

がんを対象としたPhase II試験のデザイン

国立がん研究センター
JCOGデータセンター 統計部門
町田龍之介

下記のスライドを参考に作成させて頂きました

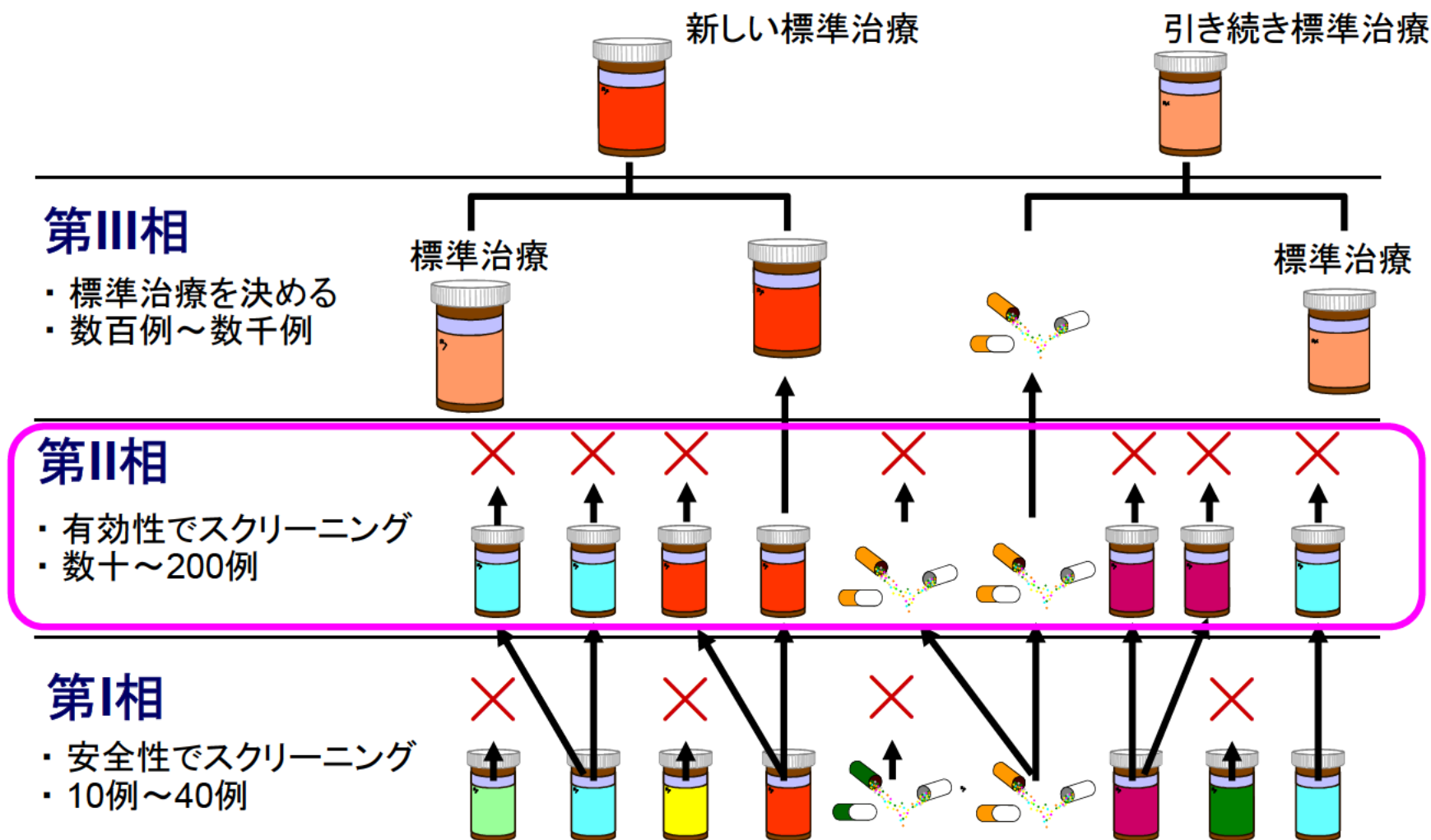
- ・水澤先生 2020.10.10 : JCOG臨床試験セミナー入門編
- ・片山先生 2013.8.10 片山先生: JCOG臨床試験セミナー中級編
- ・福田先生: 統計的考え方, 臨床試験その4_PhaseII, 特殊なPhaseII
分子標的治療の試験デザイン_BiomarkerDrivenTrialDesign

2021年10月2日 JCOG臨床試験セミナー中級編

本日のトピックス

- 第II相試験デザインを理解し、状況に応じて適切なデザインを選択できるようになること
- 単群の第II相試験
 - 閾値と期待値に基づく典型的デザイン
 - ベイズ流に基づくデザイン
- ランダム化第II相試験
 - ランダム化スクリーニングデザイン
 - ランダム化セレクションデザイン
- マスタープロトコルによる試験
 - バスケット試験, アンブレラ試験, プラットフォーム試験

がん臨床試験の開発の流れ(相)



臨床試験の原則

臨床試験は“比較”である

第III相

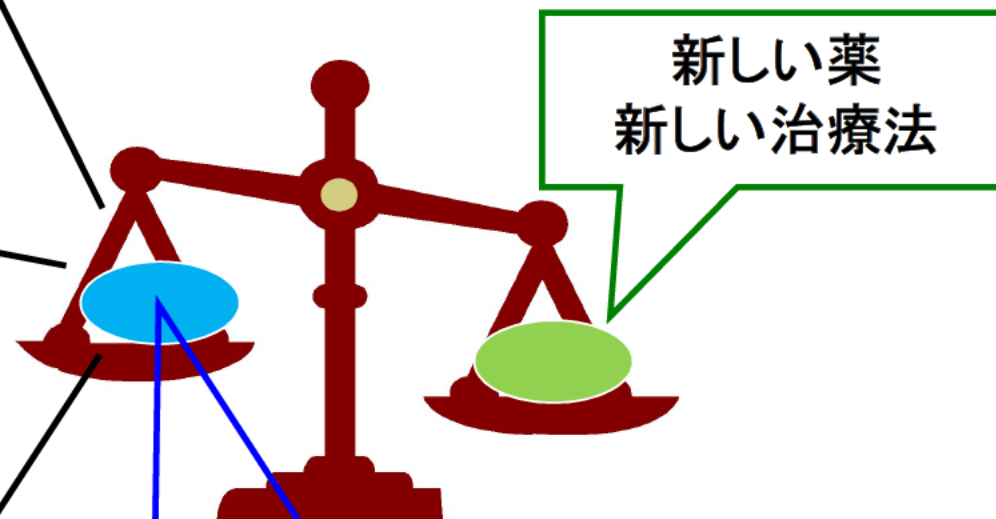
標準薬・標準治療と
総合的な有効性と安全性を
直接同時比較

第II相

過去のデータから導かれる
有効性の水準や目標と比較

第I相

過去のデータから導かれる
安全性の許容水準と比較
(許容される毒性の割合)



新しい薬
新しい治療法

比較する相手は常に
標準薬・標準治療

第II相試験の典型的な状況

- 有効性でスクリーニングしたい
 - 試験薬・試験治療: **細胞傷害性薬剤**
 - RECISTによる腫瘍縮小効果で評価(奏効割合など)
 - ✓ 第III相試験でのエンドポイント(予後)の代替
 - ✓ 客観的なエンドポイントで短期間で評価できる
- ヒストリカルコントロールが存在している
 - 従来は細胞傷害性薬剤が主流で, 過去の標準治療の腫瘍縮小効果のデータを比較相手として使用できる

第II相試験の典型的デザイン

□ 閾値と期待値を用いた単群デザイン

- 対象:がん種を限定
- 登録数:20～60例
- Primary endpoint: 奏効割合など短期的なエンドポイント

□ 閾値

- ✓その後の開発をあきらめる値
- ✓標準治療とのリスクベネフィットバランスに基づき設定

□ 期待値

- ✓ぜひ開発を続けたい値(第III相へ進める値)
- ✓臨床的に意味のある値, 期待される値に基づき設定

閾値の設定は？

- 条件設定：単群の第II相試験
 - Primary endpoint: 奏効割合
 - 標準治療の奏効割合: 20%
 - 試験治療は標準治療よりも毒性が強くなる
 - ✓ 標準治療(治療薬A), 試験治療(治療薬A+治療薬B)

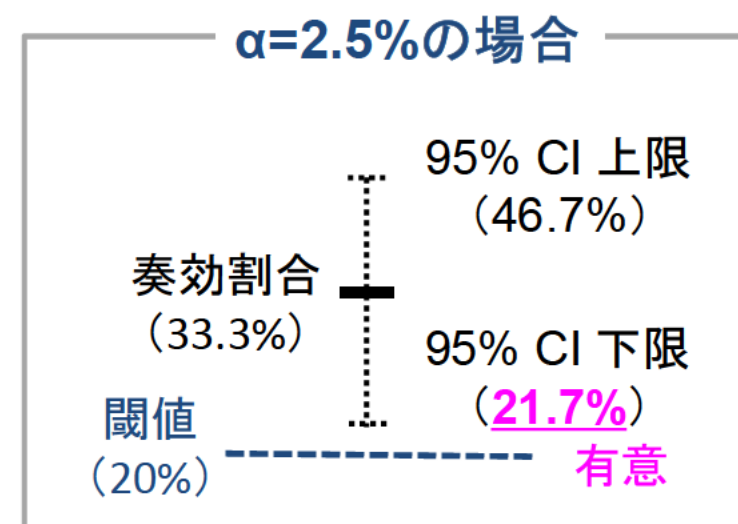
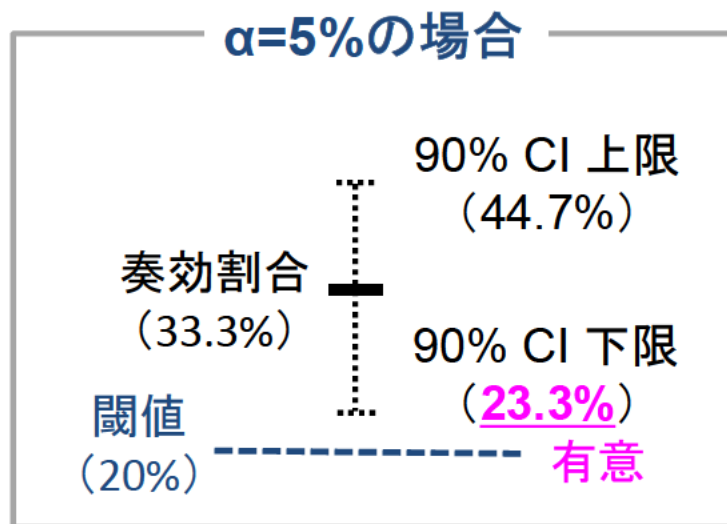
- 上記の試験の場合, どの閾値が妥当でしょうか？
 - 閾値 = 15%
 - 閾値 = 20%
 - 閾値 = 25%

期待値の設定は？

- 基本的には「そのくらいの値が得られるのであれば、ぜひ開発を続けたい」と考える値
- 統計的な観点から妥当かの判断はできない
 - グループの先生方が臨床的な観点から決定
- 考慮すべきは以下の2点
 - 標準治療と比較した**毒性とのバランス**（臨床的意味）
 - 本当にそれだけの**有効性が期待できるか**

有意かどうかの判断規準

- P値が有意水準(α)を下回れば「有意」ですが、、、
P値がなくても奏効割合の信頼区間から判断できる
- 信頼区間(CI)の下限が閾値を上回れば有意
 - 閾値20%の設定のもとで60人中20人(33.3%)が奏効



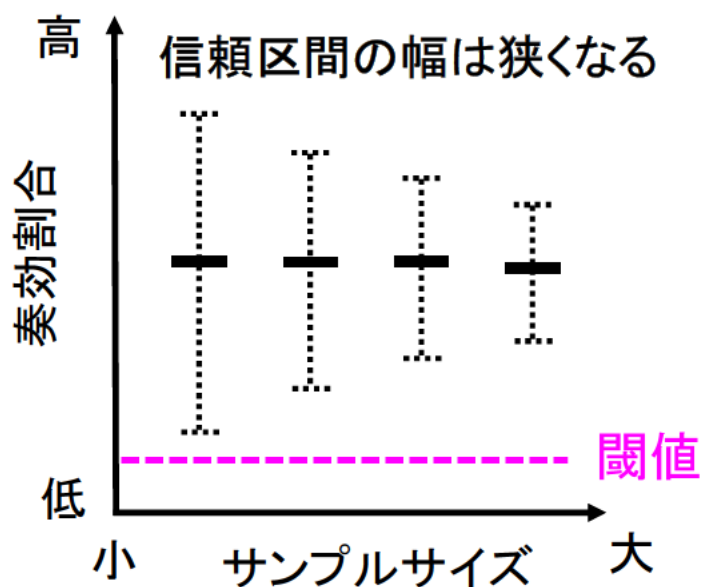
有意かどうかの判断では閾値が重要

閾値とサンプルサイズの関係

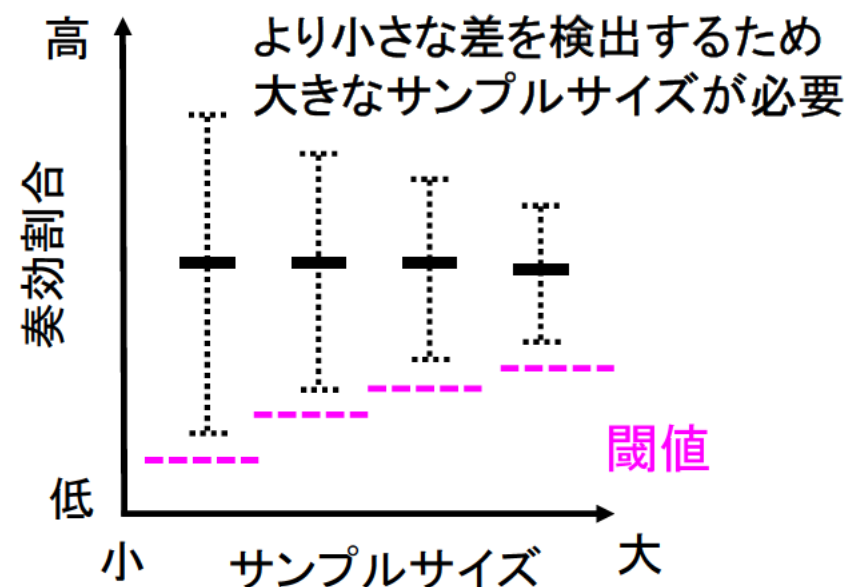
(イメージ)

- サンプルサイズ設計に必要な要因
 - 閾値, 期待値, 有意水準, 検出力の4つ

① サンプルサイズを大きくすると



② 閾値を高く設定すると

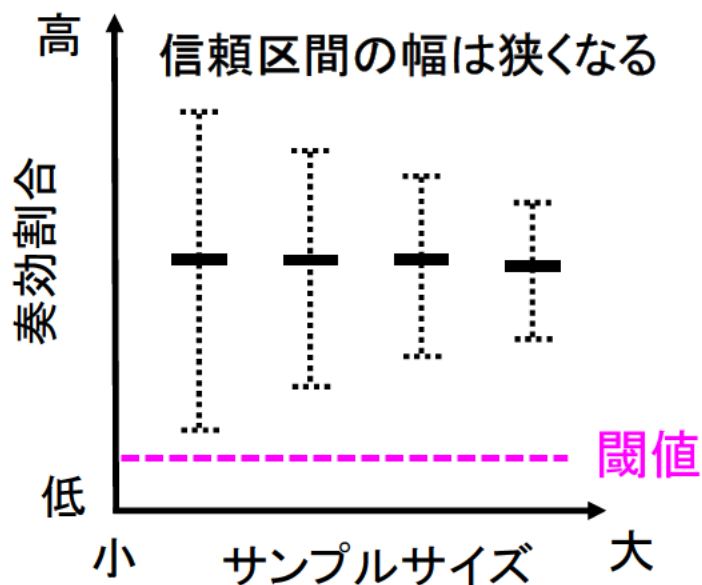


期待値とサンプルサイズの関係

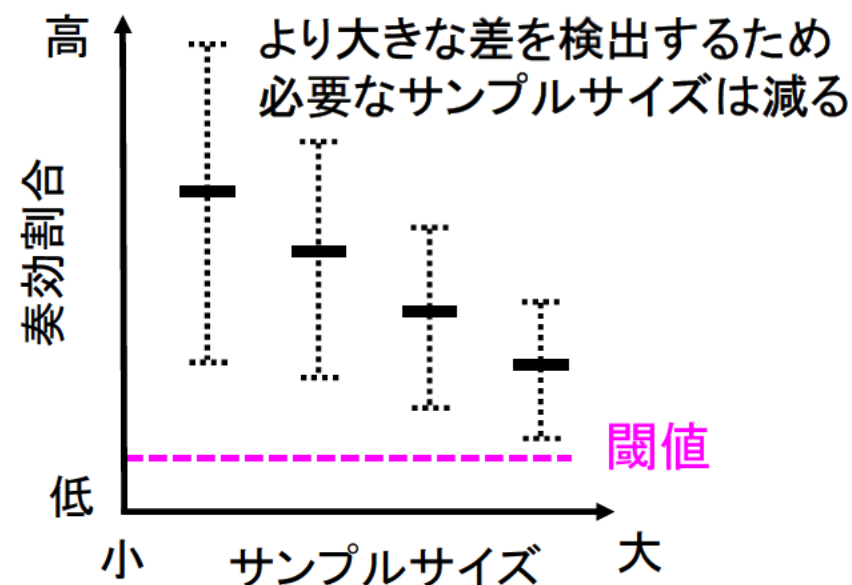
(イメージ)

- サンプルサイズ設計に必要な要因
 - 閾値, 期待値, 有意水準, 検出力の4つ

① サンプルサイズを大きくすると

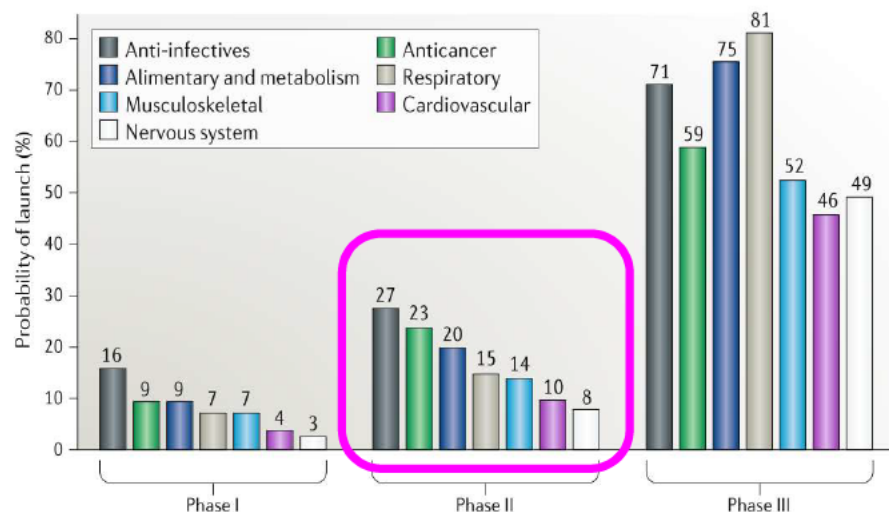


③ 期待値を高く設定すると



1段階デザイン or 2段階デザイン

- 1段階デザイン(1 stage design):
 - 途中で登録停止や解析することなく、目標被験者数まで一度に登録
 - 最後に検定して、有望かどうかの判断を行う
- がんの第II相試験で真に有効なものは20%程度



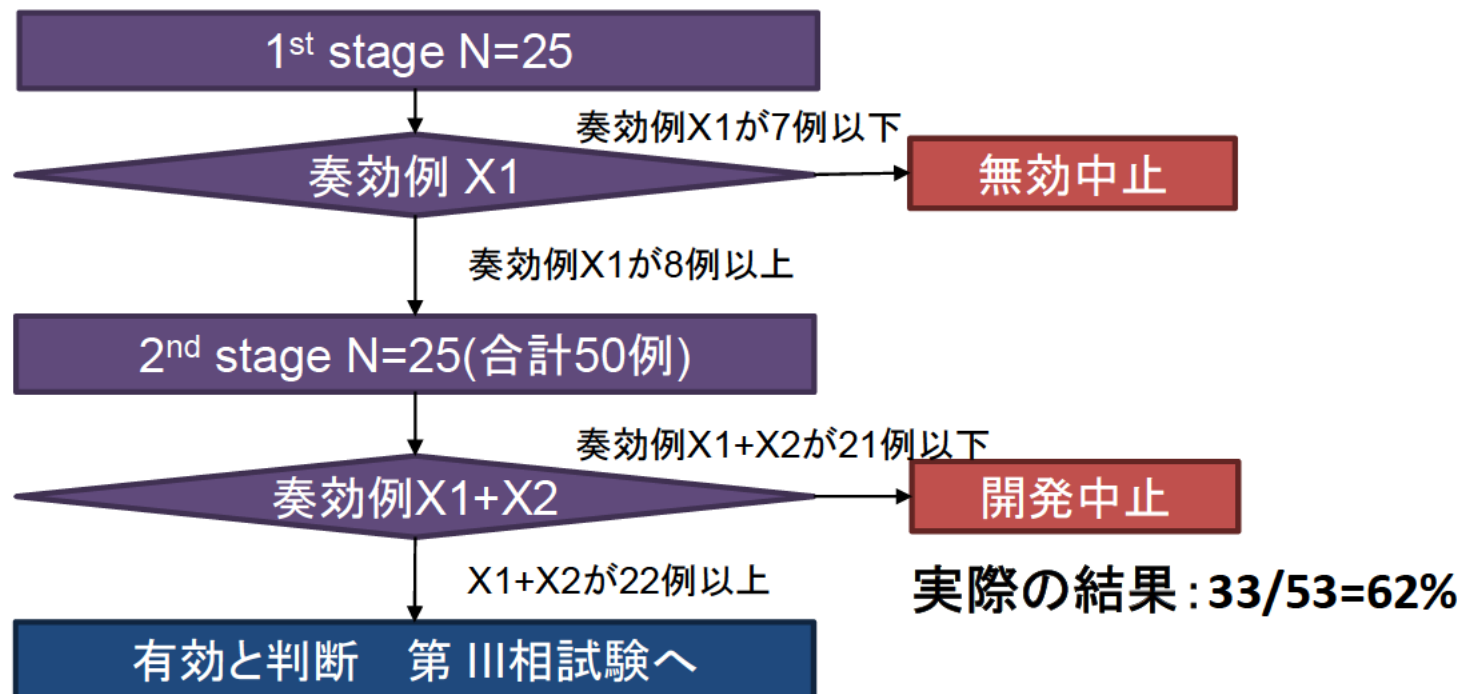
- 無効なものが多ければ被験者に対する倫理的配慮が必要
- 試験途中で登録停止・解析する2段階デザインが提案されている

Fig. 2 | Probability of launch from phases I, II and III by therapeutic area from 2010-2017.

Dowden H, Munro J: Nat Rev Drug Discov 18:495-496, 2019

2段階デザインの例 (JCOG0807)

- ❑ 切除不能・再発食道がん患者を対象とした第I/II相試験
- ❑ 第II相部分: 標準治療であるCisplatin+5-FU 併用療法 (CF療法) にDocetaxelを加えたDCF療法の有効性の評価
 - ❑ 閾値35%, 期待値50%, $\alpha=10\%$, 検出力=80%



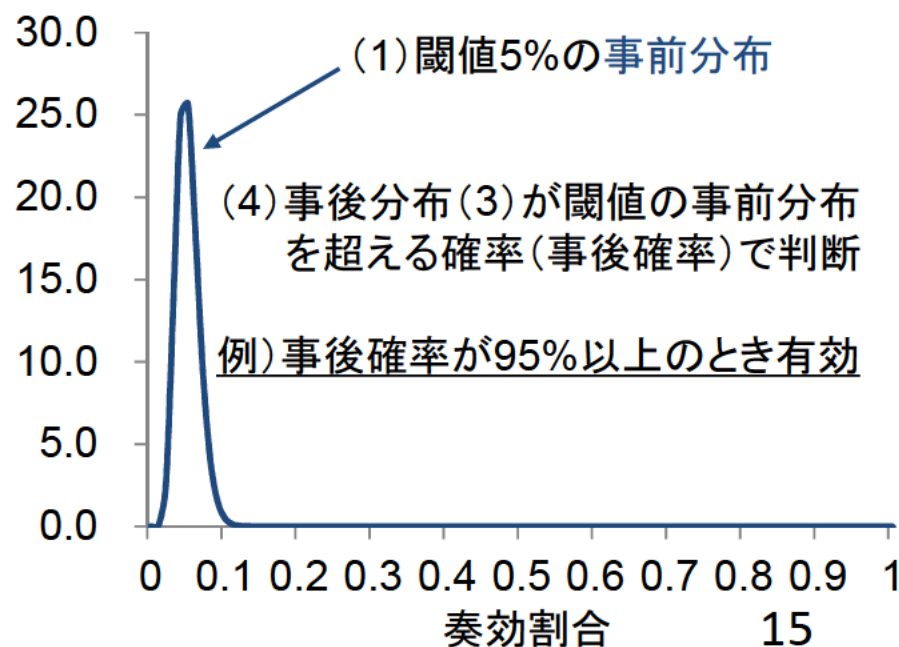
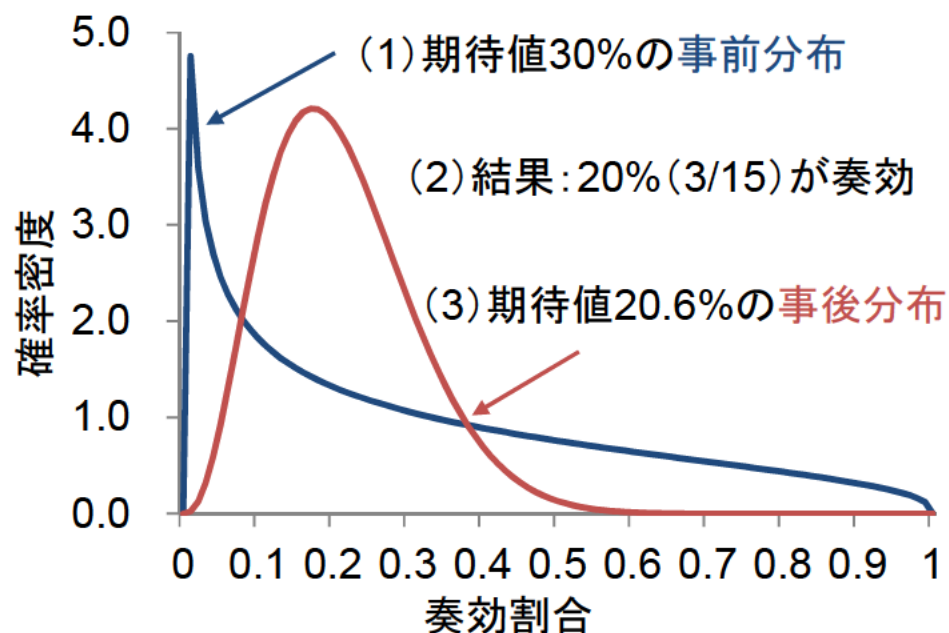
Hironaka S, et al. Cancer science. 2014;105(9):1189-95. ¹³

希少がん領域における臨床試験

- 典型的な単群デザイン
 - 目標被験者数に達した時点で意志決定（頻度流）
- 対象となる患者が少なく実現可能性がない状況
 - 試験途中で（1例ずつの）結果をモニタリングしながら柔軟な意志決定したい
 - 蓄積された情報（[事前情報](#)）を十分に活用したい
- ベイズ流のデザイン
 - 予測確率に基づく方法
 - 事後確率に基づく方法

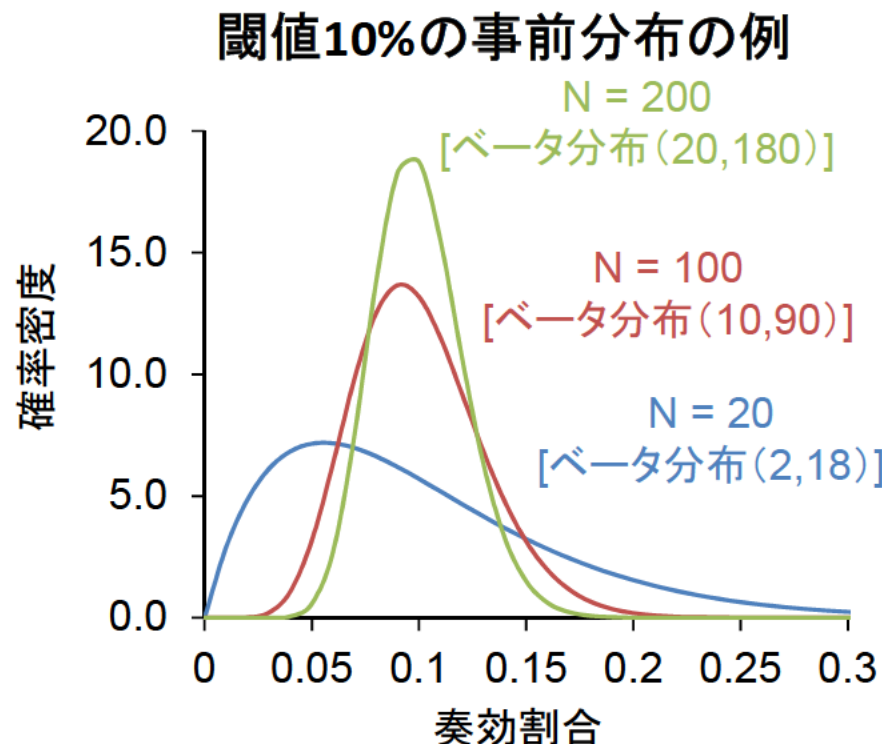
ベイズ流のアプローチ

- (1) 曖昧さがある閾値と期待値を設定 (事前分布)
- (2) 試験を実施して、連続的に試験治療のデータを得る
- (3) 試験データをもとに事前分布を更新 (事後分布)
- (4) 奏効割合が閾値よりも大きくなる (事後) 確率で判断



閾値・期待値の事前分布の設定

- 基本的な考え方は典型的な単群デザインと同じ
- 加えて閾値・期待値に関するデータについて
どのくらいの人数のデータが蓄積されているかの情報も必要



- 事前のデータ数を多く設定した場合
 - ✓ 事前分布の幅は小さくなる
 - ✓ 少ない被験者数で判断できる

⇒ α エラーの確率が極端に高くなっていないかの確認が必要(頻度論)

- 恣意性を排除することは困難
 - ✓ 解析段階でパラメータを変えた感度解析等も必要

ベイズ流デザインの例 (Thall and Simon, 1994)

□ NCCH1615 (STATICE試験)

- 対象: 根治切除不能かつ化学療法歴を有する子宮癌肉腫
- Primary endpoint: 奏効割合
- 閾値5%: 事前分布のベータ分布(10, 190)
- 期待値30%: 事前分布のベータ分布(0.6, 1.4)

□ 各対象者数で必要奏効例を超える場合に有効と判断

- 閾値奏効割合5%を超える事後確率が95%以上

解析対象数	15	16	17	18	19	20	21	22	23
必要奏効例数	3	3	3	3	3	3	3	4	4

Primary endpointは腫瘍縮小効果でよい？

□ 切除不能膵臓癌

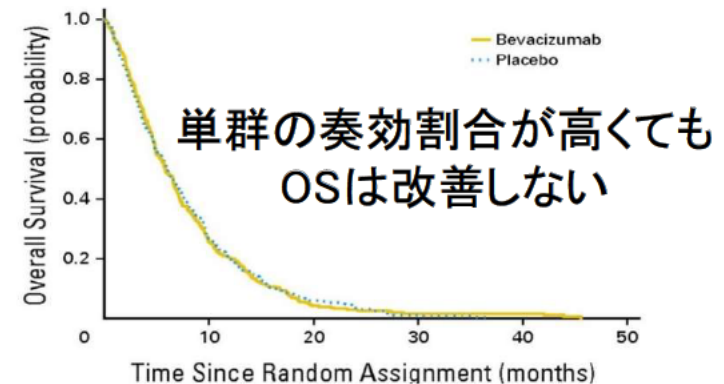
□ 単群phase II試験の結果

□ 奏効割合

- ベバシズマブ併用: 21%
- 当時の標準治療である
ゲムシタビン単剤: 10%未満

Kindler, Hedy L., et al. *JCO* 2005;23

第III相へ



Kindler, H. L., et al. *JCO* 2010; 28: 3617-3622

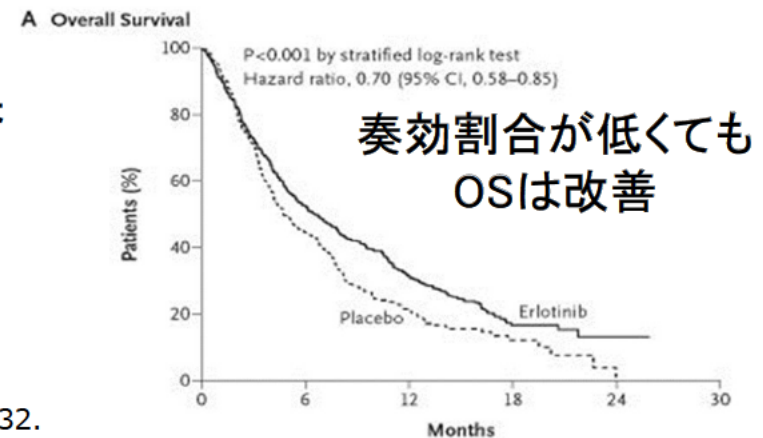
□ 既治療IIIB/IV期 非小細胞肺癌

□ ランダム化比較phase III試験の結果

□ 奏効割合

- プラセボ群: 0.7%
- エルロチニブ群: 8.2%

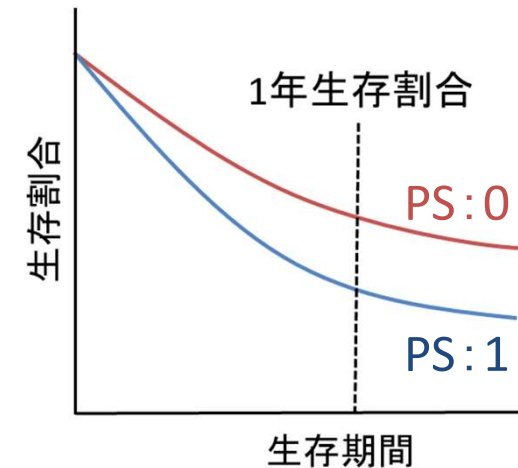
Shepherd FA et al. *N Engl J Med* 2005;353:123-132.



単群試験で年次(無増悪)生存割合は？

- 年次生存割合(例:1年生存割合)

- 登録される患者の全身状態によりバラツキが生じやすい
- 後治療の影響を受けやすい

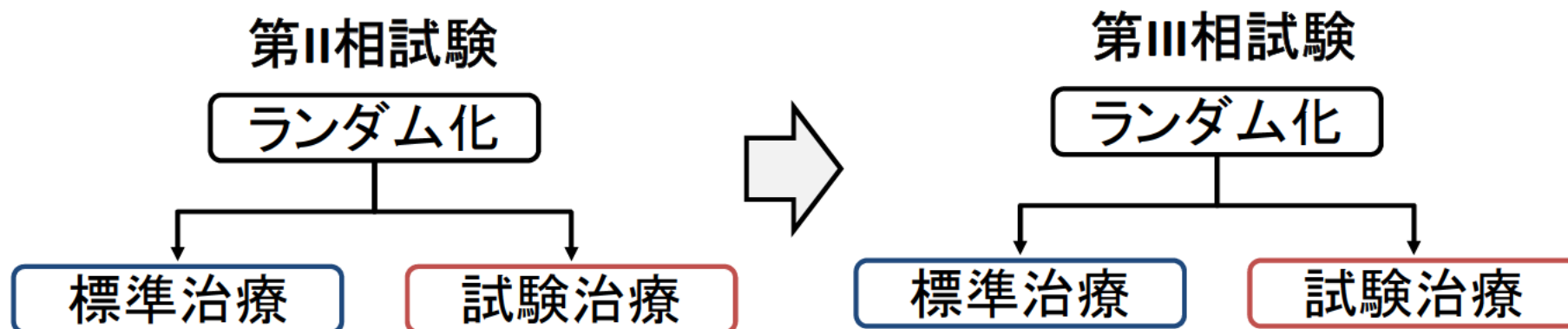


- 年次無増悪生存割合(例:6か月無増悪生存割合)

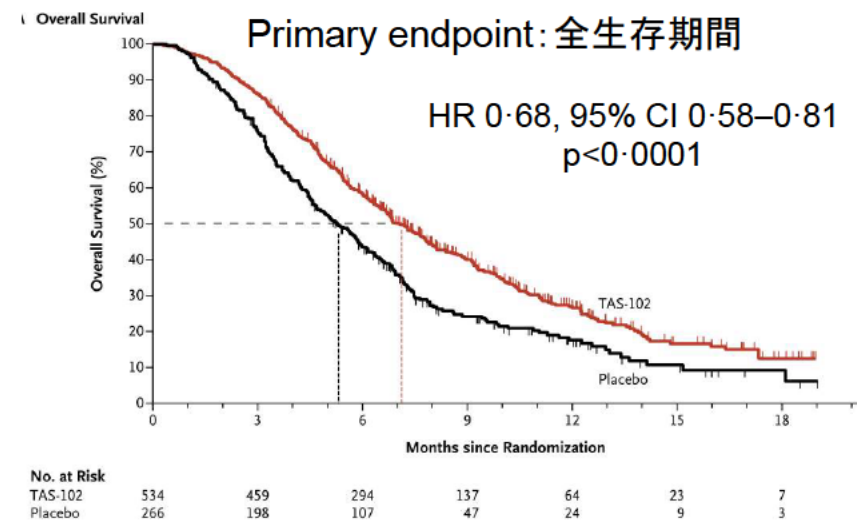
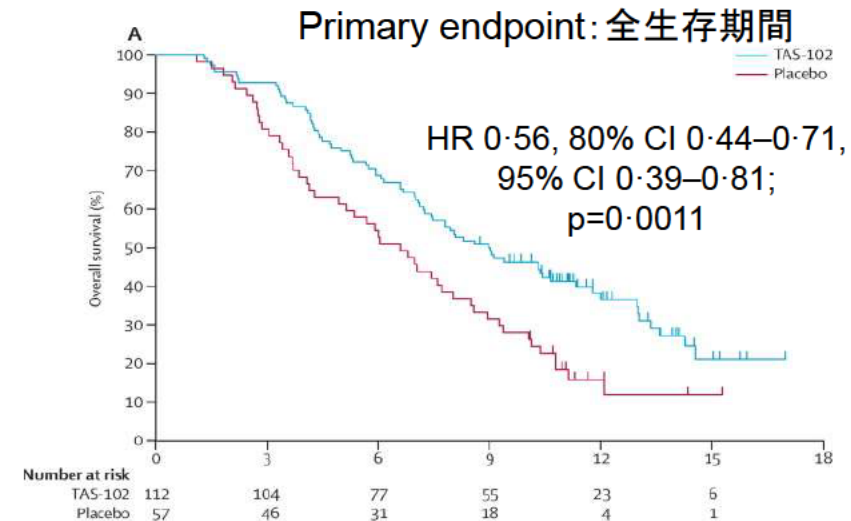
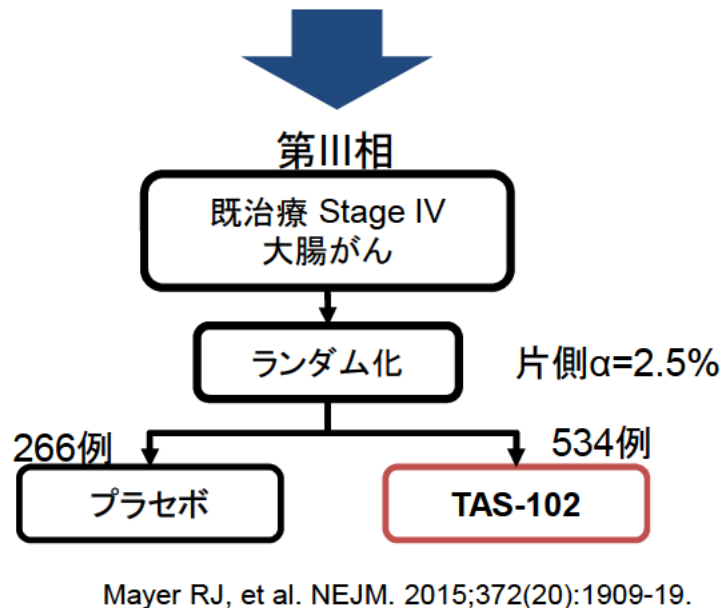
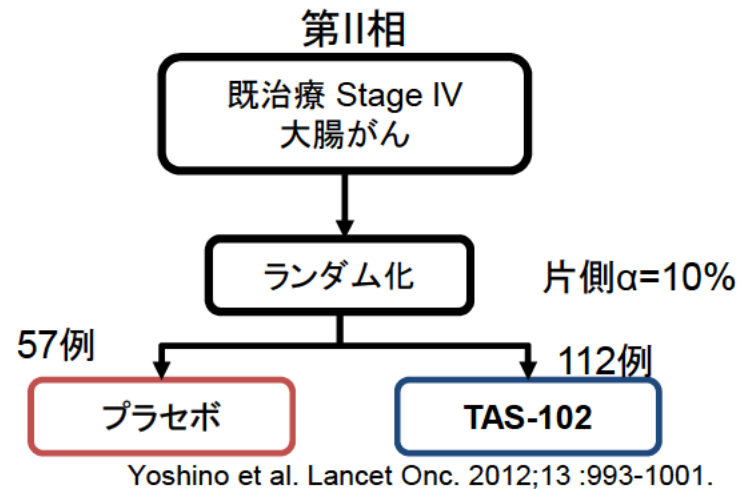
- 腫瘍縮小効果は検査間隔の影響を受けやすい
- 疾患によっては、増悪の定義が試験やグループで異なる

適切なヒストリカルコントロールがない場合

- 第III相での標準治療と試験治療をランダム化
⇒ ランダム化スクリーニングデザイン
 - サンプルサイズを抑えるため有意水準を少し高く設定
 - 有意になった場合でも第III相試験は必須
 - ✓ あくまで第III相試験の予備試験という位置づけ

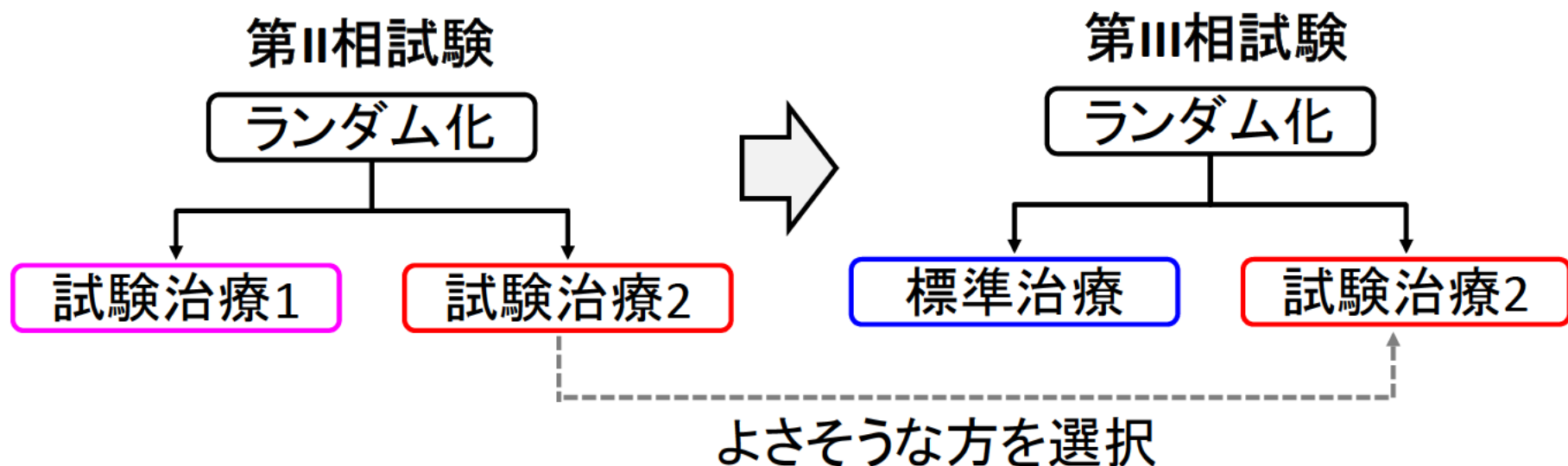


ランダム化スクリーニングデザインの例



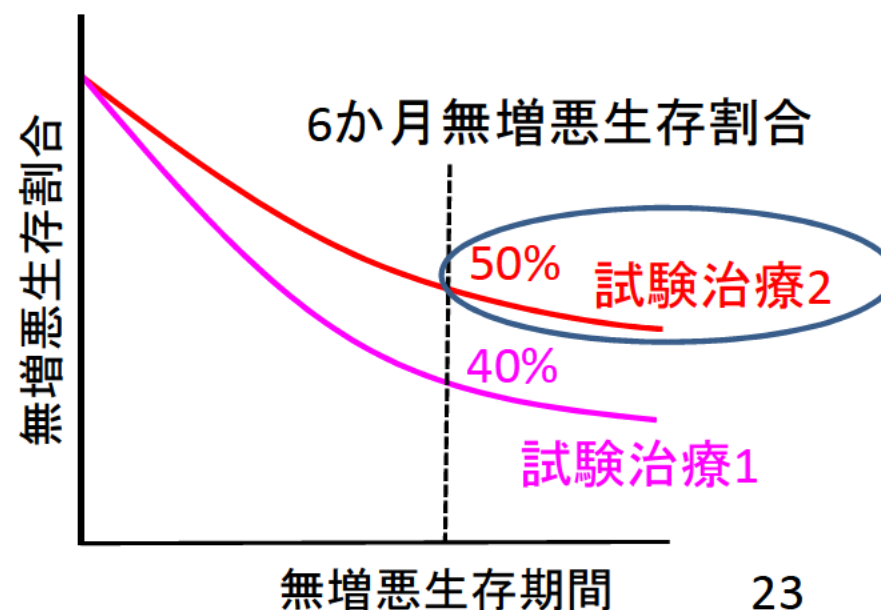
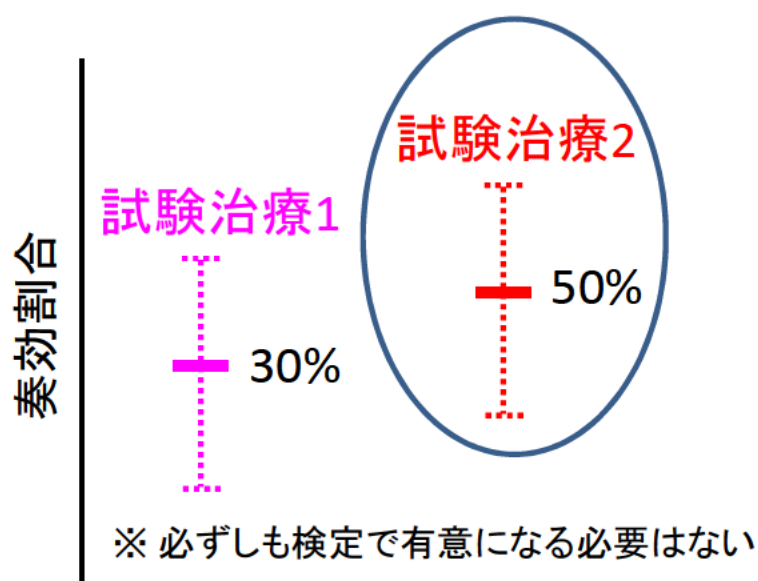
第III相の候補レジメンが複数ある場合

- いくつかある新しいレジメンの中から次の相に進めるものを“選択”するため試験治療をランダム化
⇒ ランダム化セレクション(選択)デザイン
 - 複数の試験治療どうしの比較(3つ以上も可能)



どのように第III相へ進む治療を選択するのか

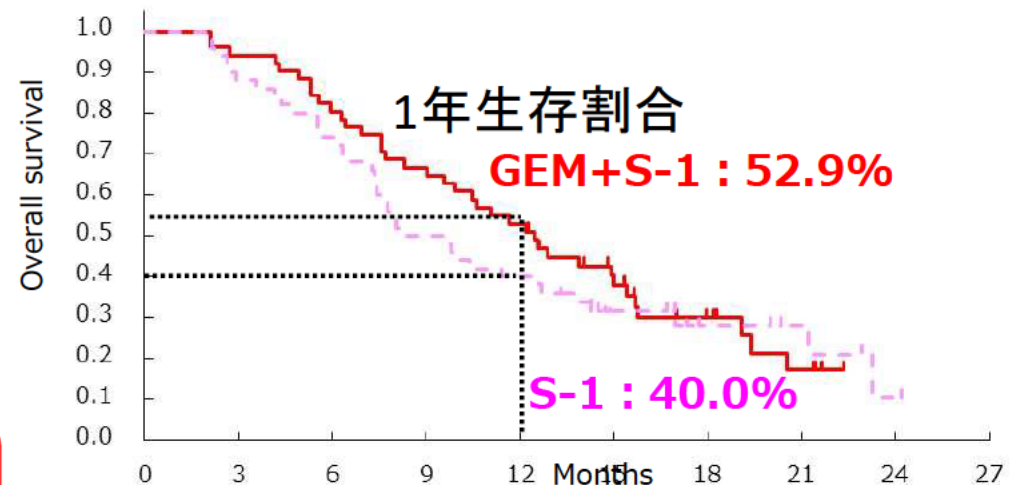
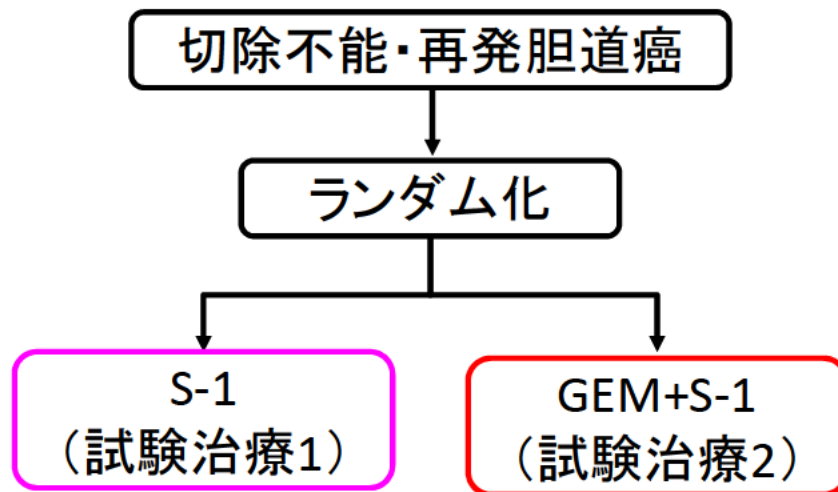
- Primary endpoint:
 - 奏効割合, 6か月無増悪生存割合, 1年生存割合など
- 判断規準:
 - 効果(点推定値)が少しでも高い方を選択する



ランダム化セレクションデザインの例

□ JCOG0805 (肝胆膵グループ)

- Primary endpoint: 1年生存割合
- 判断規準: GS療法とS-1療法のうち、1年生存割合の点推定値で上回るものをより有望な治療法とする



※ 1年生存割合が点推定値で優った**GEM+S-1療法**を標準治療の**GEM+CDDP療法**とのphase III 試験の試験治療に選択(JCOG1113として実施)

確認問題

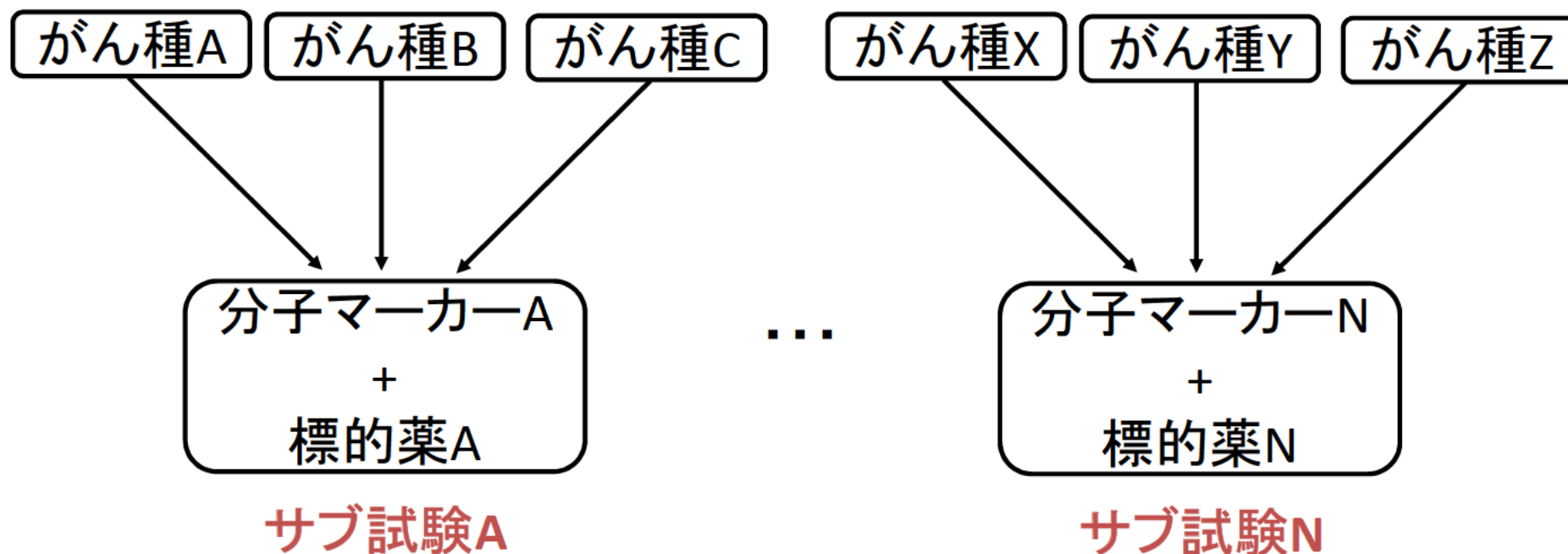
- スクリーニングデザインでは、治療効果が少しでも上回っている試験治療を次の相に進める
- 選択デザインでは、有意水準を少し大きめに設定し有意になれば次の相に進める

分子標的治療の治療開発

- 近年、次世代シーケンサーや遺伝子診断パネルの進展により、分子マーカーに着目した治療開発
- 従来のようにがん種、進行度、組織型等を定めてその各集団に対して治療開発を行う場合の課題
 - 試験間の作業重複が多く効率的でない
 - 希少な部分集団になり現実的でない
- 分子マーカーとその標的治療の組み合わせを評価するための効率的な試験デザインが提案されている

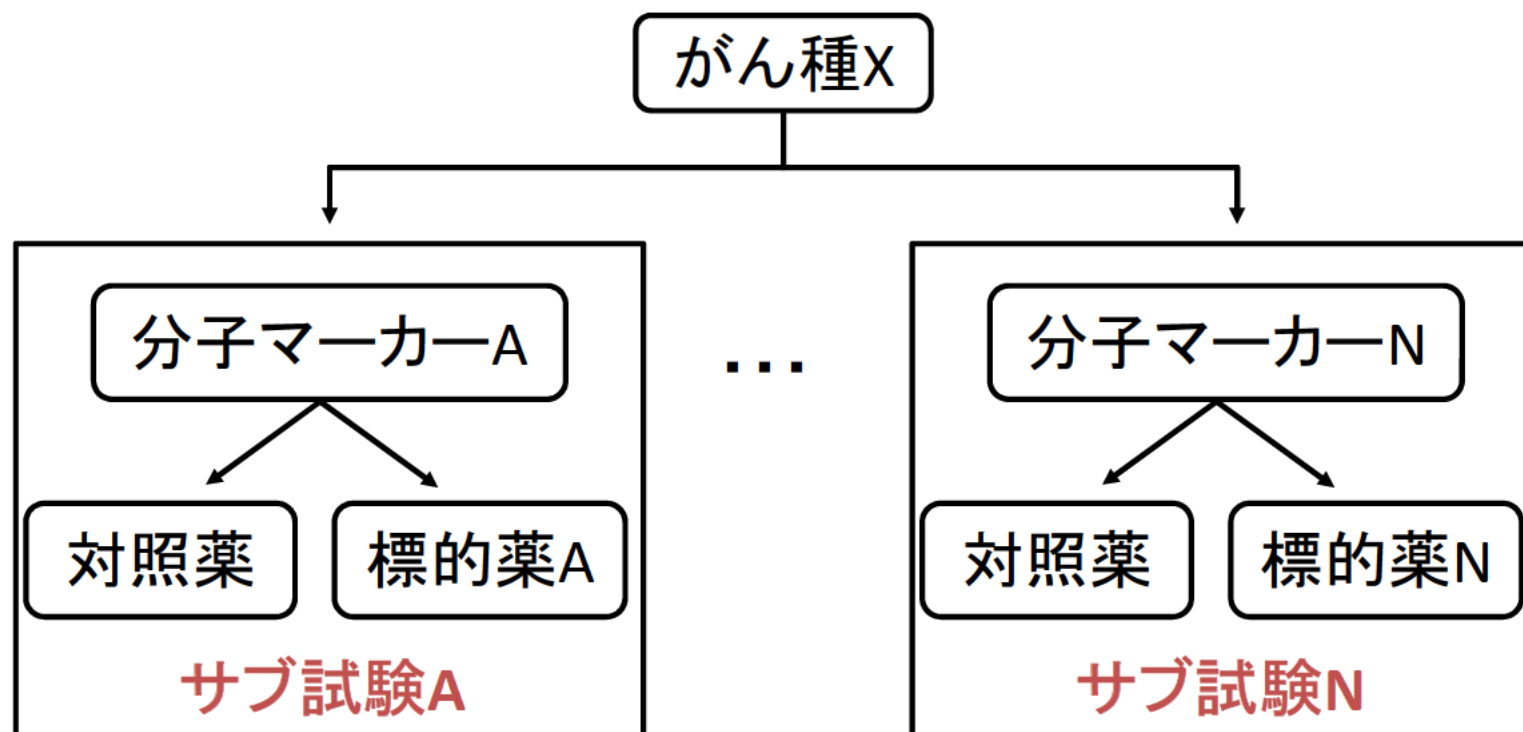
複数がん種/単一薬剤の場合

- 複数のがん種に対して
単一の標的治療を評価したい⇒バスケット試験
- ある薬剤がどのがん種に効くか



単一がん種/複数薬剤の場合

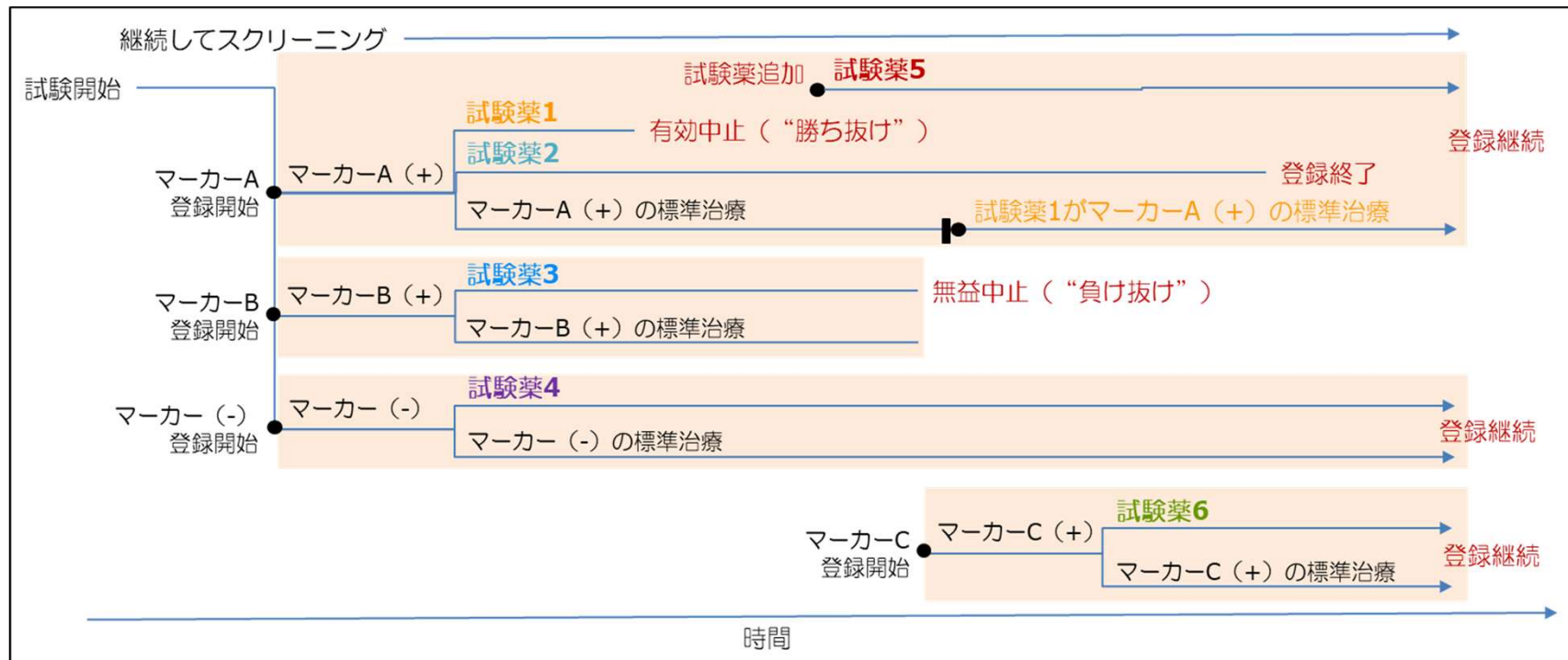
- 単一のがん種に対して
複数の標的治療を評価したい⇒アンブレラ試験
- あるがん種にどの薬剤が効くか



治療薬の追加・除外を認める場合

- 永続的に単一がん種に対して複数の標的治療を評価（アンブレラ試験）
- さらに試験中に新たな治療法や対象患者の追加や除外を認めたい⇒プラットフォーム試験

（福田先生の資料より）



マスタープロトコルの活用

- 複数の仮説を評価することを目的に作成された包括的なプロトコル(マスタープロトコル)
 - 共通プロトコルでの管理による治療開発の迅速化
 - 広範囲の分子マーカーを対象とすることで、希少がん患者を効率的に登録可能
- バスケット試験とアンブレラ試験
 - 各サブ試験をマスタープロトコルの傘下として実施
- プラットフォーム試験
 - マスタープロトコル試験の一種

第II相試験デザインのまとめ(1)

- 主に有効性による候補薬のスクリーニング
- 適切なヒストリカルコントロールが存在する場合
 - 奏効割合などをprimary endpointとした単群試験
- 適切なヒストリカルコントロールがない場合
 - 標準治療と試験治療をランダム化したスクリーニングデザイン
- 試験治療の候補が複数の場合
 - 試験治療どうしをランダム化したセレクションデザイン
 - 次は、標準治療と比較する第III相試験が必要

第II相試験デザインのまとめ(2)

- マスタープロトコルによる試験
 - バスケット試験
 - ✓複数がん種/単一薬剤
 - ✓ある薬剤がどのがん種に効くか
 - アンブレラ試験
 - ✓単一がん種/複数薬剤
 - ✓あるがん種にどの薬剤が効くか
 - プラットフォーム試験
 - ✓単一のがん種/複数の標的・薬剤(追加・除外を許容)
 - ✓長期間にわたって継続