がん臨床試験の計画に必要な統計の基礎知識
～外部対照を用いた比較における注意点～

研究支援センター 生物統計部
/東病院医薬品開発推進部門
若林 将史

2022. 12.10 (土)

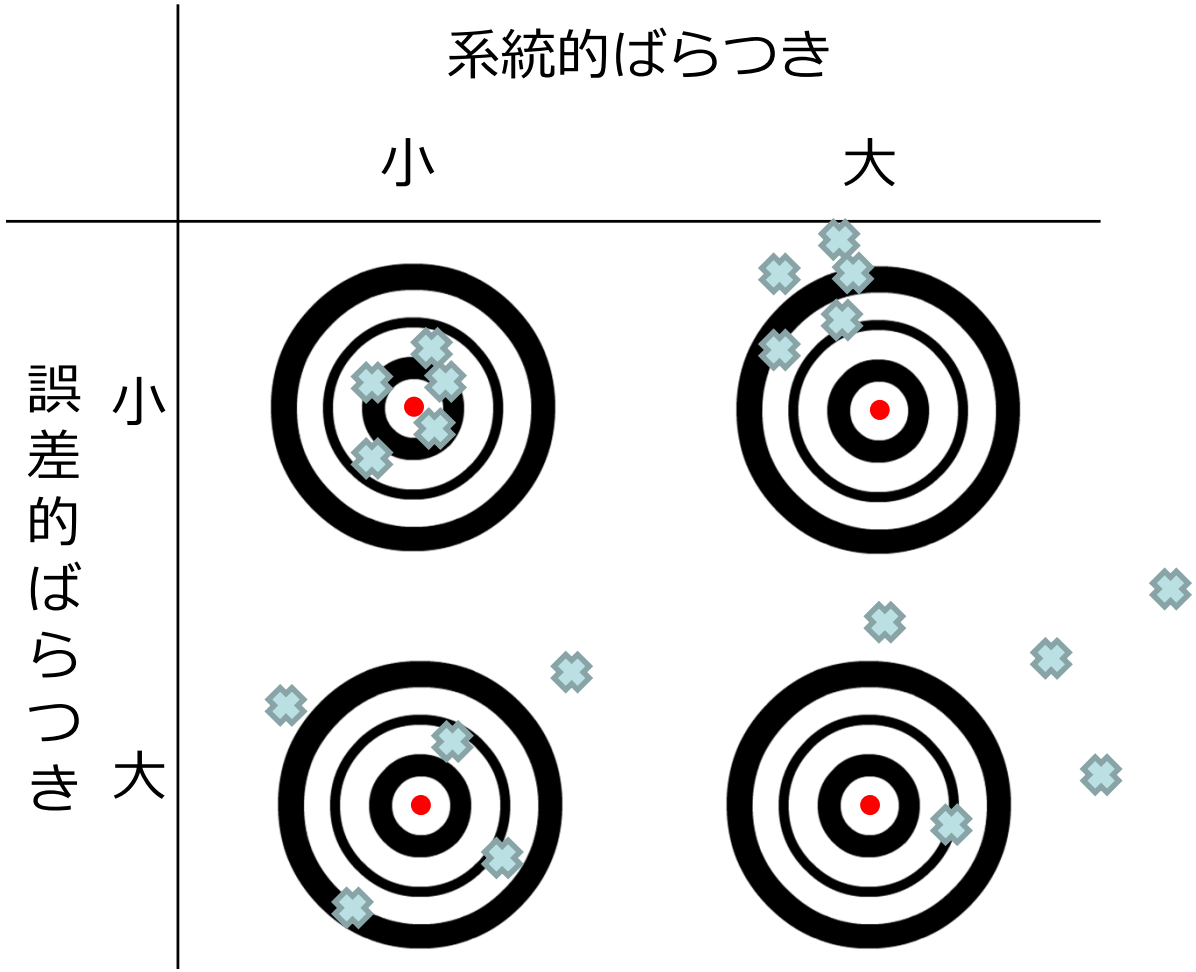
講義の目標

- ランダム化はなぜ必要か？
  - 交絡とランダム化
- 外部対照を用いた比較における問題点
  - データベースの選択
  - 種々のバイアス
- 解析段階における交絡への対処法
  - 多変量解析と傾向スコア

交絡とランダム化

2つのばらつき

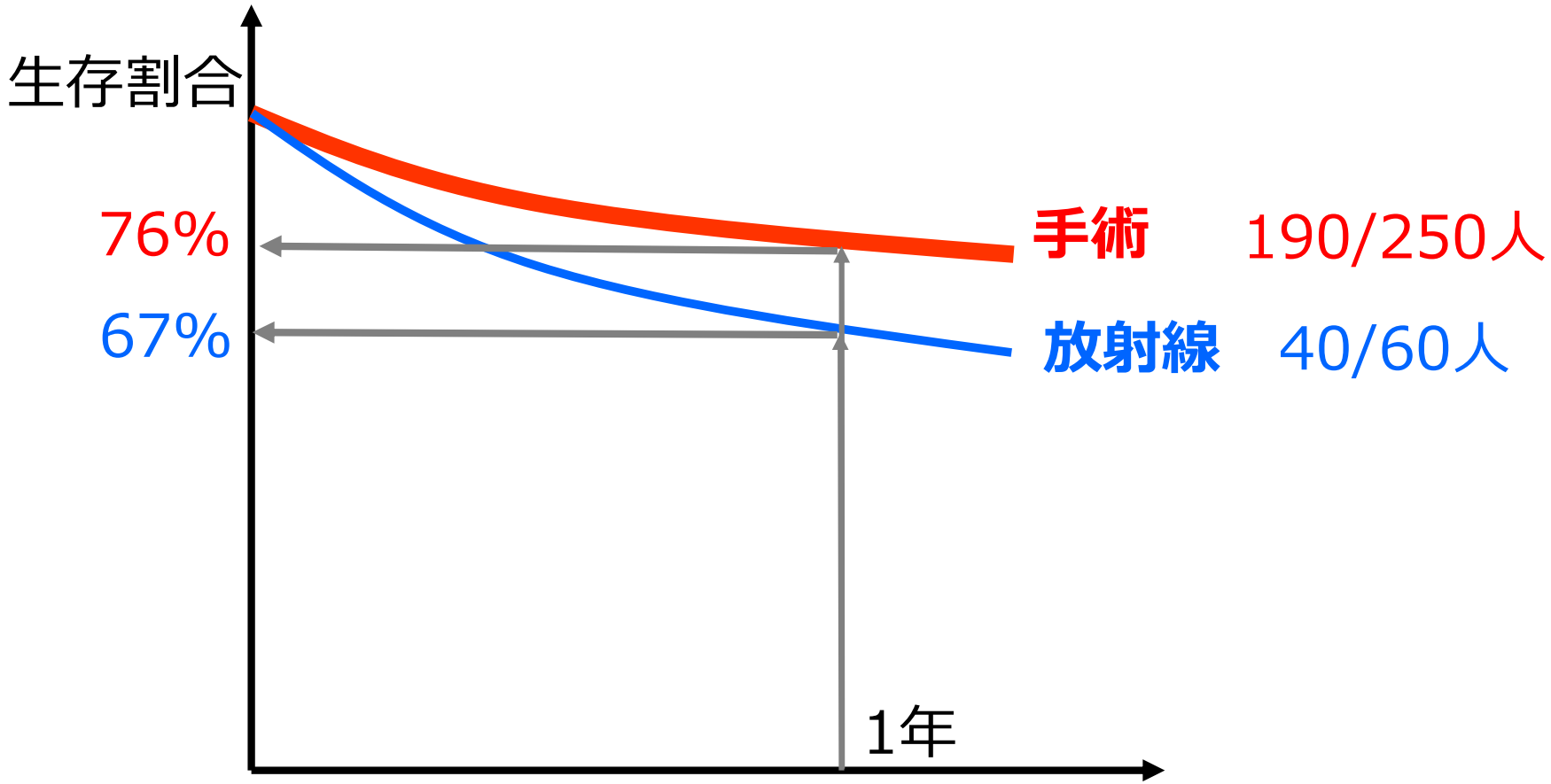
- 誤差的ばらつき
  - ランダム誤差
  - 精密度（precision）
  - サンプルサイズを大きくするほど小
- 系統的ばらつき
  - バイアス
  - 正確度（accuracy）
  - サンプルサイズを大きくしてもだめ



注：仮想例

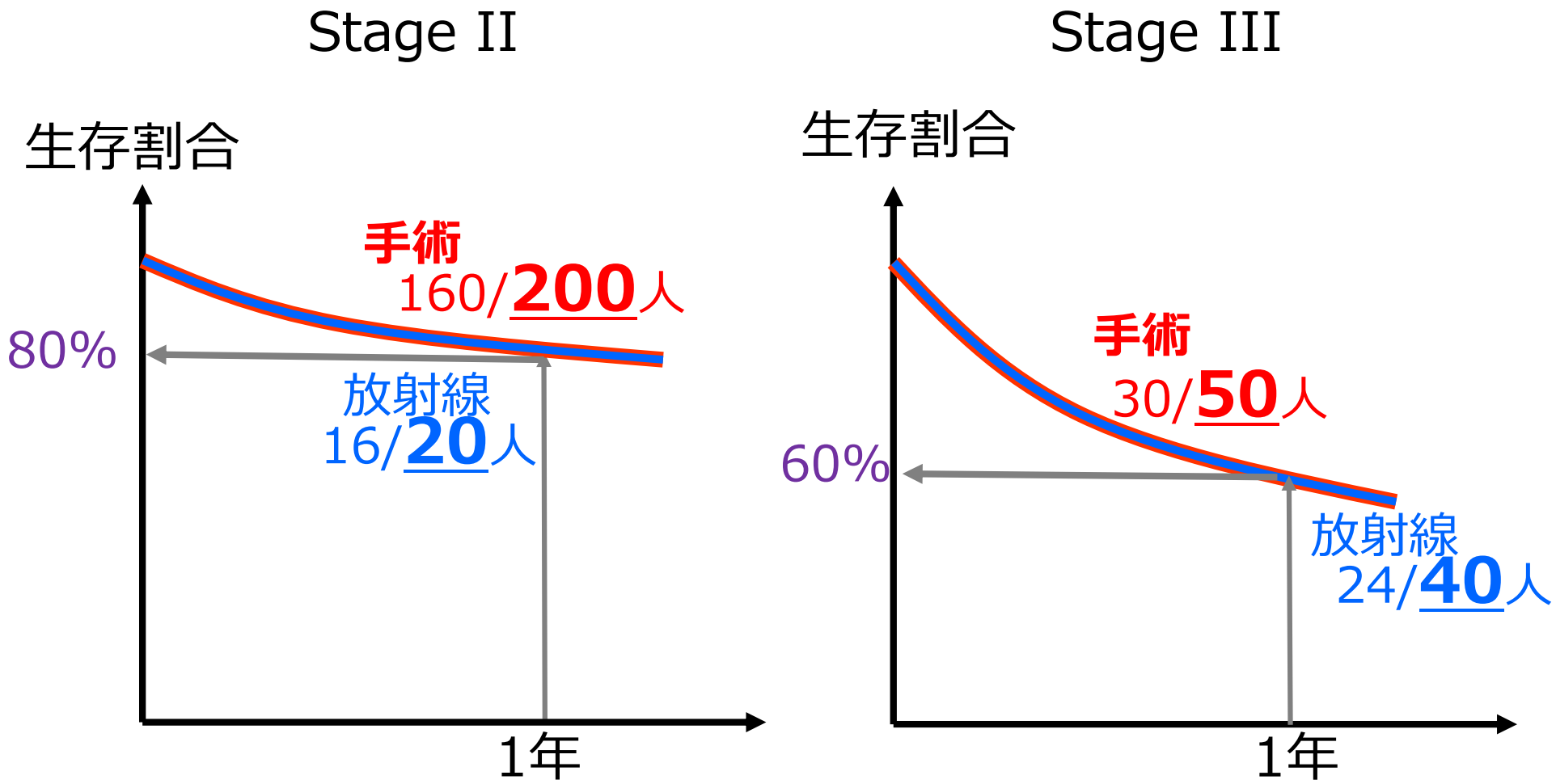
国内の学会で見かける発表

- ある病院の適格規準●●を満たす患者を**手術群**(250例)と**放射線療法群**(60例)に分けて後方視的（いわゆるレトロ）に予後に関する検討を行った。
- **手術群**は**放射線群**と比較し予後良好であった。
- この対象に**手術**をすることが推奨される。



注：仮想例

病期(Stage)で分けた場合の予後



Stageによらず**手術**と**放射線**の予後は変わらない

注：仮想例

比較したいのは手術と放射線の違いだから

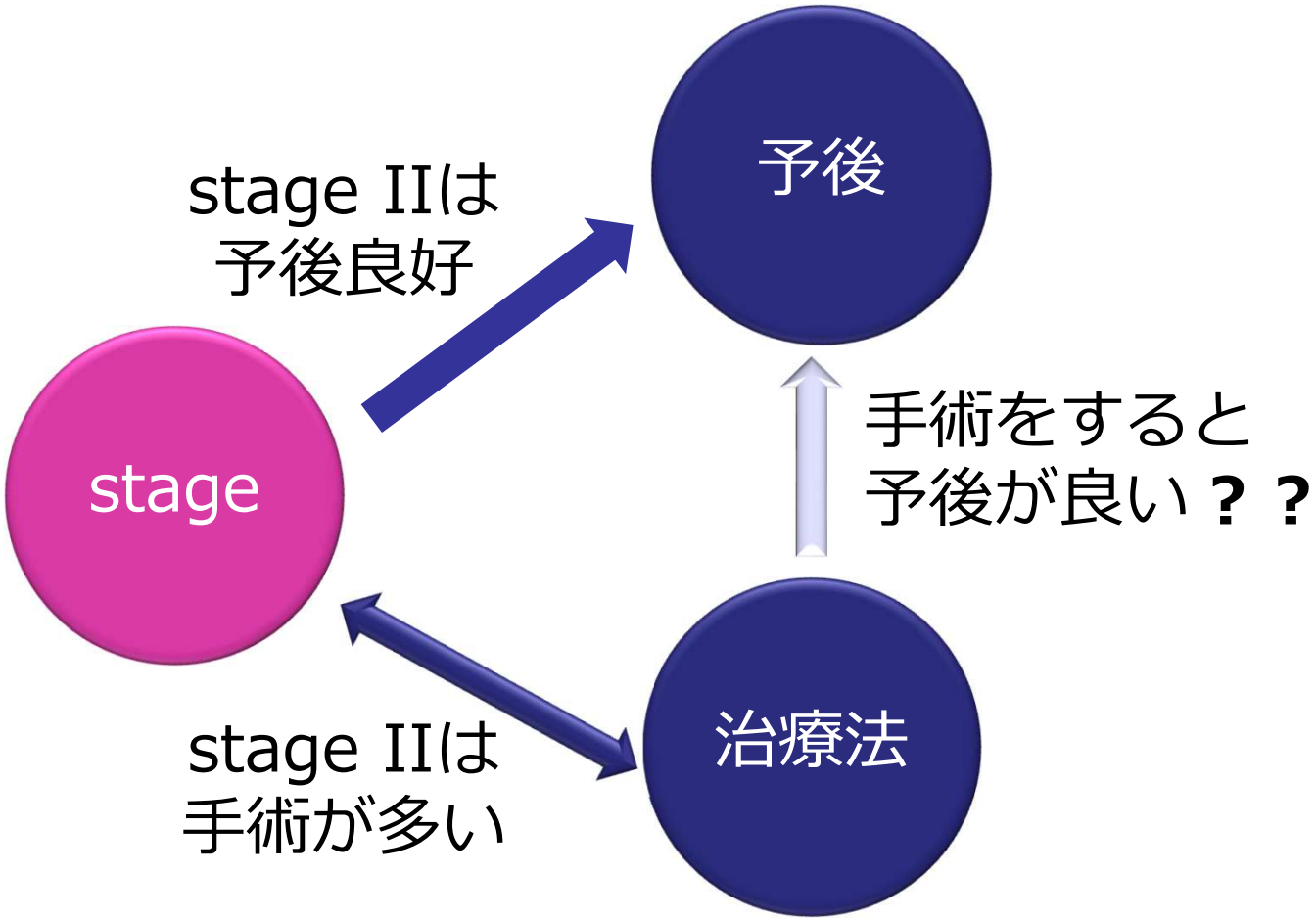
- 治療法以外の予後に影響する要因の条件が同じでなければ、”比較”にならない！！

治療法	stage II		stage III	合計
手術	**200人** (80%)	＞＞	50人	250人
放射線	20人 (33.3%)	＜＜	**40人**	60人

- 手術は放射線と比べ**「stage II」**の割合が高い
- stageによって予後が異なる(stage IIは予後良)

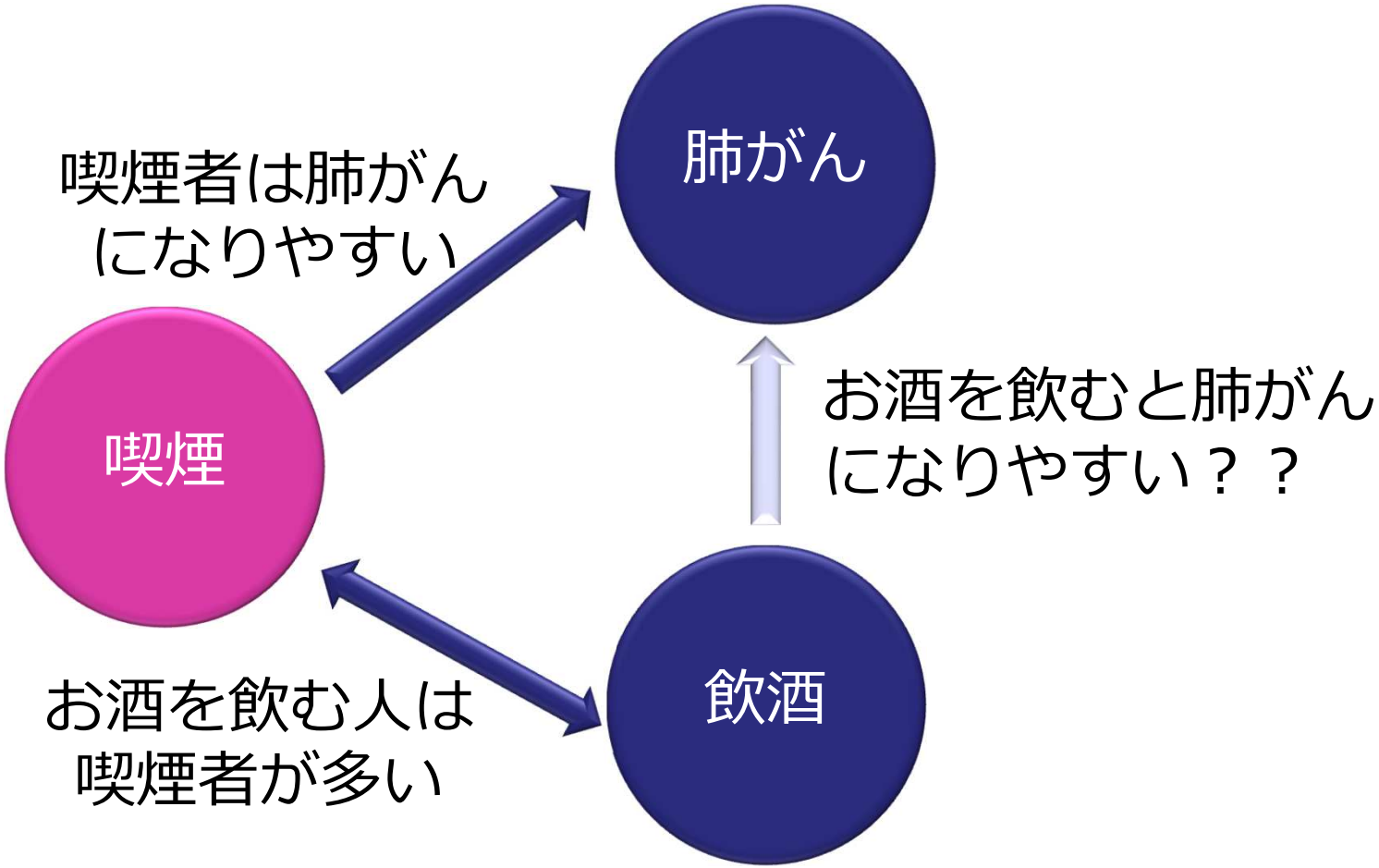
交絡についてのまとめ

- 治療法と予後に関連する第3の因子（stage）によって見かけ上の関連が生じてしまう現象のこと（**バイアス**の一種）
  - 交絡を引き起こす因子（=stage）のことを**交絡因子**という



その他の交絡の例

- 飲酒が肺がんを起こすわけではないのに、そう見えてしまう
  - この場合、喫煙が**交絡因子**



交絡がないことを保証するには

- 治療群間で予後に関係する背景因子を揃える
  - 年齢
  - 性別
  - Performance Status（全身状態）
  - その他…

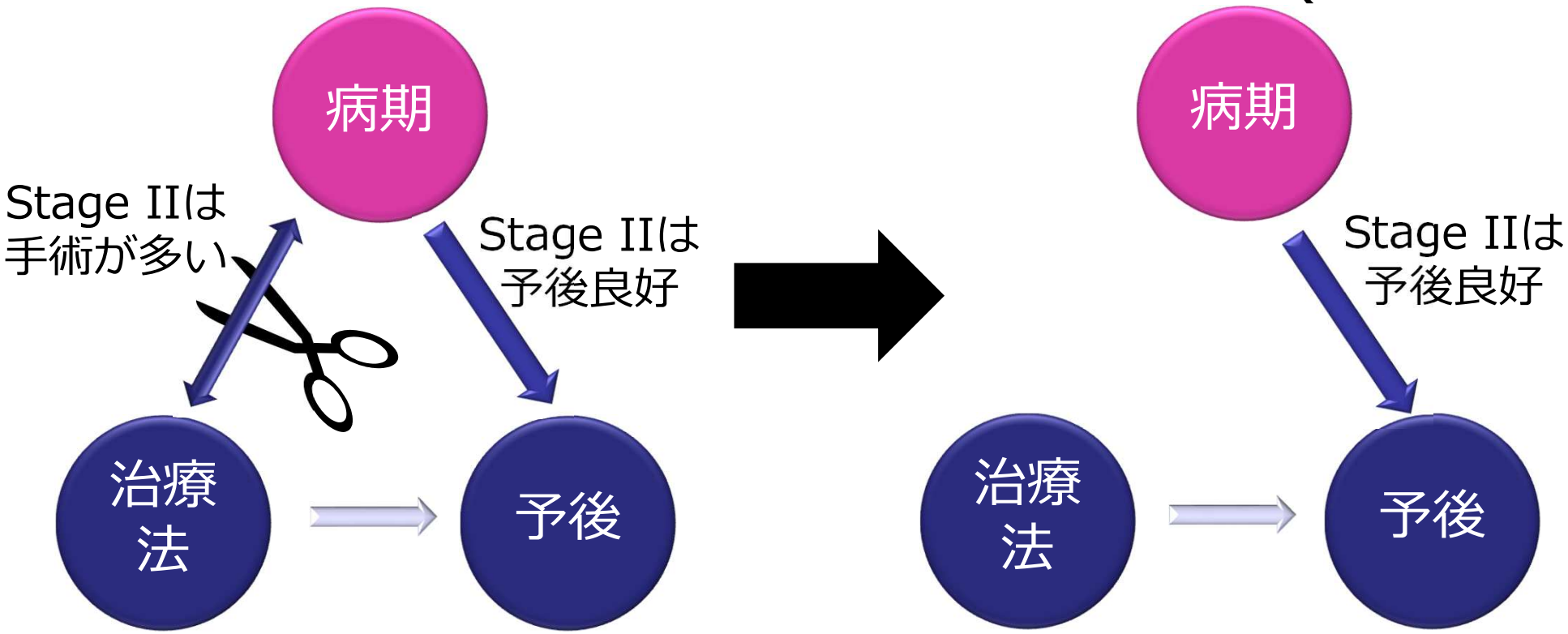
未知のものも含めて因子がたくさんあるために
全てを考慮できない

↓

患者が受ける治療法を**ランダム**に決める

ランダム化の意義

- 医師/患者の意思によらず確率に基づき治療群に患者を割付ける（ランダム化）
  - 例：候補患者に対して、背景因子がよりバランスされる群に高い確率で割付られる（最小化法）
- **治療法**と**交絡因子（病期）**の関連を除去できる
  - 治療群間で背景因子が等しくなるので、交絡が除去される
    - 治療法以外は等しい集団 → 効果に差があれば治療法の違い **(比較可能性の担保)**



希少フラクションの開発では…

一定期間で集積可能な症例数・試験期間の確保が難しい

→ランダム化比較試験の実施が困難では？

- 単群試験を検証的試験として実施？
  - ヒストリカルコントロールでよいのか？（患者選択バイアス）
- 国際共同試験としてランダム化比較試験を実施？
  - 研究のinitiativeを誰がとるのか？データセンターの機能は？
- がん種横断的な開発としてバスケット試験を実施？
  - がん種ごとに治療lineやモダリティに関するコンセンサスが必要

外部対照を用いた比較

RWDに関するガイドライン等

Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drug and Biological Products

Guidance for Industry

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Oncology Center of Excellence (OCE)

September 2022
Procedural

3495713fnl09-06-22.docx

Understanding the Need for Non-Interventional Studies Using Secondary Data to Generate Real-World Evidence for Regulatory Decision Making, and Demonstrating Their Credibility

November 25, 2019

Duke | MARGOLIS CENTER for Health Policy

A Roadmap for Developing Study Endpoints in Real-World Settings

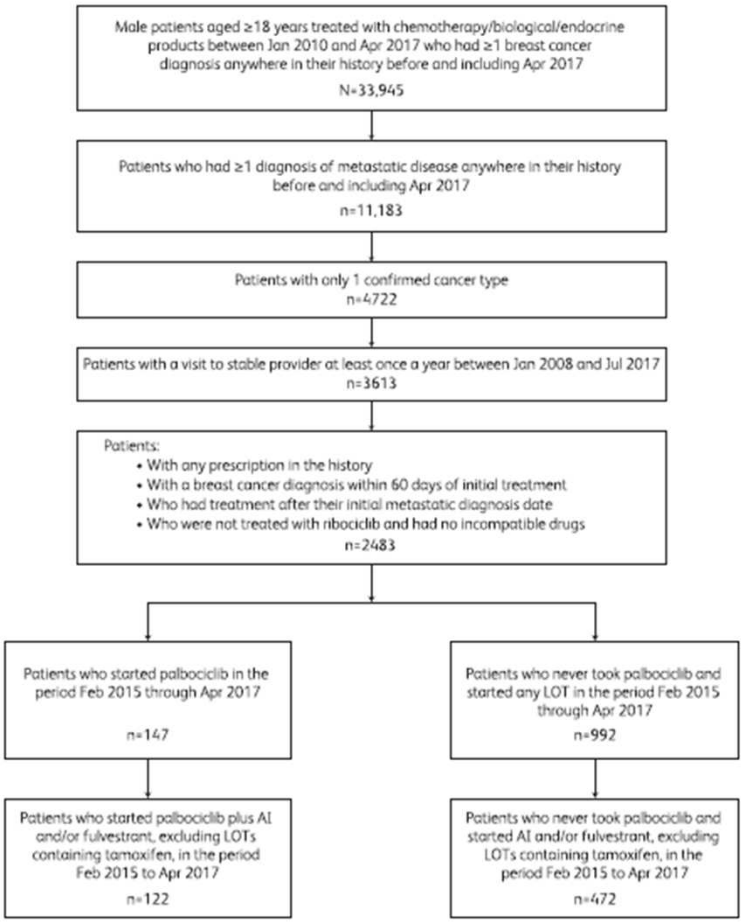
August 28, 2020

Duke | MARGOLIS CENTER for Health Policy

Albert L Kraus *et al.* (2021) Clin Pharmacol Ther;111(1):302-309.

Real-World Data of Palbociclib in Combination With Endocrine Therapy for the Treatment of Metastatic Breast Cancer in Men

(a)



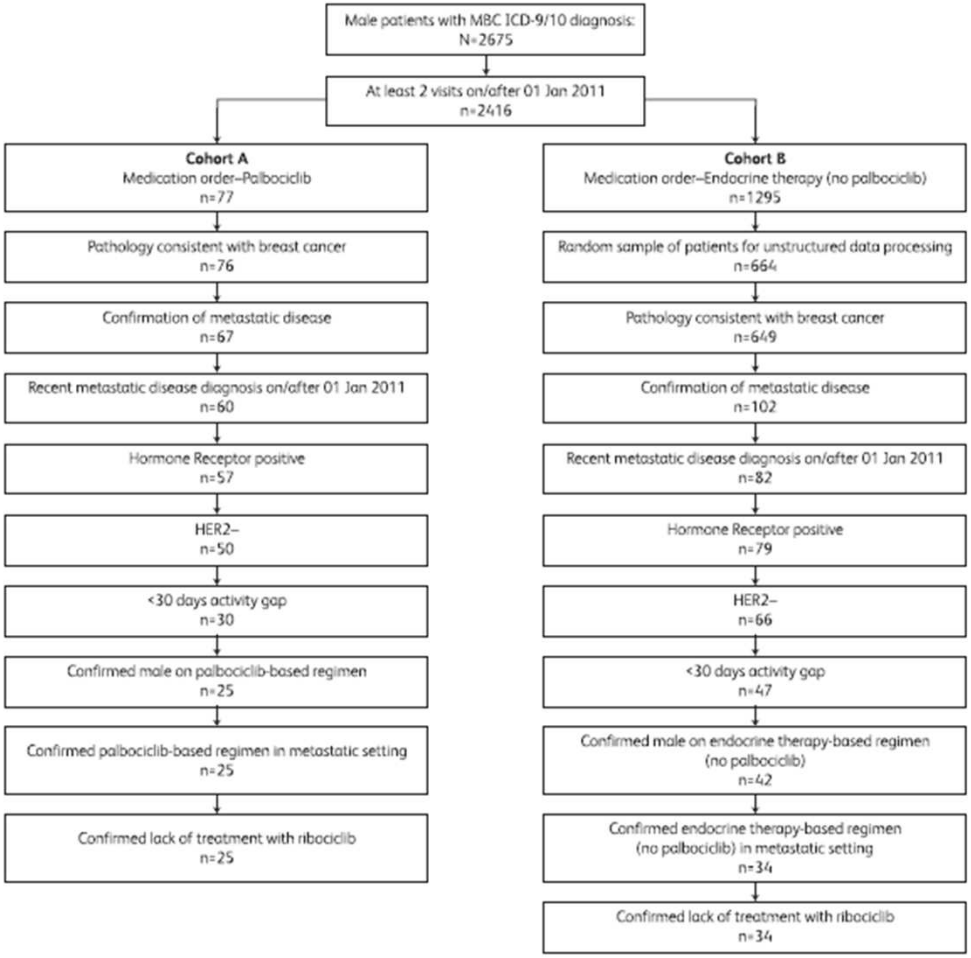


Figure 1 Cohort selection. (**a**) Pharmacy and medical records database (**b**) EHR-derived database. AI, aromatase inhibitor; EHR, electronic health record; HER2–, human epidermal growth factor receptor 2–negative; ICD, International Classification of Diseases; LOT, line of therapy; MBC, metastatic breast cancer.

Albert L Kraus *et al.* (2021) Clin Pharmacol Ther;111(1):302-309.

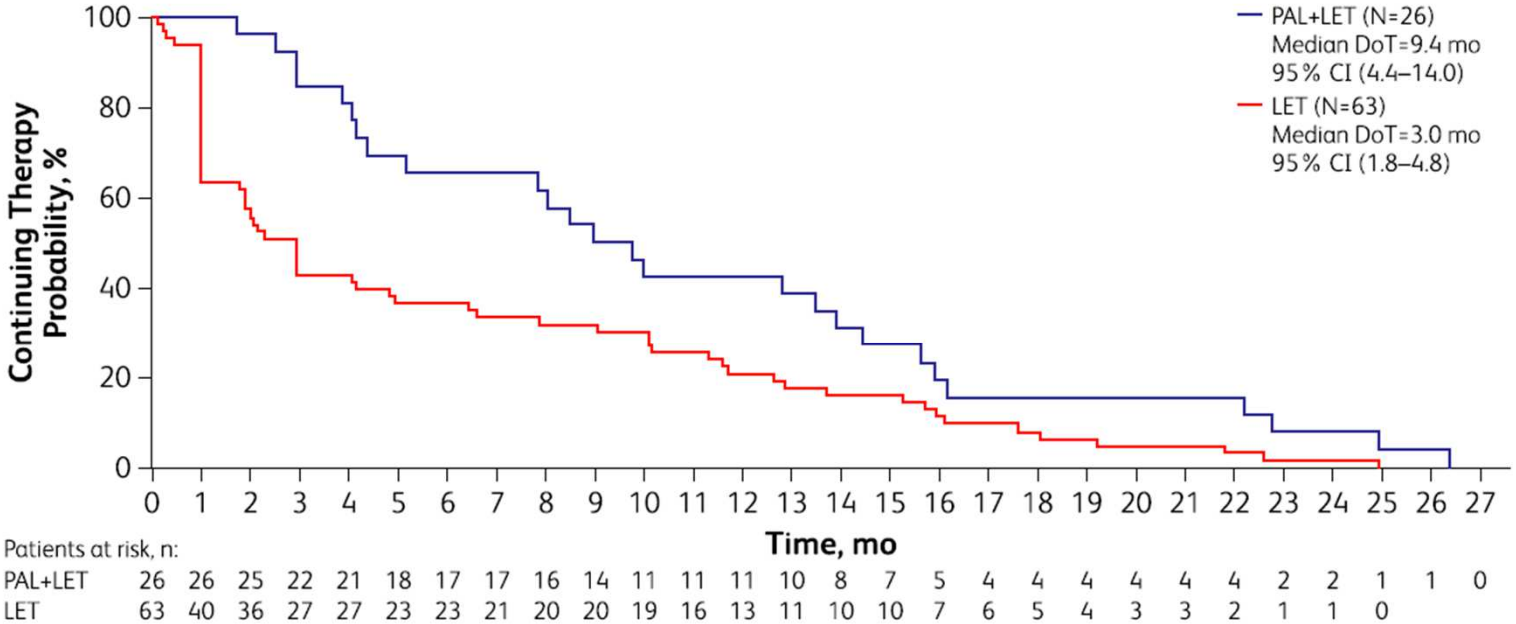


Figure 2 Pharmacy and medical records database. Kaplan-Meier plot of duration of therapy in the first-line setting comparing PAL + LET with LET alone. CI, confidence interval; DoT, duration of therapy; LET, letrozole; PAL, palbociclib. [Colour figure can be viewed at wileyonlinelibrary.com]

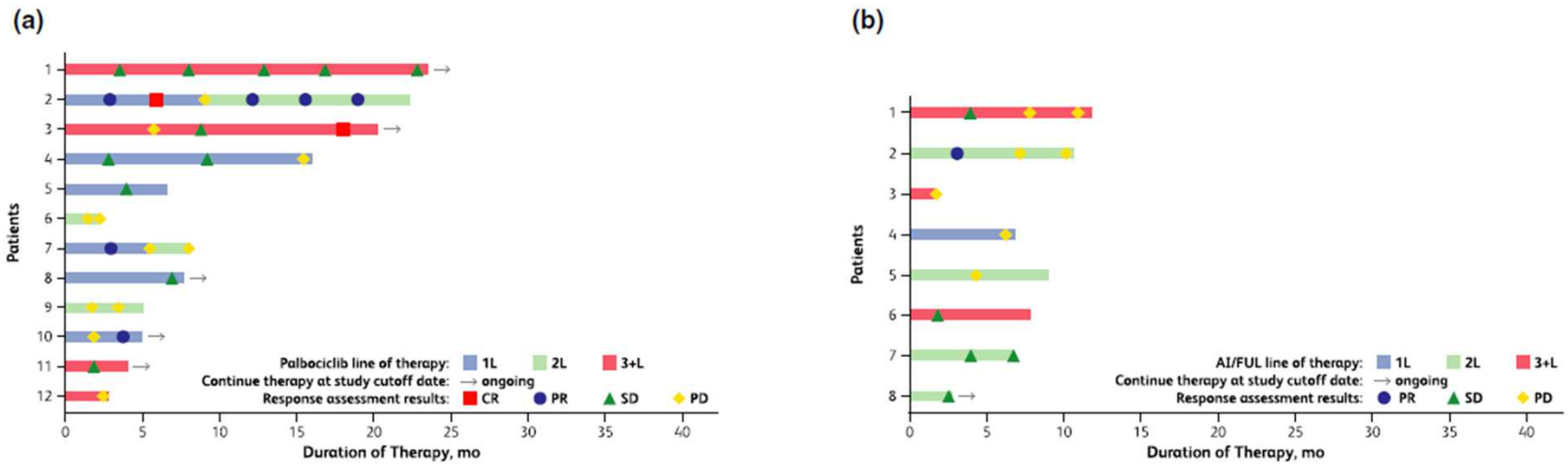


Figure 3 EHR-derived database. (**a**) Duration of therapy and response assessments in the palbociclib plus ET cohort and (**b**) duration of therapy and response assessments in the ET-alone cohort. AI, aromatase inhibitor; CR, complete response; EHR, electronic health record; ET, endocrine therapy; L, line; FUL, fulvestrant; PD, progressive disease; PR, partial response; SD, stable disease. [Colour figure can be viewed at wileyonlinelibrary.com]

Albert L Kraus *et al.* (2021) Clin Pharmacol Ther;111(1):302-309.

CONCLUSIONS

Collective data on palbociclib, including mechanism of action, PKs, pivotal phase II and III randomized clinical trials in patients with HR+/HER2– MBC, and real-world data in men with MBC presented in this report, led to the expansion of the approved palbociclib indication to include men with HR+/HER2– MBC in the United States. Real-world data contributed to the totality of evidence that men with MBC can benefit from the addition of palbociclib to ET, with a safety profile consistent with previous observations in women. The novel experience described here highlights a potential new pathway to expand a treatment's indication and increase patient access based on multiple sources of real-world evidence in rare tumors, where rarity makes large randomized studies challenging.

- Palbociclibの男性乳がんに対する適応拡大
- RWDの利活用の可能性を示した

研究のlimitation(Discussionより抜粋)

The findings from the present analysis further support palbociclib as a treatment for HR+/HER2– MBC in men in the real-world setting. However, one of the limitations of this retrospective analysis is the small dataset, due to the rarity of breast cancer in men. Additionally, inherent limitations associated with EHR data collected during the patient journey and medical and pharmacy claims data collected for billing/reimbursement purposes rather than for research purposes include potentially incomplete/missing data. Moreover, medication compliance was not assessed in this analysis, and confounding factors (e.g., imbalance of baseline characteristics) cannot be excluded. Additionally, in the current study, real-world response was evaluated based on clinical assessment and radiologist notes, whereas the palbociclib phase III clinical trials assessed tumor response according to the Response Evaluation Criteria in Solid Tumors (RECIST) criteria.[9,11] However, an analysis of female patients who received letrozole as first-line therapy for HR+/HER2– MBC in the real-world setting vs. the control arm of PALOMA-2 showed concordance between real-world outcomes (e.g., real-world response rate) and RECIST-based outcomes (e.g., objective response).[30] As is common for retrospective observational studies, other potential limitations include selection bias, limited ability for statistical comparisons in small populations, and potential confounding bias due to the lack of randomization.

レジストリデータの活用

- 東病院の医師主導治験として、単群試験の外部対照としてレジストリデータを活用を計画した試験例

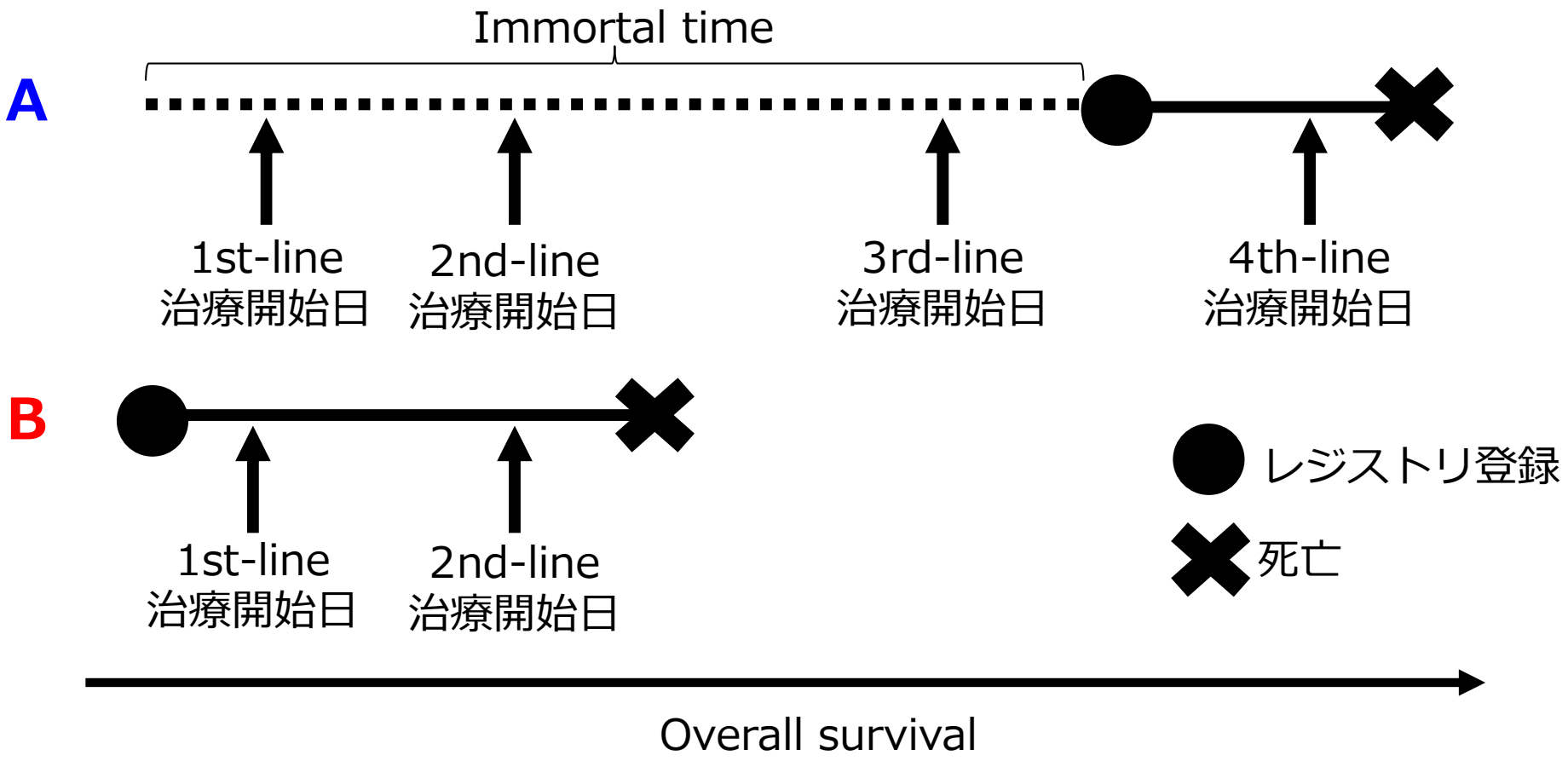
試験名称（略称）	試験登録番号	対象	主要評価項目
EPOC1602 (TRIUMPH)	UMIN000027887	HER2陽性大腸がん	ORR
EPOC1703 (BIGBANG)	UMIN000031857	BRAF non-V600変異大腸がん	ORR
EPOC1708 (METBEIGE)	JapicCTI-194677	MET遺伝子増幅を有する大腸がん	ORR
EPOC1805 (TiFFANY)	JapicCTI-194624	ctDNAにFGFR遺伝子異常の固形がん	ORR
EPOC1806 (HERALD)	JapicCTI-194707	ctDNAでHER2遺伝子増幅を有する固形がん	ORR
EPOC2101 (NEXUS)	jRCT2031220025	BRAFV600E変異を有する切除可能大腸がん	1y-PFS

外部対照の問題点

- データの収集プロセスの違い
  - 個人情報の取り扱い、品質保証の管理体制、患者同意の有無
- データの収集目的の違い（二次利用を見据えていない？）
  - 解析に必要なデータの欠測
  - 従うガイドラインの違い等、本体研究とは異なる定義でのデータ収集（例：薬剤コード、有害事象コード）
- データの収集時期の違い
  - いわゆる時代間差（後治療、術式、BSCの成績・効果の向上）
- バイアス
  - 交絡
  - Immortal time bias 等

Immortal time bias

Immortal timeが存在することで生じるバイアス



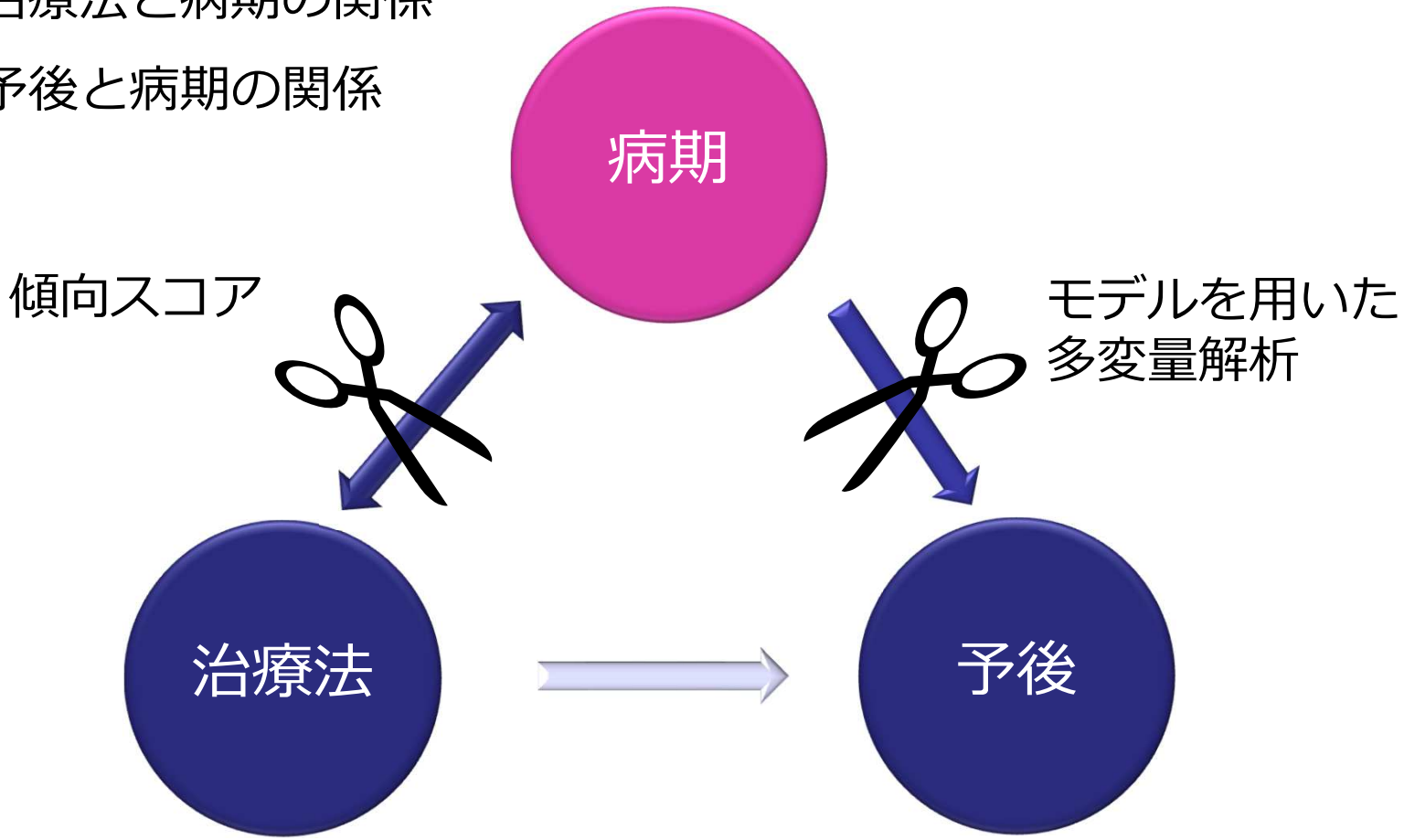
登録日からのOS：**A**<**B**、1st-line治療開始日からのOS：**A**>**B**

レジストリに登録されるまで生存しているselectionがかかる

外部対照を用いる場合の
交絡への対処法

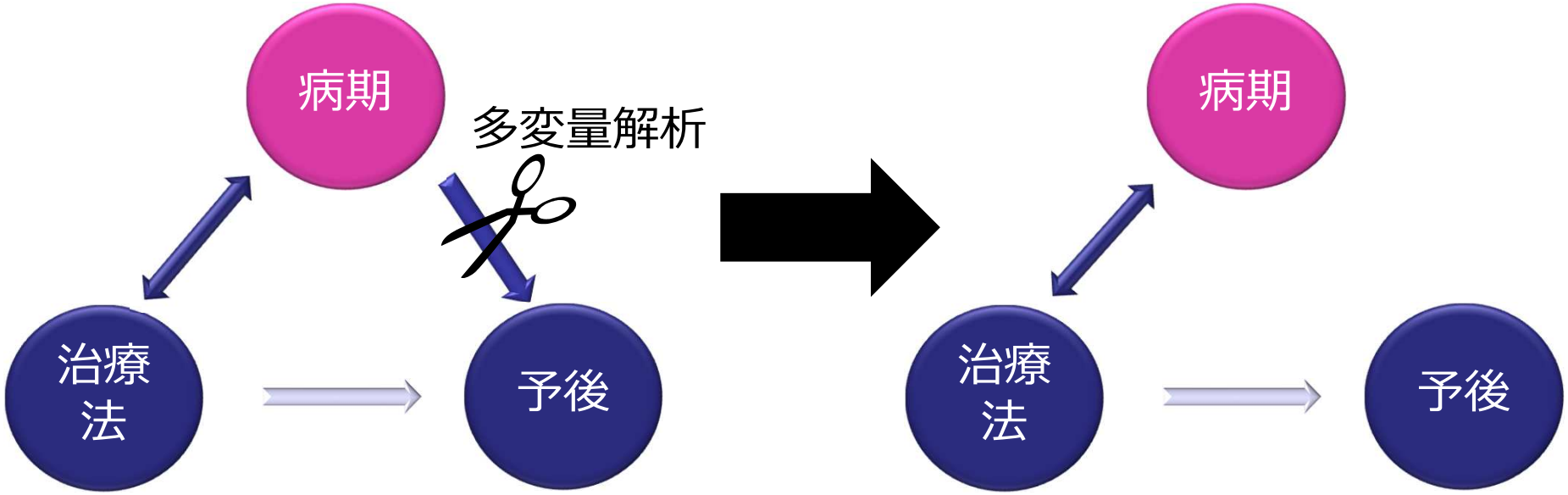
解析段階で交絡を除去するには？

- ランダム化はあくまでデザイン段階（データ収集時）の対処法
- 解析段階で少なくともどちらか一方の関係を除去できないか？
  - 治療法と病期の関係
  - 予後と病期の関係



統計モデルを用いた多変量解析

- ロジスティック回帰やCox回帰の使用
  - 複数の交絡因子を同時に考慮できる
  - モデルの仮定が正しければ強力な方法
- オーバーフィッティングの問題
  - イベント数に対して説明変数をモデルに組み入れ過ぎていないか？
- 多重共線性の問題
  - 相関の強い変数を同時にモデルに組み入れていないか？



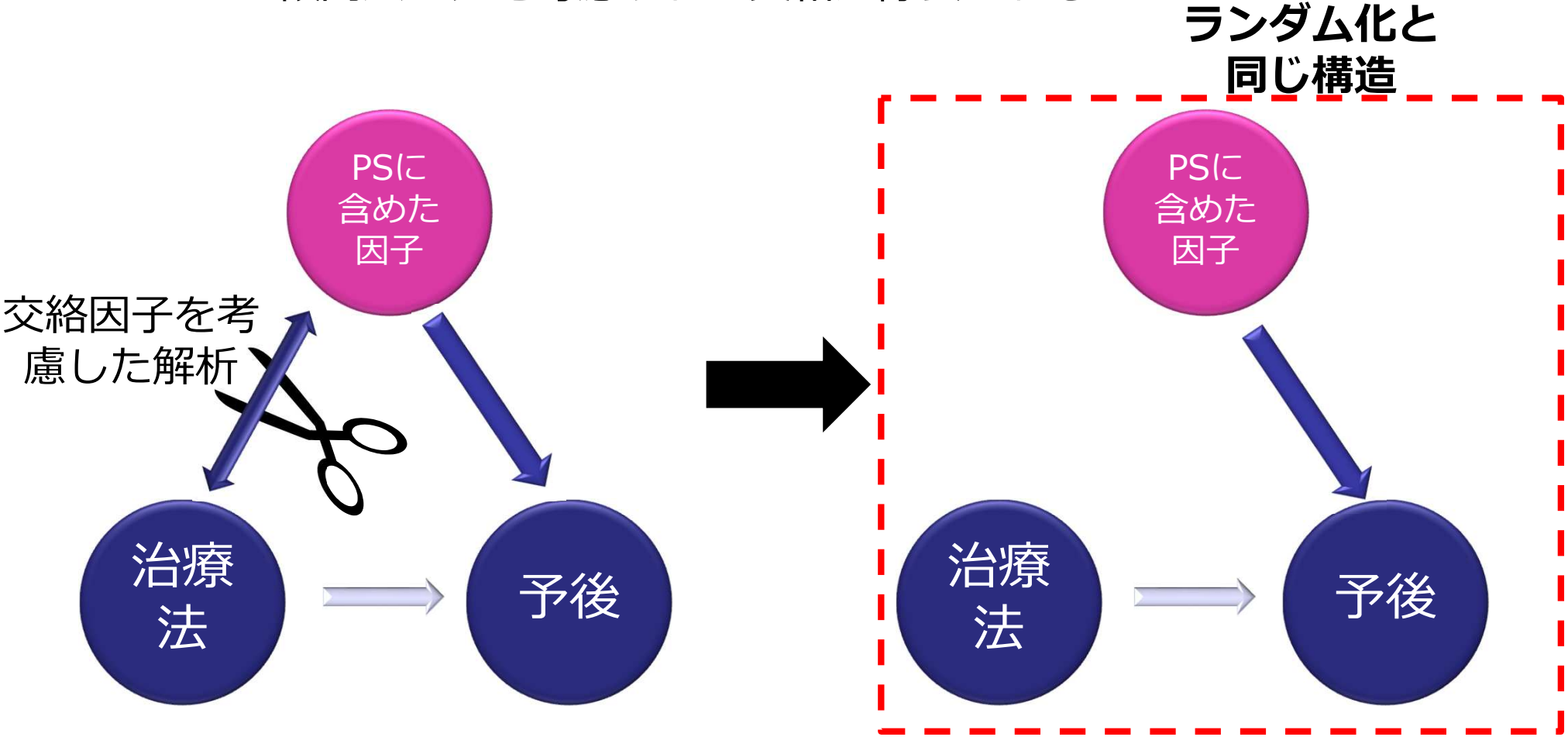
Rosenbaum P.R., Rubin D.B. Biometrika. 1983;70:41–55.

傾向スコア（Propensity Score）

- **傾向スコア**とは
  - ある背景因子を持つ対象が、一方の治療法に割り付けられる確率
  - 取り得る値の範囲：0～1
- 治療法の選択に関係する複数の因子を**1つのスコアに要約**する
  - 1つ1つの背景因子は違っても、傾向スコアが同じであれば、一方の治療法への割り付けられやすさは同じとみなせる
  - 傾向スコアが同じ対象を集団で考えれば、治療群間で背景因子が平均的に同じであることが期待できる

傾向スコアで調整した比較の意義

- **<u>治療法</u>**と**<u>傾向スコア[PS]（≒背景因子）</u>**の関連を除去できる
  - 交絡を除去する構造がランダム化と同じになる
  - 傾向スコアを考慮すれば交絡が除去される



傾向スコアを用いた解析法

- 2段階の手順で行う

① 傾向スコアの推定

- 患者ごとの傾向スコアを推定する

② 傾向スコアを用いて調整した解析

- マッチング
- 層別解析
- 多変量解析
- IPTW

本講義では説明を割愛

患者ごとの傾向スコアの推定

（治療Aと治療Bの比較を仮定した場合）

- 傾向スコア（PS）は治療Aを受ける確率で、0<PS<1の値を取る
  - ≒0：治療Aを受ける確率はほぼない（ほぼ間違いなく治療Bを受ける）
  - ≒1：ほぼ間違いなく治療Aを受ける
- 通常は**ロジスティック回帰モデル**を用いて推定する

$$\log\left(\frac{PS}{1-PS}\right) = \beta_1 \cdot 年齢 + \beta_2 \cdot stage + \beta_3 \cdot 性別 + \cdots + \beta_i \cdot 併存症$$

$$\Downarrow$$

$$PS = \frac{\exp(\beta_1 \cdot 年齢 + \beta_2 \cdot stage + \beta_3 \cdot 性別 + \cdots + \beta_i \cdot 併存症)}{1+\exp(\beta_1 \cdot 年齢 + \beta_2 \cdot stage + \beta_3 \cdot 性別 + \cdots + \beta_i \cdot 併存症)}$$

- 傾向スコアは**「予後(アウトカム)」の情報を用いずに**推定可

*傾向スコア=治療Bを受ける確率として推定することも可能

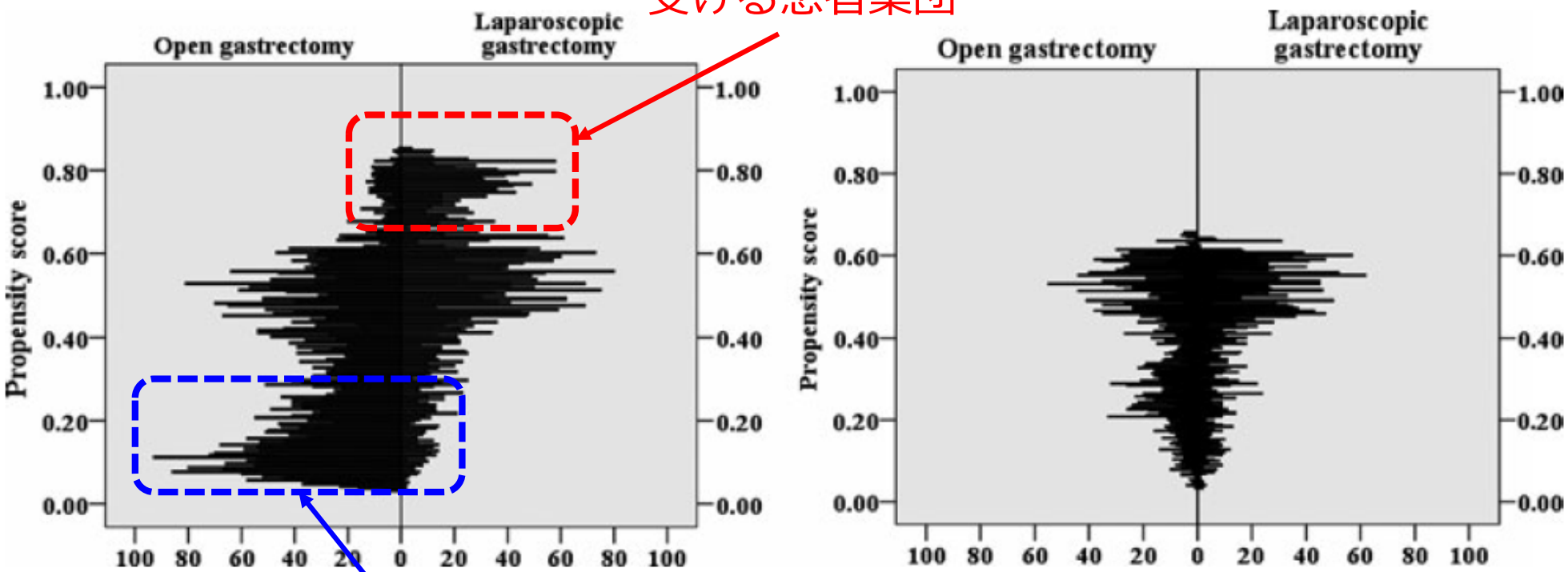
マッチングの例

- 早期胃がんに対して開腹切除術(OP) vs 腹腔鏡切除術(LAP)

全患者を対象にした
傾向スコアの分布

ほぼ間違いなくLAPを
受ける患者集団

マッチング後の
傾向スコアの分布

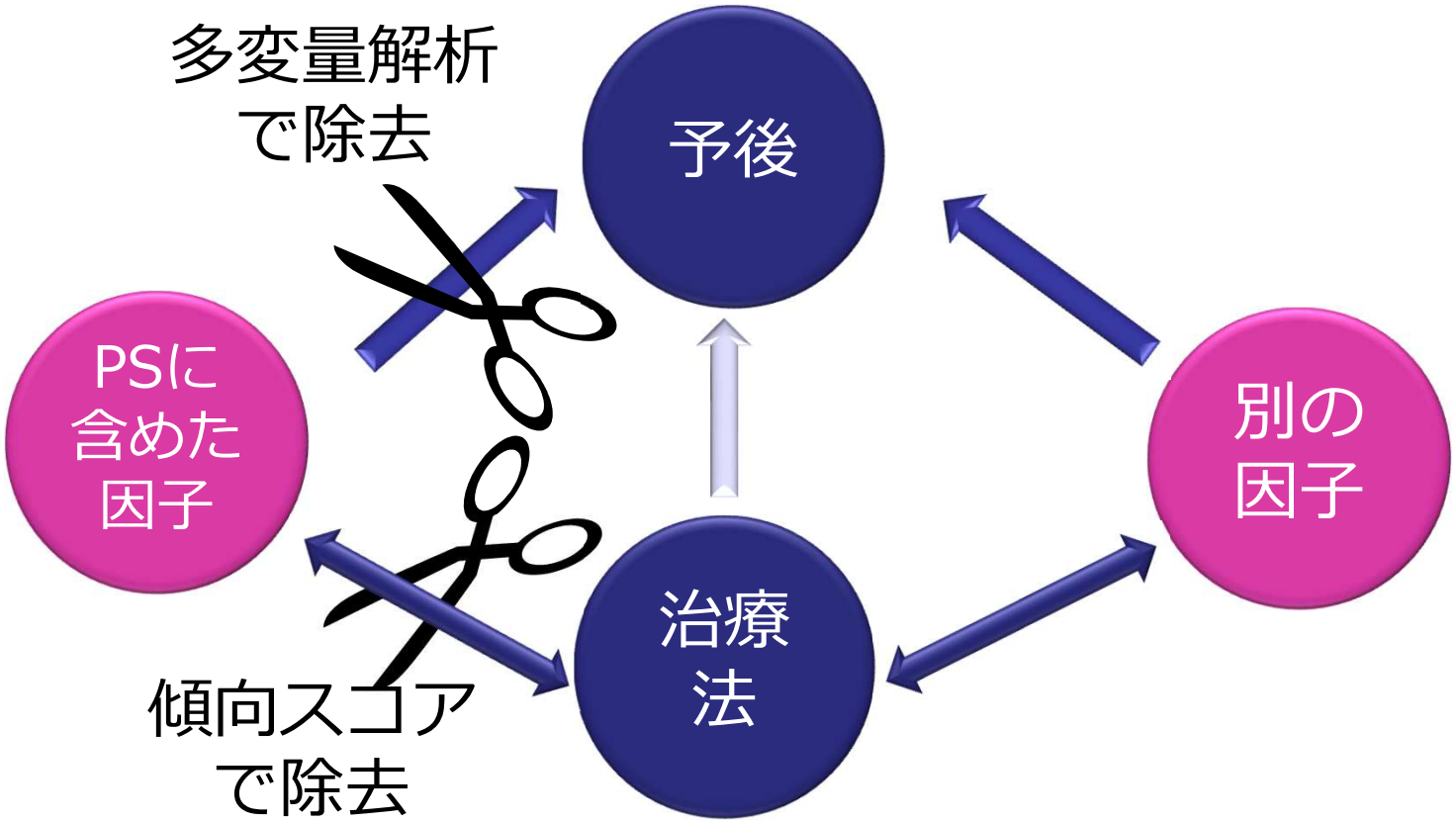


ほぼ間違いなくOPを受ける患者集団

Ann Surg 2013;257: 640–646

傾向スコア & 多変量解析の欠点

- **傾向スコア(PS)の解析で調整していない因子の影響が除去できない**
  - 調整していない因子が影響していたら、バイアスが入ったまま



傾向スコア解析のメリット(vs.多変量解析)

1. **推定した傾向スコアが妥当かを評価することが容易**
   - 群間で患者別の背景因子の分布の確認
   - 背景因子がほとんど揃わなければ、揃っている対象だけで解析 or 調整した解析は不能と判断できる
2. **傾向スコアはアウトカムを用いずに推定可能**
3. **アウトカムがレアな時でも柔軟に対応可能**
   - 目安として多変量解析では1つの変数に対し10イベント必要(とされている)
   - モデルを使った多変量解析の場合も多くの変数を入れなくて良い
4. **傾向スコア推定のためのモデルの誤特定に対し頑健**
   - コンピュータの乱数を用いたシミュレーションでは、傾向スコア推定のためのモデルが間違っていても治療効果の推定値にバイアスが入りにくい

（注）ランダム化に代わる方法ではない

- 未知 or 既知でもデータとして収集されていない交絡因子の存在
  - 傾向スコアと多変量解析による対応の限界
  - 未知の交絡因子もバランスできるランダム化には勝てない

- サンプルサイズが少数の場合はかえってバイアスを生む可能性
  - モデルの仮定やマッチングの限界
  - Limitationとして理解しつつ調整しない解析にとどめるのも選択肢の1つ（前述した男性乳がんの例）

- 交絡に対する対処法は他にもある
  - 操作変数法、Bayes流の統計手法
  - 内部対照（ランダム化した場合の対照群）と外部対照の両方を活用する方法（Hybrid control）

外部対照を用いた研究の報告

Reportingガイドライン

- 臨床研究に関するreportingガイドライン
  - https://www.equator-network.org/ (EQUATOR network)
    - CONSORT, STROBE, PRISMA, STARD, TRIPOD, 等
  - 研究の完全(complete)、明快(clear)、透明性(transparency)の確保

- Real World Dataを外部対照として利活用した場合の記載方法
  - 外部対照の説明も含めて研究の科学的妥当性を示す必要がある
  - 直接参考にできるガイドラインが無い
    - 臨床試験のCONSORT？ or 観察研究に関するSTROBE？

Schulz KF *et al.* (2010) BMC Medicine 8:18.

CONSORT声明

CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials

Kenneth F Schulz*[1], Douglas G Altman[2], David Moher[3] for the CONSORT Group

Abstract

The CONSORT statement is used worldwide to improve the reporting of randomised controlled trials. Kenneth Schulz and colleagues describe the latest version, CONSORT 2010, which updates the reporting guideline based on new methodological evidence and accumulating experience.

To encourage dissemination of the CONSORT 2010 Statement, this article is freely accessible on bmj.com and will also be published in the Lancet, Obstetrics and Gynecology, PLoS Medicine, Annals of Internal Medicine, Open Medicine, Journal of Clinical Epidemiology, BMC Medicine, and Trials.

CONSORT声明

- 臨床試験における報告事項のチェックリスト
  - タイトル・抄録 (Title and Abstract)
  - はじめに (Introduction)
    - 背景・目的
  - 方法 (Method)
    - 試験デザイン、参加者、介入、アウトカム、症例数、ランダム化、割振りの隠蔵機構、実施、ブラインディング、統計学的方法
  - 結果 (Result)
    - 参加者の流れ、募集、ベースライン・データ、解析された人数、アウトカムと推定、補助的解析、害
  - 考察 (Discussion)
    - 限界、一般化可能、解釈
  - その他の情報 (Other infomation)
    - 登録、プロトコール、資金提供者

外部対照の利活用時に特に留意・検討が必要と考えられる項目

まとめ

- 予後に関係する背景因子が治療群間で偏ることで、見かけ上の関連が生じる or 見かけ上関連が消える現象を**交絡**という
  - 交絡は**ランダム化**によって排除できる

- RWDを外部対照として用いる場合の注意点
  - 目的に合ったデータベースの選択（Fit for purpose）
  - 種々のバイアスの存在（交絡やimmortal time bias 等）

- 解析段階での交絡への対処法の例
  - 多変量解析と傾向スコア