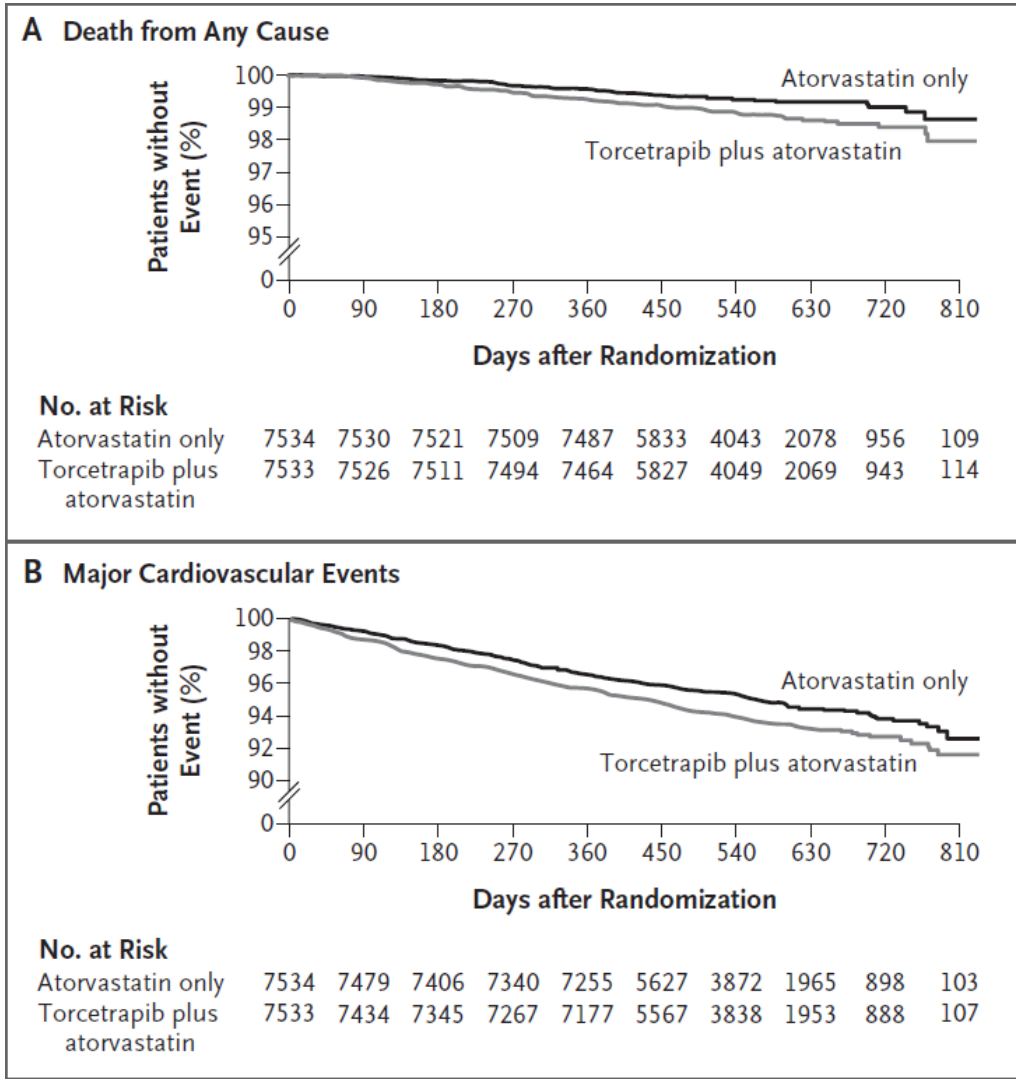
代替評価項目 Surrogate endpoints	臨床（真の）評価項目 Clinical/True endpoints
血中コレステロール値	心疾患の発生
血圧値	心疾患、脳血管障害の発生
血糖値	糖尿病合併症の発生
腫瘍縮小効果	癌死
不整脈	心臓死
心機能指標	心不全の発生
血管狭窄度	狭心症、心筋梗塞
眼圧	緑内障の視野狭窄

疾患、評価項目と治療との関係（理想）



代替評価項目が疾患過程の唯一の因果経路上に存在し、真の評価項目への治療効果が、完全に代替評価項目への効果を介して伝達される

ILLUMINATE（失敗例）



対象：心血管系疾患既往患者
試験薬：コレステリルエステル転送蛋白(CETP) 阻害薬: torcetrapib + atorvastatin

変化（12か月）

HDL-C: ＋0.5 mg/dl vs. ＋34.2mg/dl

LDL-C: ＋0.9 mg/dl vs. －21.5mg/dl

SBP: ＋0.9 mmHg vs. ＋5.4 mmHg

中間解析で早期試験中止！

全死亡：0.8% vs. 1.2%

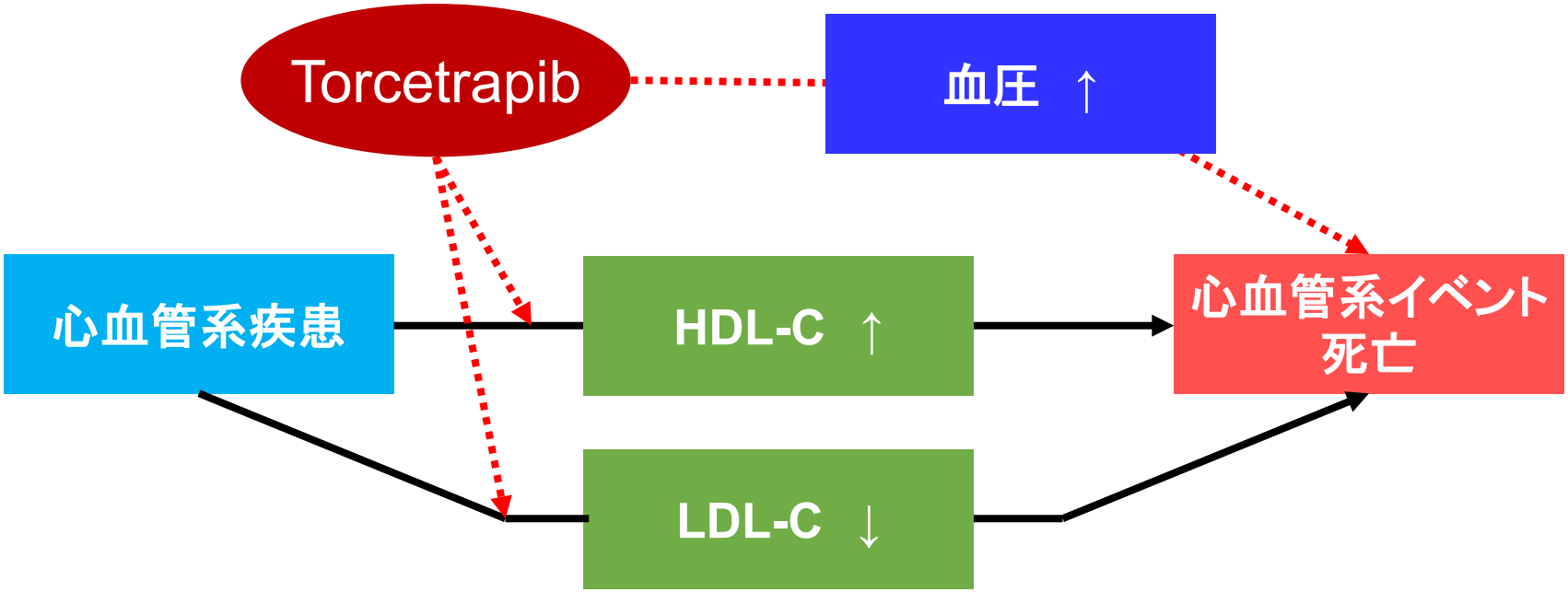
（ハザード比　1.58）

心血管イベント: 5.0% vs. 6.2%

（ハザード比　1.25）

NEJM 2007;357:2109-2122

疾患、評価項目と治療との関係
- ILLUMINATE -



治療が、代替評価項目を介す経路と介さない経路の両方に作用する可能性がある

Estimand（エスティマンド）

- 試験目的によって提起される、「臨床的疑問を反映した治療効果」の正確な叙述

Estimandの5つの要素

- 関心のある治療　Treatment
- 対象集団　Population
- 変数（評価項目）　Variable/Endpoint
  - 被験者ごとに得られる測定値
- その他の中間事象　Other intercurrent events
  - 中間事象：治療開始後に発現し、関心のある臨床的疑問に関連した変数の測定を不可能とする事象や変数を試験治療の効果として希釈する際に影響を与える事象
- 集団レベルの変数要約　Population-level summary

ICH E9ガイドライン (補遺)

中間事象をとり扱うストラテジー
Strategies for addressing intercurrent events

- 治療方針 Treatment policy
  - 中間事象を考慮しない
- 仮想　Hypothetical
  - 中間事象が起きなかったら、という仮想的状況を考える
- 複合変数　Composite variable
  - 中間事象を評価項目の一部にとり込む
- 治療下　While on treatment
  - 中間事象までの測定値のみを考える
- 主要層　Principal stratum
  - 潜在的な中間事象によって定義された集団を考える

日本製薬工業協会　医薬品評価部会　データサイエンス部会
「Estimandの治療実施計画書への実装」 2023.3

計量生物学 Vol. 41, No. 1, 37 − 54 (2020)

総　説

臨床試験におけるランダム化の意義と限界

Significance and limitations of randomization in clinical trials

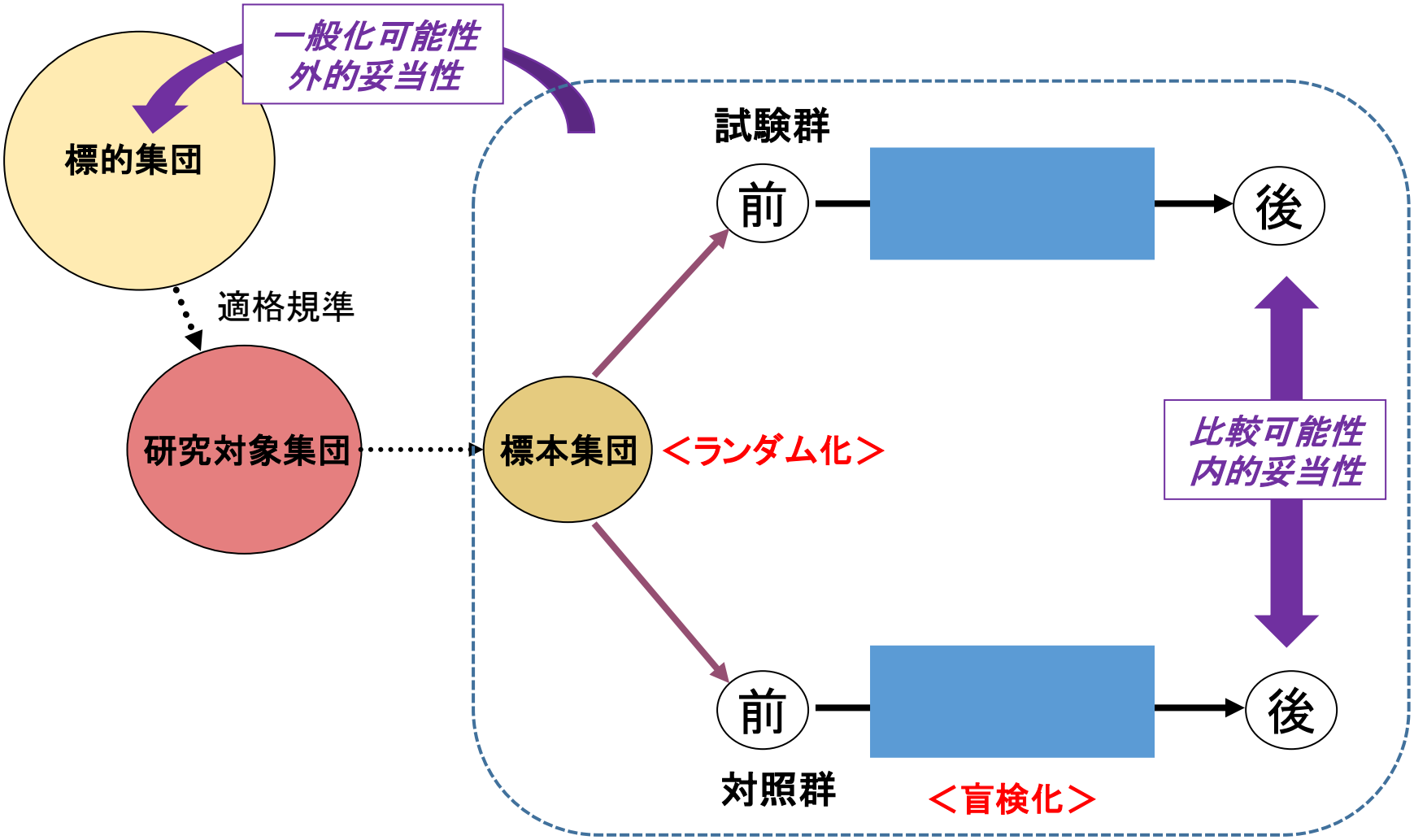
手良向聡[*1]
Satoshi Teramukai[*1]

[*1] 京都府立医科大学大学院医学研究科生物統計学
[*1] Department of Biostatistics, Graduate School of Medical Science, Kyoto Prefectural University of Medicine
e-mail : steramu@koto.kpu-m.ac.jp

https://www.jstage.jst.go.jp/article/jjb/41/1/41_37/_article/-char/ja/

ランダム化対照試験

RCT: Randomized controlled trial



実験計画におけるフィッシャーの3原則

- 局所管理　Local control
  - 系統誤差を減少させる

- ランダム化　Randomization
  - 系統誤差を偶然誤差に転化する

- 繰り返し　Replication
  - 偶然誤差を精確に推定する

実験に伴う2種類の誤差

- 偶然誤差　Random error
  - 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
  - 繰り返し測定を行えば推定可能
- 系統誤差　Systematic error（バイアス　Bias）
  - 圃場の肥沃度や日当たりの不均一性のように確率変数と見なせない誤差
  - 繰り返しには関係なく結果を歪める原因

臨床試験におけるランダム化の意義

1. 試験群間の比較可能性を高める
   - 交絡を制御する
2. 推論のための統計学的な基礎を与える
   - デザインに基づく解析が可能となる
3. 割付の隠匿化を行って選択バイアスを防ぐ
   - 医師の選択と割付とを無関係（独立）にする

→ 1.「試験群間の患者特性をバランスさせる」ことが重視され、2. と3. の意義はあまり理解されていない

手良向聡. 臨床試験におけるランダム化の意義と限界. 計量生物学 2020;41:37-54.

割付の隠匿化
Allocation concealment

- 割付結果が予見できてはいけない
- 割付結果を知った後に被験者を選択してはいけない

- 盲検化（マスク化）とは異なる
  - 盲検化は、治療開始から治療（評価）終了まで、割付内容がマスクされている

ランダム化の方法

- 固定的ランダム化 FR: Fixed randomization
  - 割付確率は試験開始時に固定
    - 単純ランダム化 Simple randomization
    - 並べ替えブロックランダム化 Permuted block randomization
    - 層別ランダム化 Stratified randomization
- 適応的ランダム化 AR: Adaptive randomization
  - 割付確率は試験進行中に変動
    - 割付数適応的（偏コイン法） Number-adaptive
    - 共変量適応的（最小化法） Covariate-adaptive
    - 反応適応的 Response-adaptive

注：層別ランダム化に用いた層別因子、適応的ランダム化に用いた割付調整因子は統計解析の際に考慮する → 層別解析

固定的ランダム化
- 層別並べ替えブロックランダム化 -

施設1　軽症　BAAB BABA AABB……

　　重症　ABBA ABAB BBAA……

　　ブロック（サイズ4）

施設2　軽症　……

　　重症　……

施設3　軽症　……

　　重症　……

……　……　……

層別因子：施設と重症度

適応的ランダム化
- 最小化法 -

割付調整因子	カテゴリー	A群	B群	次の被験者
施設	1	2	1	
	2	1	3	
	3	1	0	
	4	3	4	
	5	2	2	←
	…	…	…	
重症度	軽症	8	7	
	重症	7	8	←

A群の和：2+7=9 ＜ B群の和：2+8=10

↓

次の被験者はA群に確率p (>0.5)で割り付ける

盲検化（マスク化）

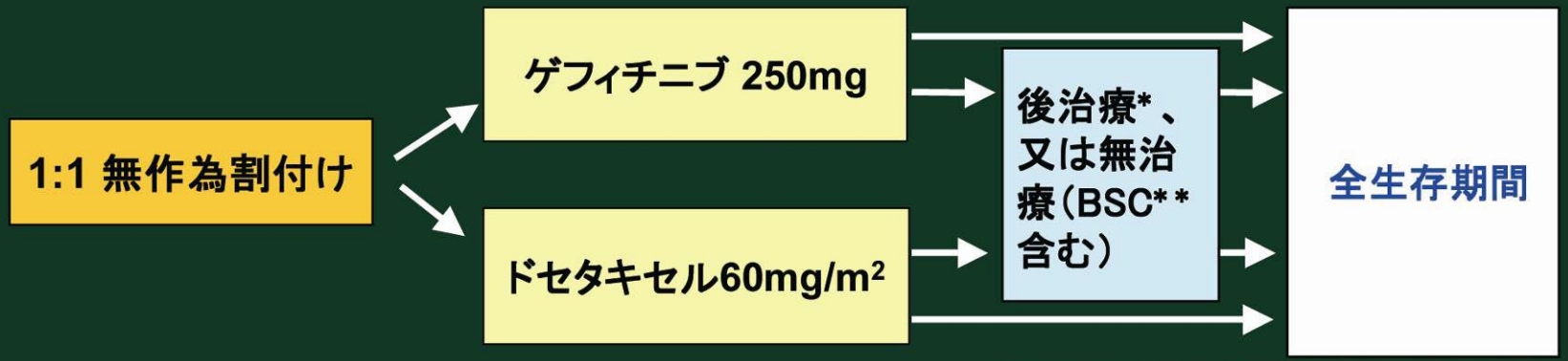
- 治療内容を知ることによるバイアスの防止策
  - 特に評価項目への影響が重要

種類	被験者	医師
非盲検 Unblinded/ オープンラベル Open-label		
単盲検 Single-blinded	盲検化	
二重盲検 Double-blinded	盲検化	盲検化

統計的な研究仮説

- 優越性試験　Superiority trial
  - 試験治療が対照治療に対して優れるかどうかを検定することを目的とした試験
- 非劣性試験　Non-inferiority trial
  - 試験治療が対照治療よりも事前に規定された非劣性マージンΔ以上は劣らないことを示すことを目的とした試験
  - 試験治療は効果以外に利点（投与が簡便、副作用の発現が少ない、安価など）を有することが前提
  - いくつかの懸念
    - 試験の質が悪いと劣っているものを劣らないと判断
    - 非劣性マージンΔの決め方

試験デザイン



＊後治療：割付け治療の中止後に行われる治療
＊＊BSC：Best Support Care（緩和療法）

主要評価項目

- 全生存期間
  - 共変量を考慮しない比例ハザードモデルに基づいたハザード比の信頼区間の上限が1.25以下であれば非劣性が結論づけられる
  - 最終解析の目標死亡例：296例

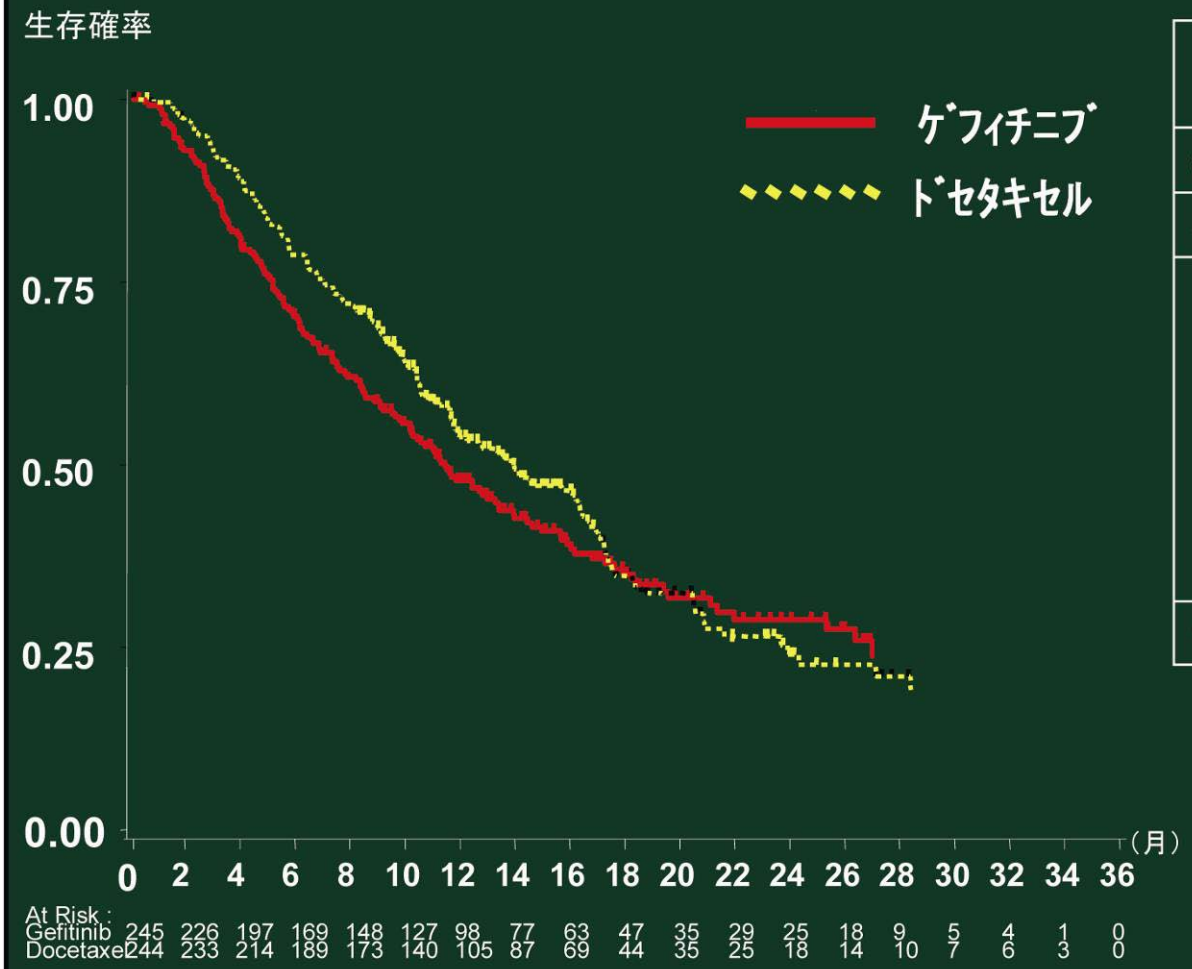
主な選択基準

- 進行/転移性（ⅢB期/Ⅳ期）又は術後再発の非小細胞肺癌患者
- 1又は2レジメンの化学療法治療歴（少なくとも1レジメンは白金製剤を含む）
- 年齢20歳以上
- 全身状態：WHO Performance Status（PS）が0～2

3

Maruyama R et al., Journal of Clinical Oncology 26 (2008), 4244-4252

主要評価項目－ 全生存期間（ITT*）



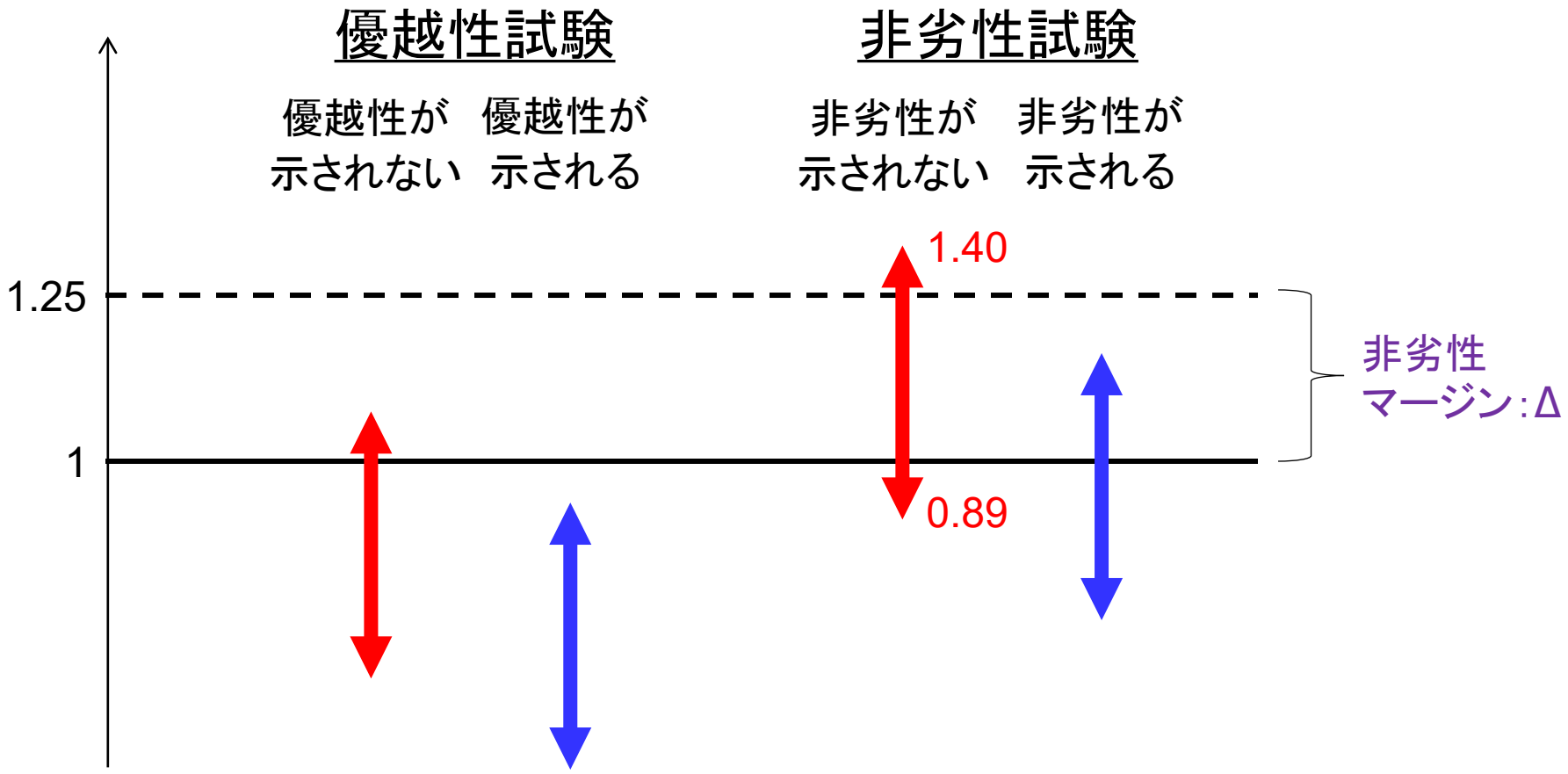
	ゲフィチニブ割付群	ドセタキセル割付群
症例数	245	244
イベント数	156	150
プロトコールで規定された主解析：共変量を考慮しないCox回帰分析 ハザード比（95.24%信頼区間）= 1.12（0.89, 1.40）、p=0.330 統計学的に非劣性は証明されなかった		
1 年生存率	48%	54%

*Intention-To-Treat：無作為割付された全ての患者のうちGCP違反の1例を除く

13

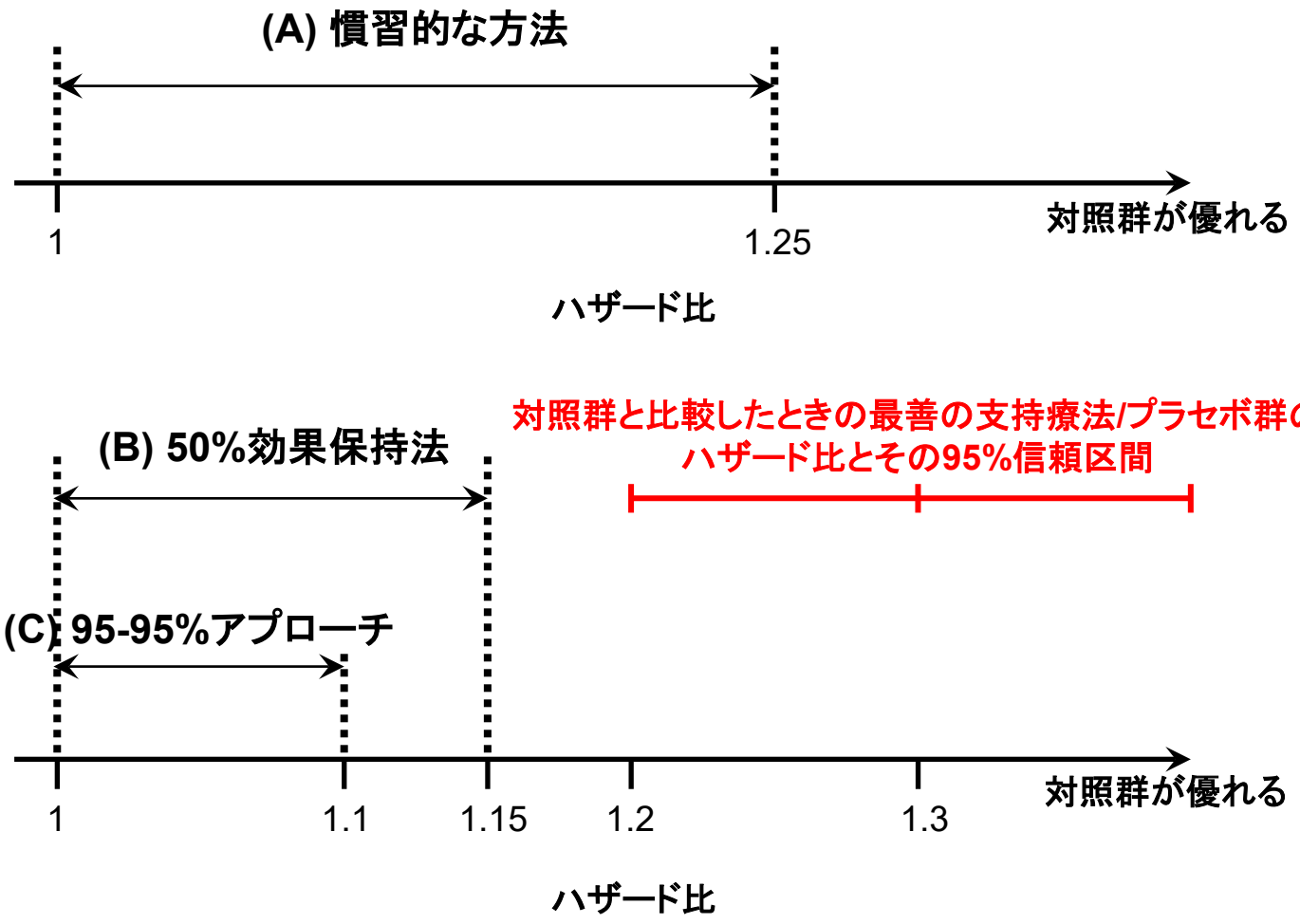
Maruyama R et al., Journal of Clinical Oncology 26 (2008), 4244-4252

優越性と非劣性の判断基準



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%信頼区間を表す

非劣性マージンの決め方



Tanaka S et al., CCR 2012;18:1837-1847

統計学的考察

① 標本サイズ（目標症例数）の設定根拠

Sample size determination (SSD)

② 解析対象集団 Analysis sets

③ 解析項目・統計手法 Statistical analysis methods

- 該当するのであれば以下の詳細

④ 適応的デザイン（中間解析、標本サイズ再設定）

Adaptive design (Interim analysis, Sample size re-estimation)

⑤ サブグループ解析

Subgroup analysis

① 臨床試験の症例数

- 科学性に重きを置くと、有意であるという結果を得やすくするために症例数を多くしたくなる
- 一方、倫理性に重きを置くと、試験の対象となる患者数を可能な限り減らしたくなる
  - しかし、症例数不足の試験を行って統計的に有意でないという結果が得られた場合に、真に効果がなかったのか、それとも真に効果があったにもかかわらず検出力不足のため有意でなかったのかを区別できないことは、無駄な試験を行ったという意味で非倫理的である

多すぎても少なすぎてもいけない

① 標本サイズ設定の一般論

- 一時的な仮定に基づく概算である
- ある仮定の下で、
  - 何人の対象者が必要か？（標本サイズ算出）
  - n人の対象者しか参加できない場合、実施する価値があるか？（検出力解析）
- 仮説検定を用いると、定式化が容易である

① 仮説検定

- 仮説の設定　Hypothesis
  - 帰無仮説　Null hypothesis : H_0
  - 対立仮説　Alternative hypothesis : H_1
- 検定統計量の選択　Test statistic
  - 通常は、S/N比（シグナルとノイズの比）
- 有意水準の設定　Significance level

計画段階

解析段階

- 統計的有意性の評価　Statistical significance
  - 帰無仮説とデータの乖離の指標（P値）を計算
  - P値と有意水準を比較 → 帰無仮説を棄却/受容

① 仮説検定における2種類の誤り

検定結果	母集団（真）	
	差なし	差あり
帰無仮説を受容 （差なし、と判断）	正しい	第II種の過誤 （βエラー）
帰無仮説を棄却 （差あり、と判断）	第I種の過誤 （αエラー）	正しい

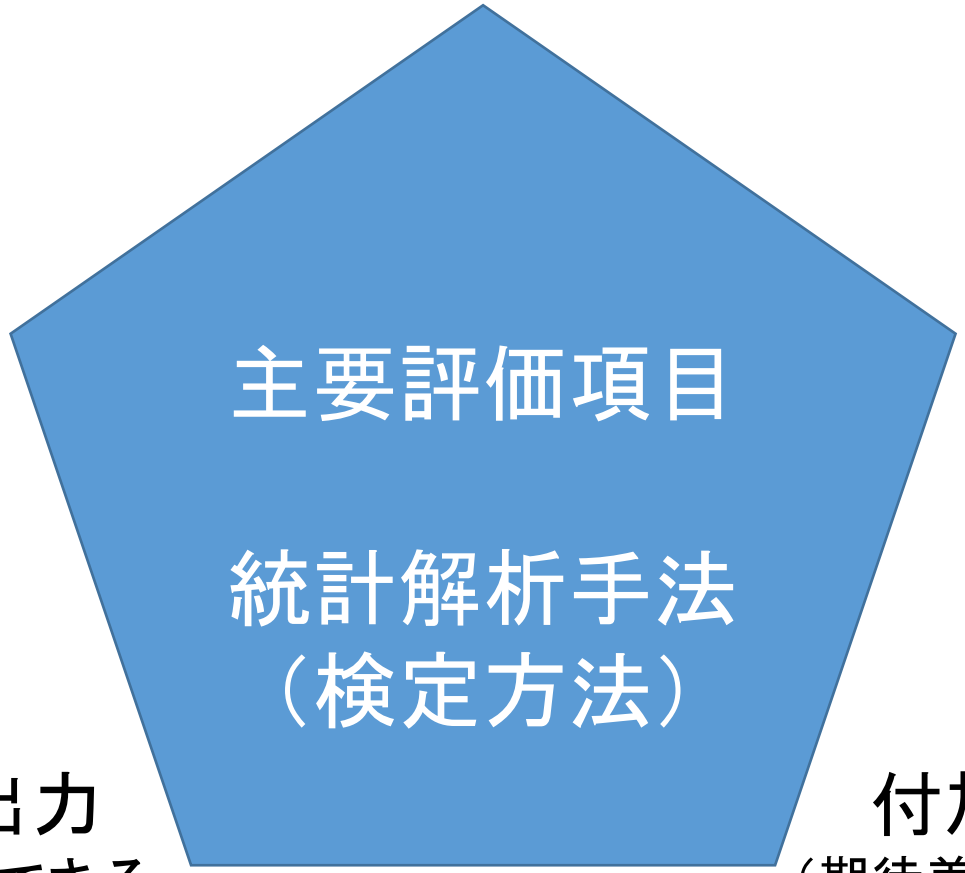
1-β：検出力 (power)

① 第I種・第II種の過誤確率

- 第I種の過誤（αエラー）＝消費者リスク
  - 効果のない医療技術が承認されるリスク
  - 規制当局（消費者の代表）が決める
    - 通常、片側0.025～0.1（両側0.05～0.2）

- 第II種の過誤（βエラー）＝生産者リスク
  - 効果のある医療技術が承認されないリスク
  - スポンサーが決めてよいが、大きくすると失敗のリスクが高まる
    - 通常、0.05～0.2

① 標本サイズ設定に必要な情報

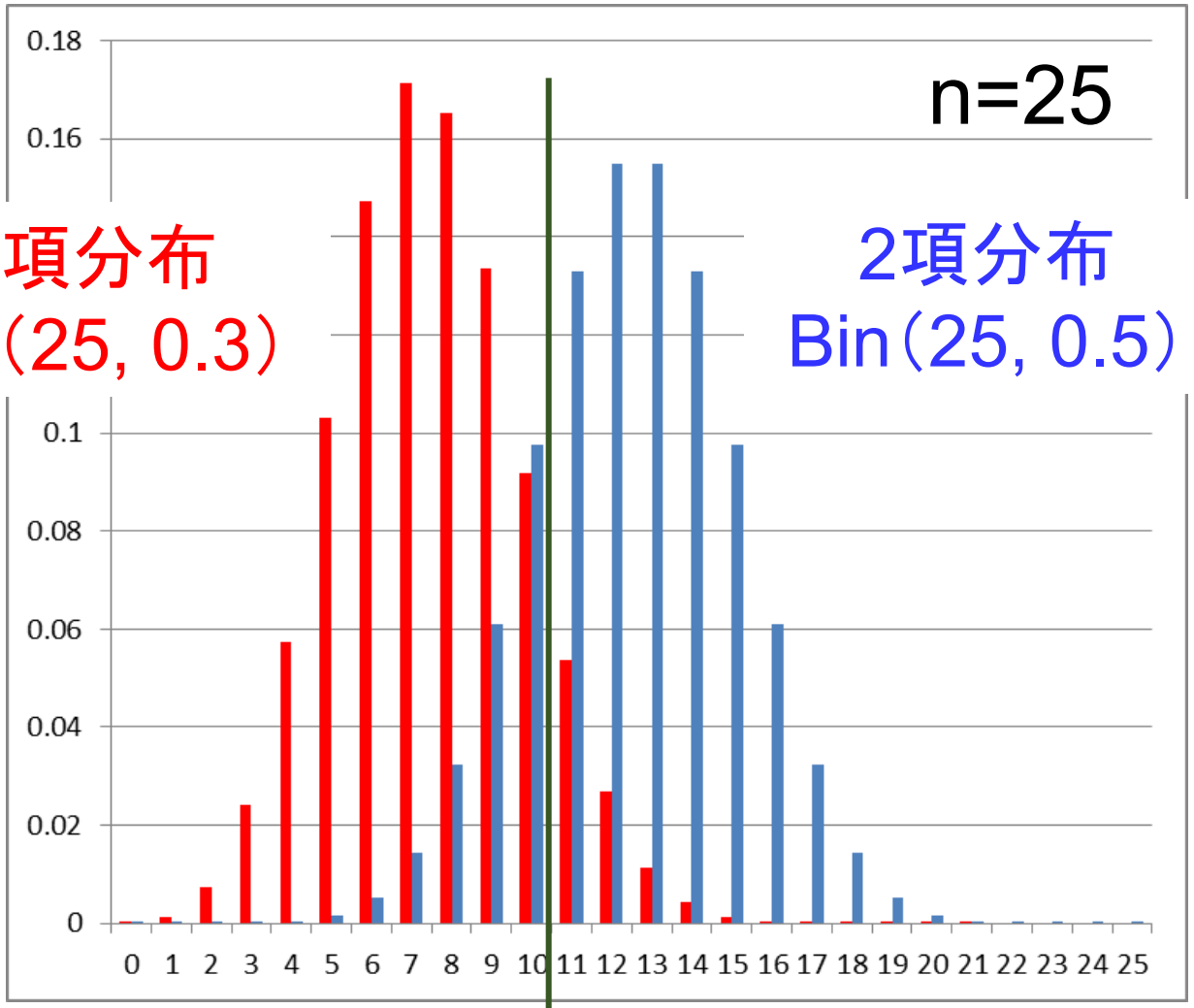


① 標本サイズ設定(2値変数、単群)

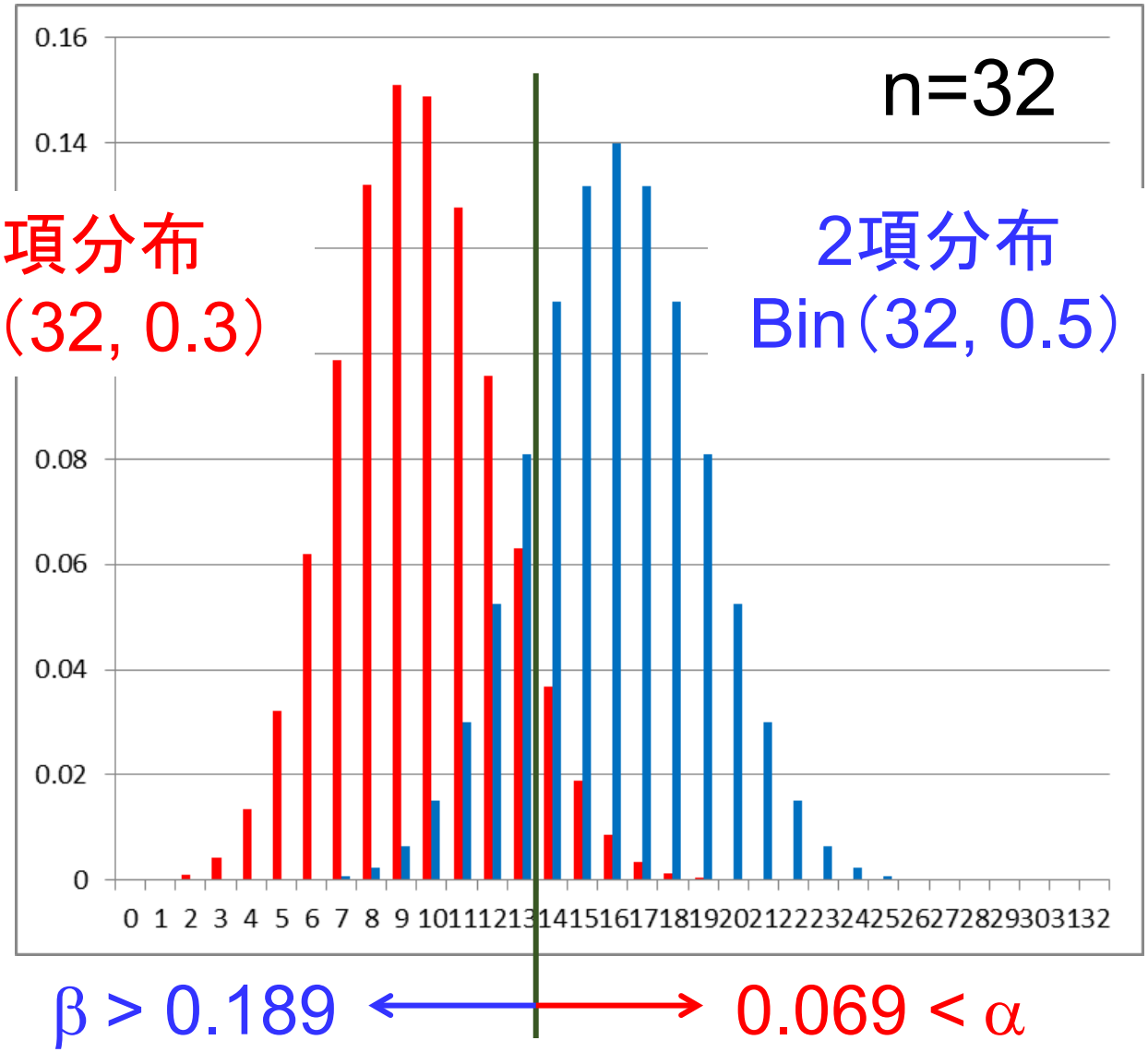
- θ: 成功確率
- 帰無仮説 H_0: $\theta \leq p_0$
- 対立仮説 H_1: $\theta > p_0$ ($\theta = p_1$)
  - p_0: 閾値
  - p_1: 期待値

- 2項検定

$p_0=0.3,\ p_1=0.5,\ \alpha=0.1,\ \beta=0.2$



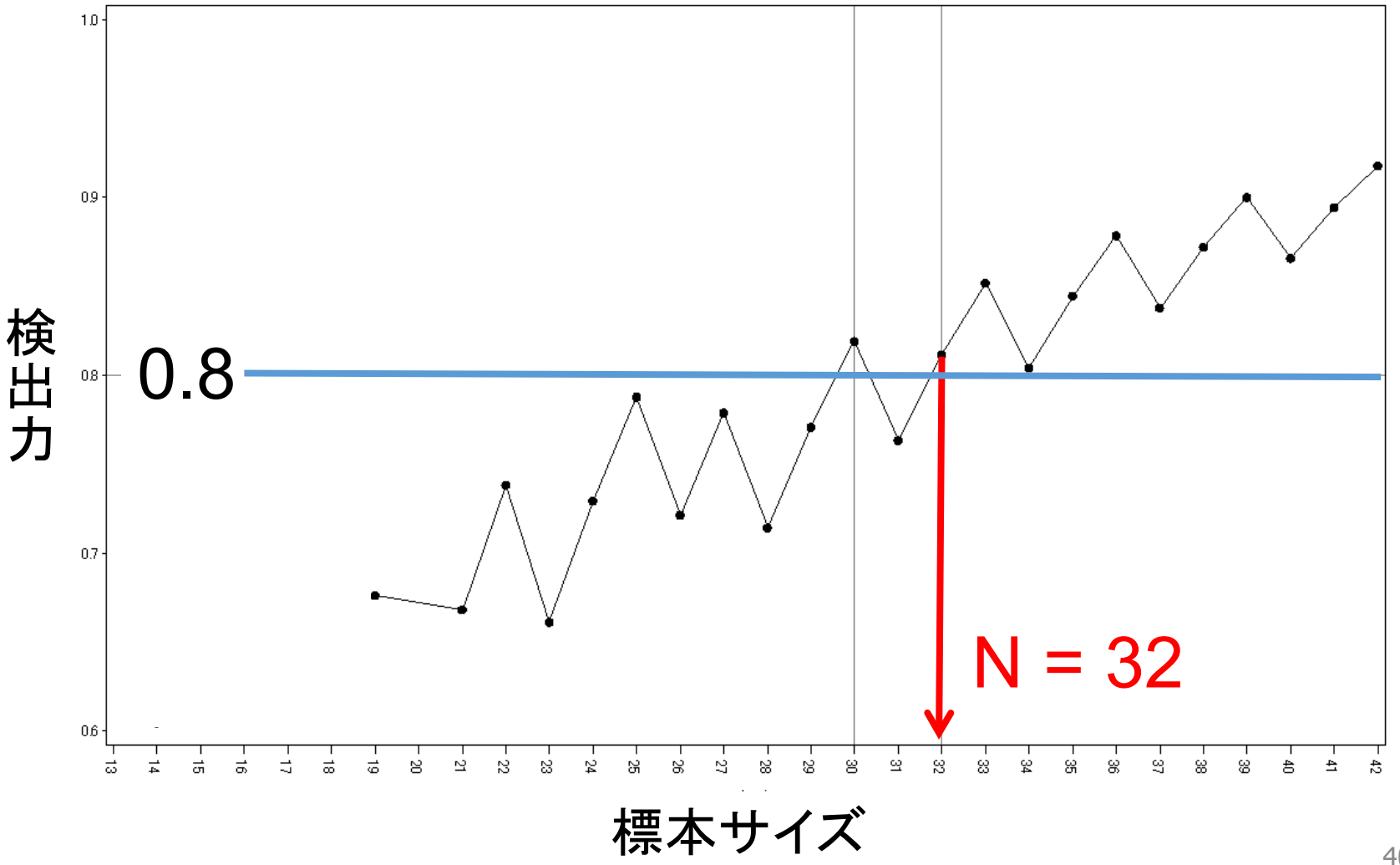
$p_0=0.3$, $p_1=0.5$, $\alpha=0.1$, $\beta=0.2$



無効と判断 u = 14 有効と判断

標本サイズと検出力の関係（2項検定に基づく方法）

$p_0=0.3$, $p_1=0.5$, $\alpha=0.1$, $\beta=0.2$

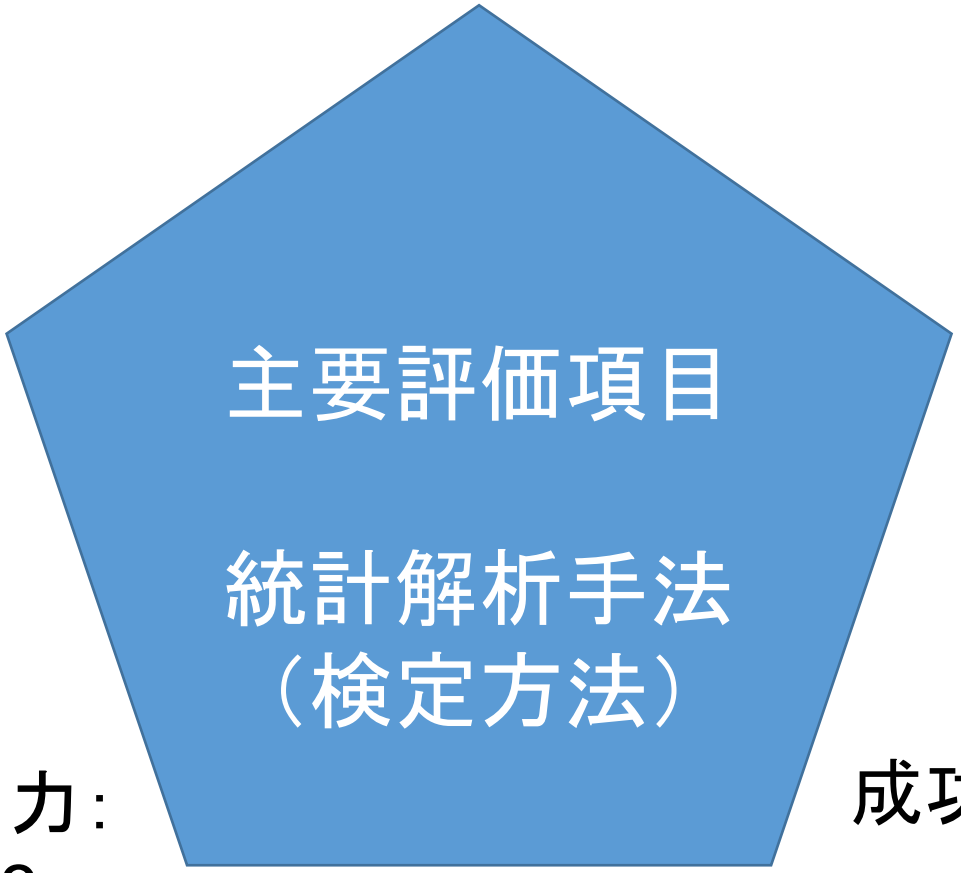


標本サイズ設定(2値変数、単群)

必要標本サイズ:32

有意水準:
0.1(片側)

帰無仮説と
対立仮説
(期待差:0.2)



検出力:
0.8

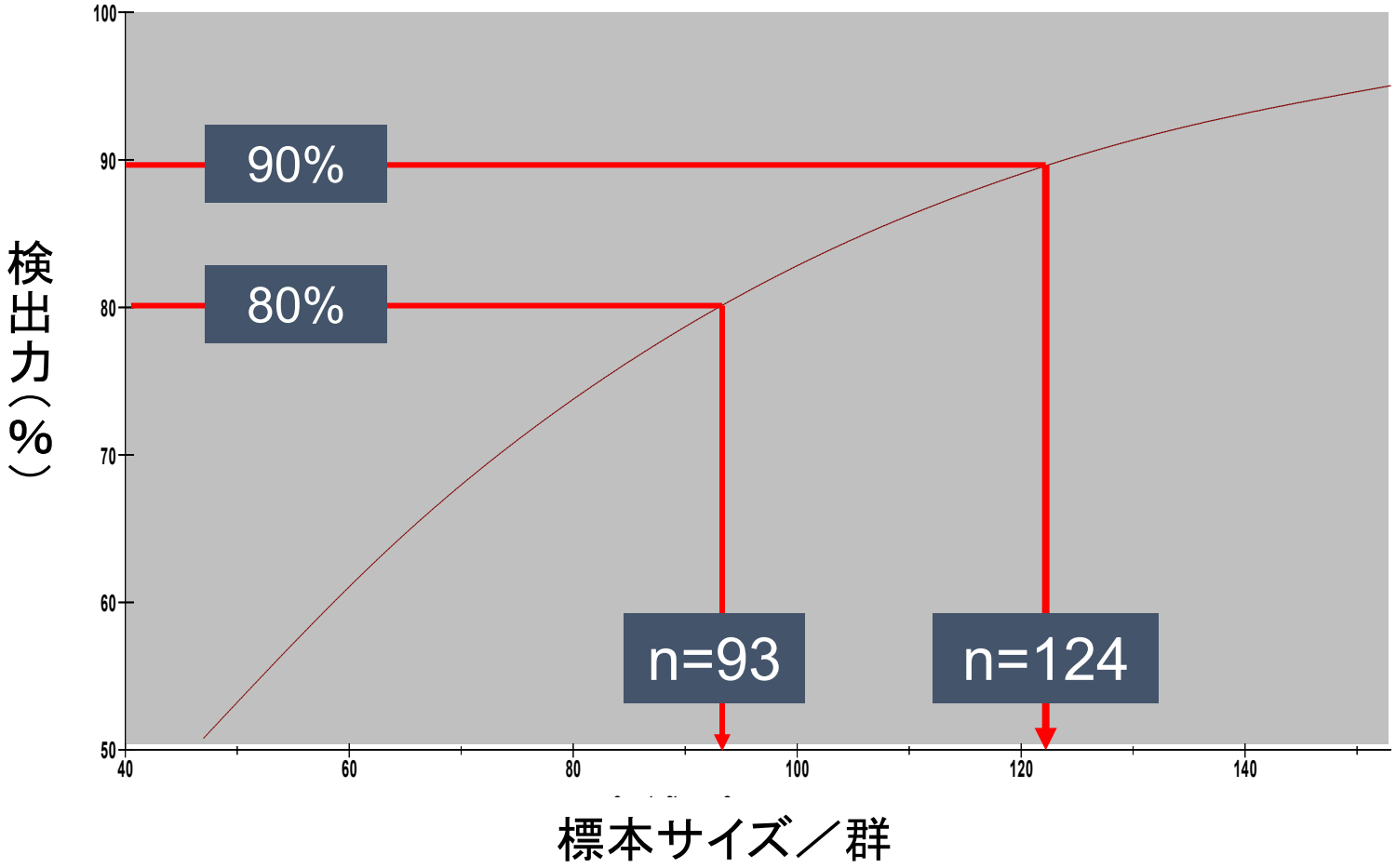
成功確率の閾値:
0.3

① 標本サイズ設定（2値変数、2群）

- 対照群の成功確率：0.3
- δ（期待差）：0.2
- カイ二乗検定
- α：0.05（両側）

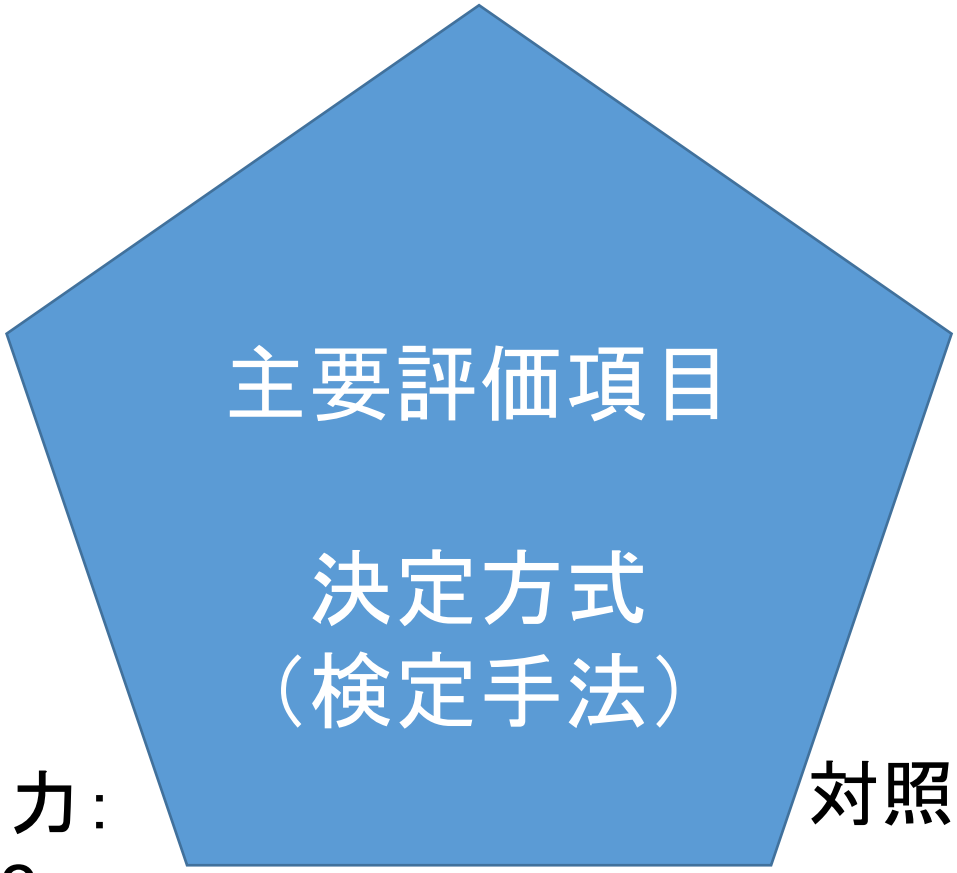
n（標本サイズ）と1-β（検出力）との関係は？

標本サイズと検出力の関係



標本サイズ設定(2値変数、2群)

必要標本サイズ:93/群=186



有意水準:
0.05(両側)

帰無仮説と
対立仮説
(期待差:0.2)

検出力:
0.8

対照群の成功確率:
0.3

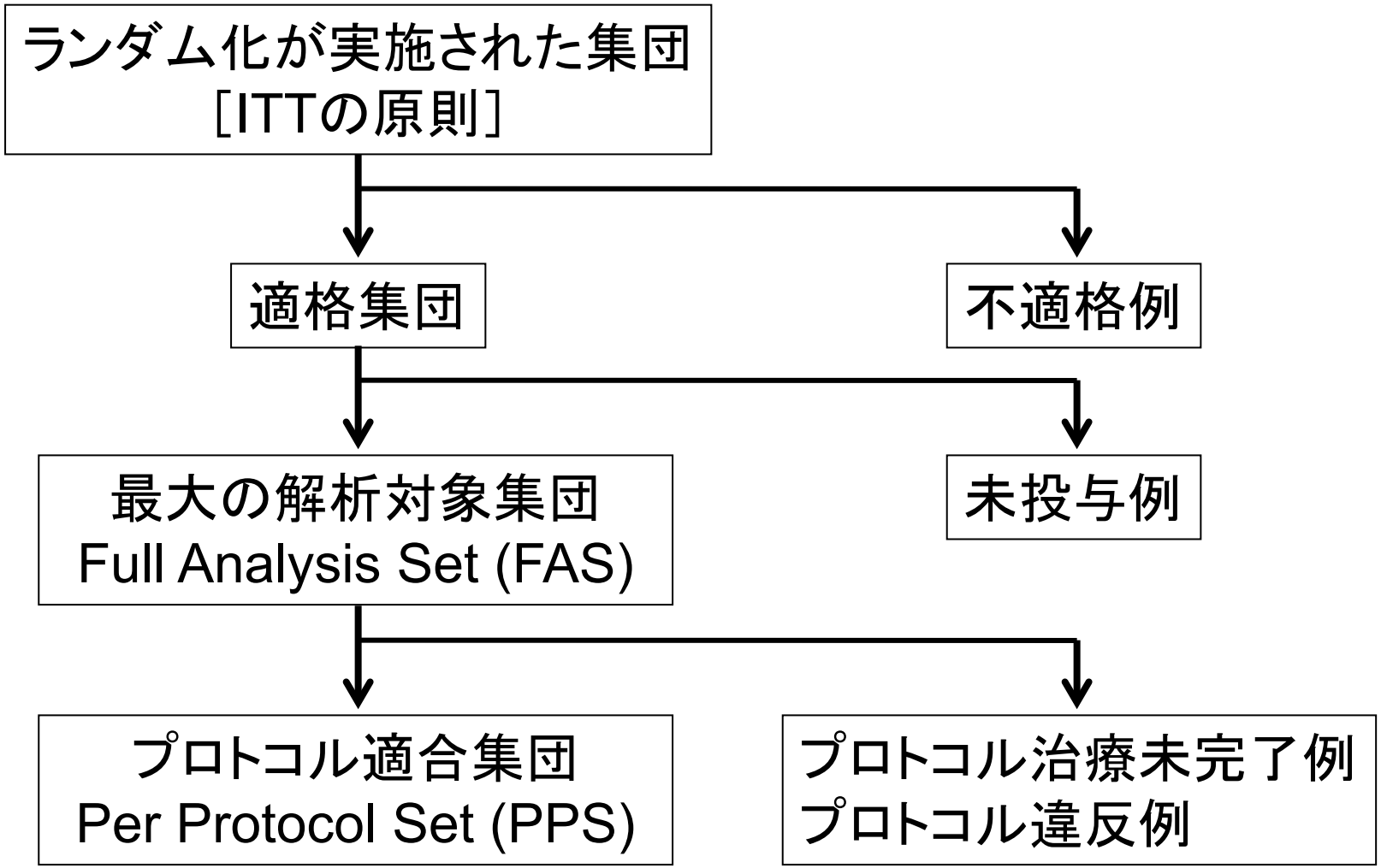
② ITTの原則

- 治療方針による効果は、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）に基づくことにより、最もよく評価できる、ことを主張する原則

ICH E9ガイドライン

  - 理想的環境における作用を評価するのではなく、治療遵守の程度なども含めた実践的な技術を評価する
  - ただし、これはPhilosophyであり、様々な解釈がある

② 解析対象集団


ランダム化が実施された集団
[ITTの原則]
適格集団
不適格例
最大の解析対象集団
Full Analysis Set (FAS)
未投与例
プロトコル適合集団
Per Protocol Set (PPS)
プロトコル治療未完了例
プロトコル違反例

③ 標準的な統計手法

目的	連続変数 Continuous	分類変数 Categorical	時間－イベント変数 Time-to-event
分布の記述 Description	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 （カプラン・マイヤー法）
要約統計量 Summary statistic	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年生存確率、 中央生存期間
検定（単純） Test (simple)	t検定、分散分析、 ウィルコクソン検定	カイ二乗検定、 フィッシャー正確検定、 ウィルコクソン検定	ログランク検定、 一般化ウィルコクソン検定
検定（層別） Test (stratified)	共分散分析	マンテル・ヘンツェル 検定	層別ログランク検定
回帰モデル Regression model	一般線形モデル	ロジスティックモデル	コックス比例ハザードモデル

④ 適応的デザイン

- A clinical trial design that allows for prospectively planned modifications to one or more aspects of the design based on accumulating data from subjects in the trial
- 当該試験における被験者の集積データに基づいて、デザインの1つ以上の側面に対する前向きに計画された変更を許容する臨床試験デザイン

FDA. Guidance for Industry.
Adaptive designs for clinical trials of drug and biologics, 2019.11.

④ 適応的デザインの型

- 比較データを用いない場合
  - 盲検下での標本サイズ再設定など
- 比較データを用いる（割付情報を開示する）場合
  - 中間解析（グループ逐次法など）
  - 標本サイズ再設定
  - 患者集団の選択（適応的エンリッチメントなど）
  - 治療群（用量群）の選択
  - 患者割付（適応的ランダム化）
  - 評価項目の選択
  - 上記の組合せ

④ 中間解析の定義・目的

- 臨床試験の中間時点で蓄積されたデータを評価すること
- なぜ実施するのか？
  - 倫理的ー ヘルシンキ宣言
    - 被験者の不利益を最小にする
  - 経済的 ⇒ 無益性の評価など
    - 社会的コストを最小にする
  - 管理的 ⇒ 適応的デザインなど
    - 試験の効率を向上させる

④ 中間解析の統計手法

- グループ逐次法（α消費関数）
  - 中間時点（通常は数回以内）で仮説検定を繰り返し、中止規準を満たすかどうかを決める方法
  - O’Brien-Fleming（オブライエン・フレミング）法が代表的
  - 主に有効性の評価に利用
- 確率打ち切り法
  - 中間データに基づいて最終結果を評価・予測する方法
  - 条件付き検出力、予測検出力、ベイズ流予測確率などを計算
  - 主に無益性の評価に利用

④ 多重性の問題

同じデータに対して、検定を繰り返し行うと、検定の回数に比例して、第I種の過誤（差がないものをあると言う誤り）が大きくなる、という問題

例：20本に1本の割合で不良品のワインが混入している貯蔵庫からランダムに2本ワインを選んだとき、2本のうちどちらかが不良品である確率は？

答え：2本ともまともなワインである確率 ＝（1本がまとも）×（1本がまとも）
＝ 0.95 × 0.95 ＝ 0.9025

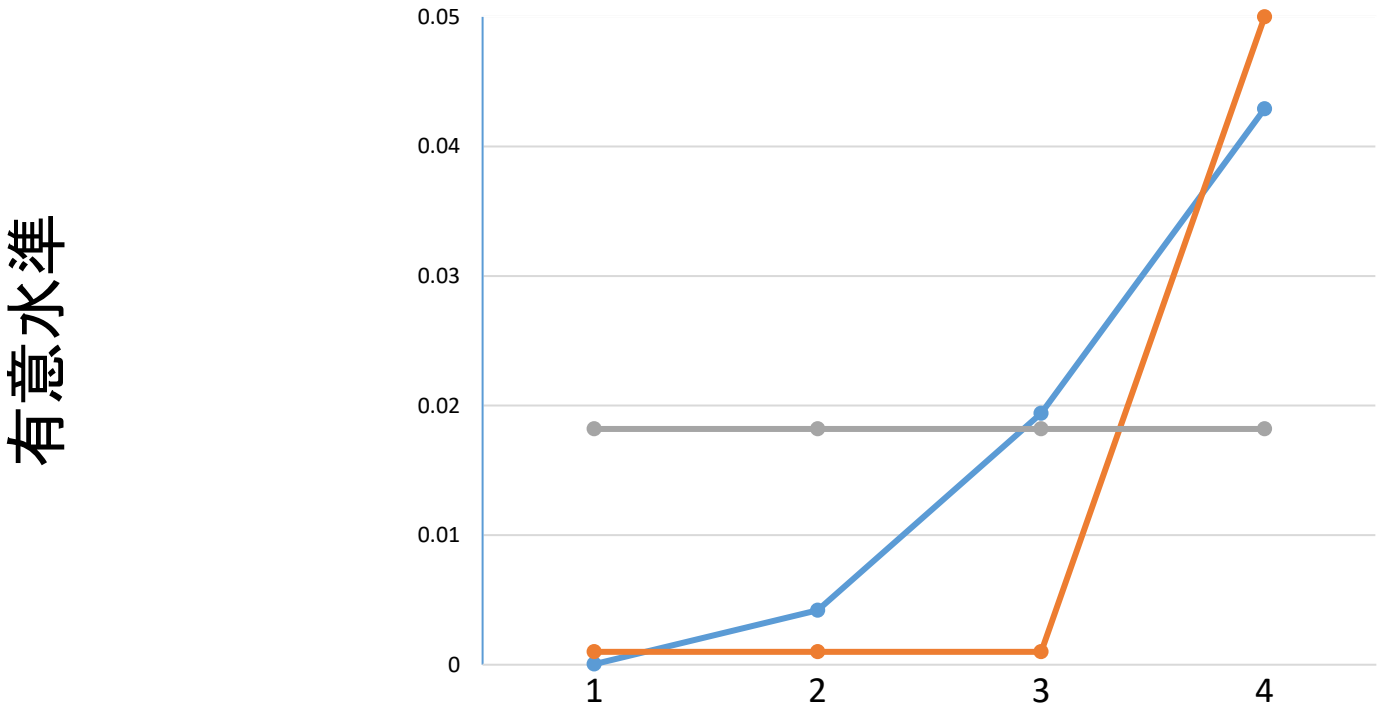
どちらか1本が不良品の確率 ＝ 1 －（2本ともまとも）
＝ 1 － 0.9025 ＝ 0.0975

第1種の過誤確率（中間解析のように各検定が独立でない場合を想定）

名目 有意水準	繰り返し検定の回数					
	1	2	3	4	5	10
1%	1.0	1.8	2.4	2.9	3.3	4.7
5%	5.0	8.3	10.7	12.6	14.2	19.3

Armitage P, *et al*. JRSS A. 1969;132:235-244.

④ グループ逐次法



オブライエン・フレミング法	0.00005	0.0042	0.0194	0.0429
ヘイビトル・ピト法	0.001	0.001	0.001	0.05
ポコック法	0.0182	0.0182	0.0182	0.0182

全体の有意水準を0.05、中間解析を3回（解析を計4回）行う計画の場合の3種類のグループ逐次法の各解析時点での調整された有意水準

⑤ サブグループ解析

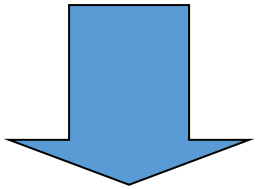
- 全体の集団のある一部のグループ（部分集団）を対象とした解析

注：「層別解析」は「stratified analysis」の訳であり、混乱するので使用しない方がよい

- 層別検定（Mantel-Haenszel検定）
- 層別ログランク検定
- 層別比例ハザードモデル

⑤ サブグループごとの検定の問題

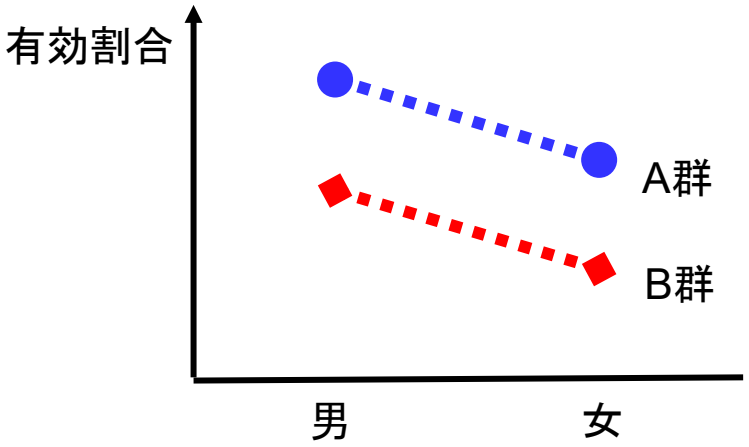
- 多数の検定の実施
  - 偽陽性の確率（αエラー）が上昇
- 各サブグループの大きさが小さい
  - 偽陰性の確率（βエラー）が上昇



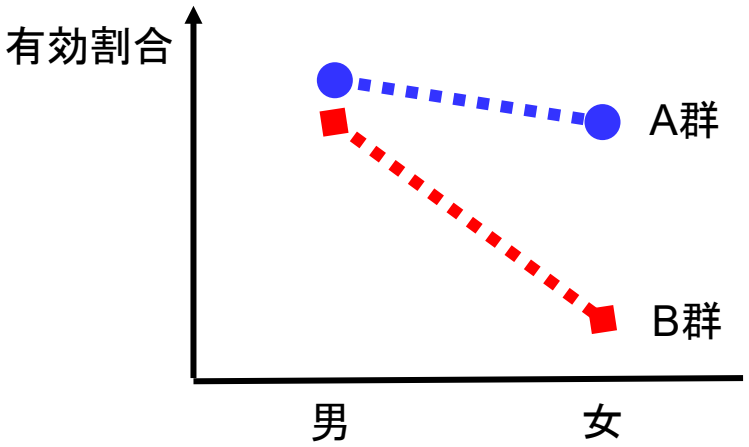
ほとんどの結果は誤り

⑤ 交互作用

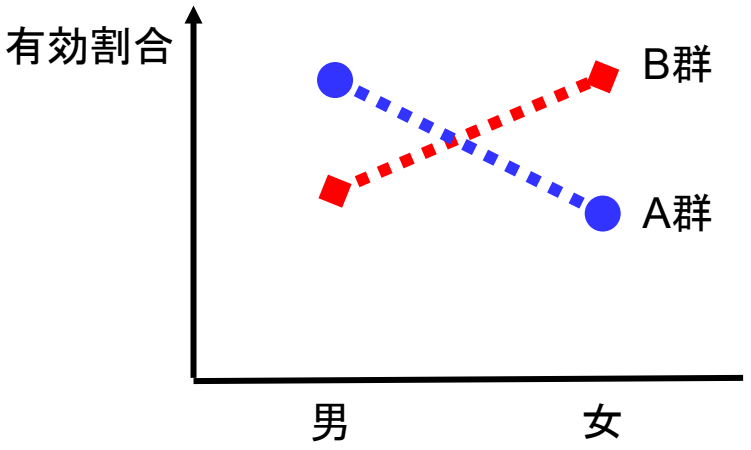
交互作用なし：各サブグループにおける効果の大きさが等しい



量的交互作用あり：
各サブグループにおける効果の大きさは異なるが効果の方向は同じ



質的交互作用あり：
各サブグループにおける効果の大きさは異なり、効果の方向も異なる



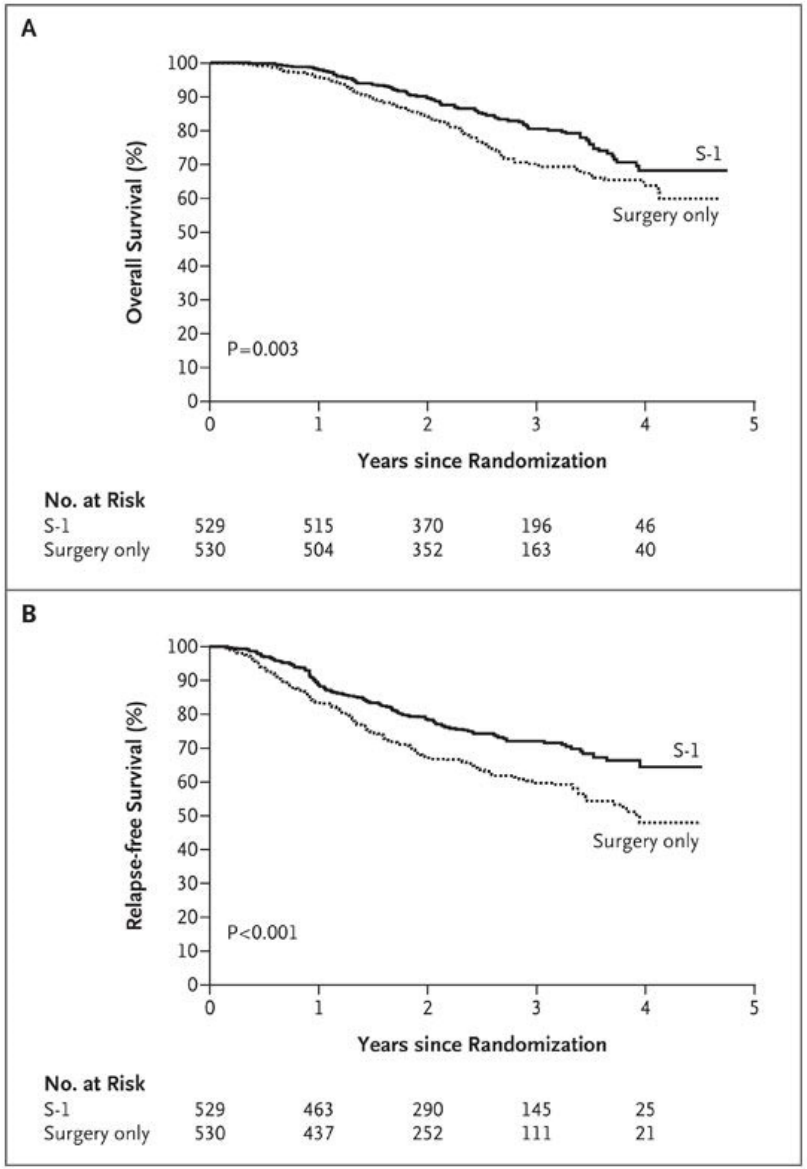
⑤ サブグループ解析の方法

- 計画時
  - 関心のあるサブグループを事前に特定（試験実施計画書/統計解析計画書に記載）しておく

- 解析時
  - 交互作用の検定*を行い、それが有意な場合のみ各サブグループでの検定結果を解釈する

* 帰無仮説：すべてのサブグループ間で効果の大きさが等しい

サブグループ解析の事例（胃癌臨床試験）



全生存期間

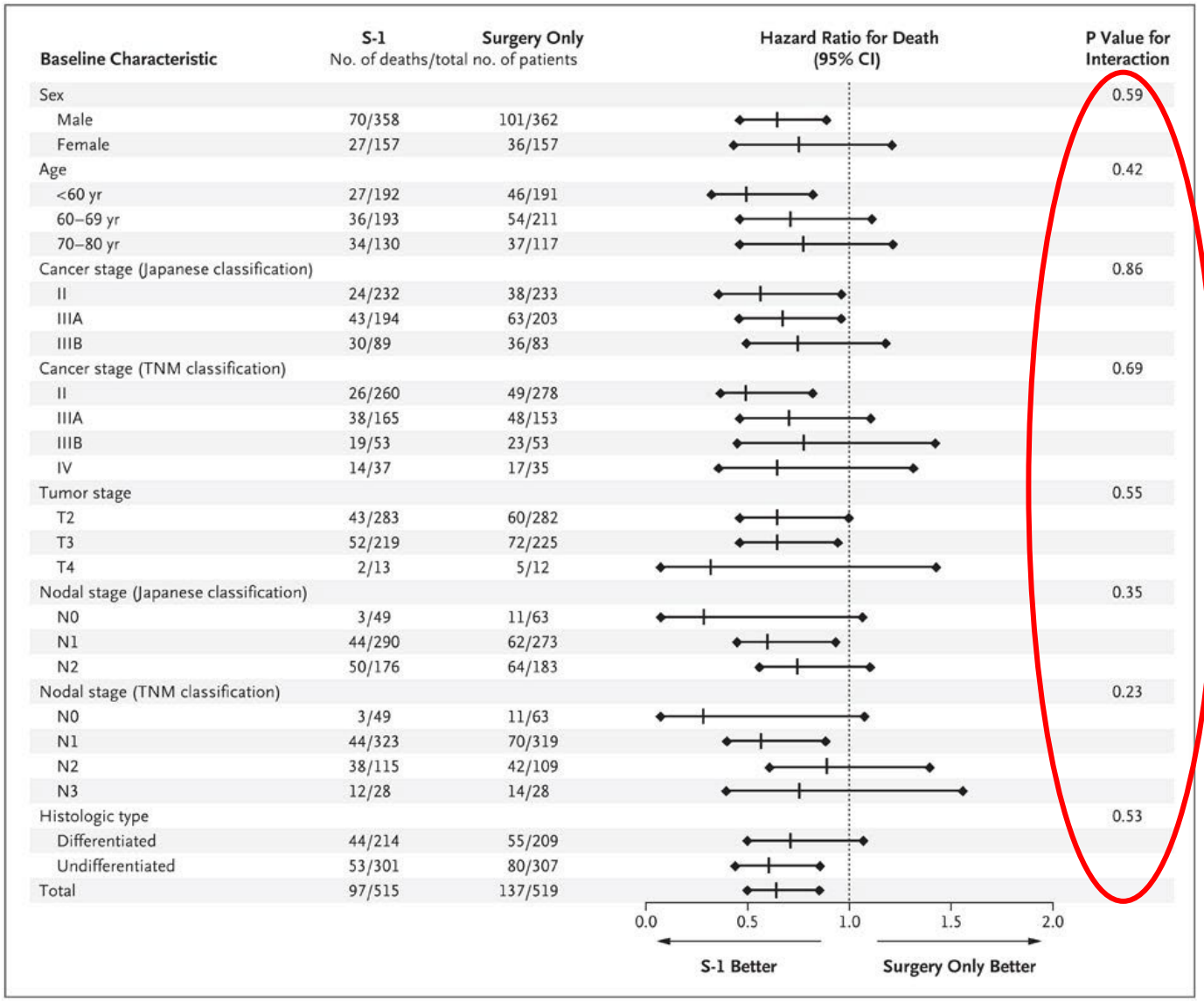
ハザード比　0.68
(95%信頼区間　0.52～0.87)
P=0.003

無再発生存期間

ハザード比　0.62
(95%信頼区間　0.50～0.77)
P<0.001

Sakuramoto S, et al. NEJM 2007;357:1810-20. (ACTS-GC)

サブグループ解析：フォレストプロット



Baseline Characteristic	S-1 No. of deaths/total	Surgery Only no. of patients	P Value for Interaction
Sex			0.59
Male	70/358	101/362	
Female	27/157	36/157	
Age			0.42
<60 yr	27/192	46/191	
60–69 yr	36/193	54/211	
70–80 yr	34/130	37/117	
Cancer stage (Japanese classification)			0.86
II	24/232	38/233	
IIIA	43/194	63/203	
IIIB	30/89	36/83	
Cancer stage (TNM classification)			0.69
II	26/260	49/278	
IIIA	38/165	48/153	
IIIB	19/53	23/53	
IV	14/37	17/35	
Tumor stage			0.55
T2	43/283	60/282	
T3	52/219	72/225	
T4	2/13	5/12	
Nodal stage (Japanese classification)			0.35
N0	3/49	11/63	
N1	44/290	62/273	
N2	50/176	64/183	
Nodal stage (TNM classification)			0.23
N0	3/49	11/63	
N1	44/323	70/319	
N2	38/115	42/109	
N3	12/28	14/28	
Histologic type			0.53
Differentiated	44/214	55/209	
Undifferentiated	53/301	80/307	
Total	97/515	137/519	

交互作用の検定はすべて有意ではない
（効果はすべてのサブグループで一様である）

Sakuramoto S, et al. NEJM 2007;357:1810-20. (ACTS-GC)

おわりに

- 結果の質を保証するためには、科学的に妥当な試験デザインに基づいた試験実施計画書を作成しなければならない
- 試験実施計画書は、専門領域外の者が審査および参照することがあるため、系統立った分かりやすい記述が求められる
- 試験実施計画書の記載事項は多岐にわたるため、様々な分野の専門家（医師、看護師、薬剤師、プロジェクトマネジャー、試験統計家、データマネジャー、臨床試験コーディネーターなど）が協同して作成する必要がある

参考図書

- Guosheng Yin. 手良向聡、大門貴志訳. 臨床試験デザイン. メディカル・パブリケーションズ. 2014.
- 手良向聡. なぜベイズを使わないのか!?ー臨床試験デザインのために. 金芳堂. 2017.
- 手良向聡. 臨床試験におけるランダム化の意義と限界. 計量生物学 2020;41:37-54.

無料ダウンロード可
（https://www.jstage.jst.go.jp/article/jjb/41/1/41_37/_article/-char/ja）