

## NCNP-TMC 臨床研究実践講座ワークショップ

# 交絡、誤差、バイアス、 その対処法

東京大学大学院医学系研究科 公共健康医学専攻 生物統計学分野  
国立精神・神経医療研究センター 情報管理・解析部 生物統計アドバイザー  
柏原 康佑

2015.1.23

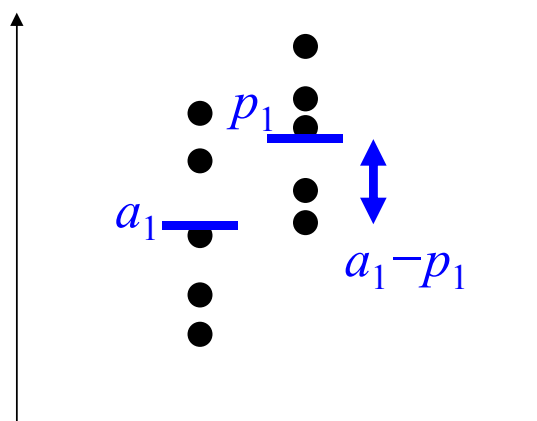
## 講義の内容

---

- バラツキとバイアス
- バラツキへの対処法
- 主なバイアスとその対処法

## 例: 大うつ病に対するSSRIの効果を検証するランダム化比較臨床試験

HAM-D  
6週時変化量

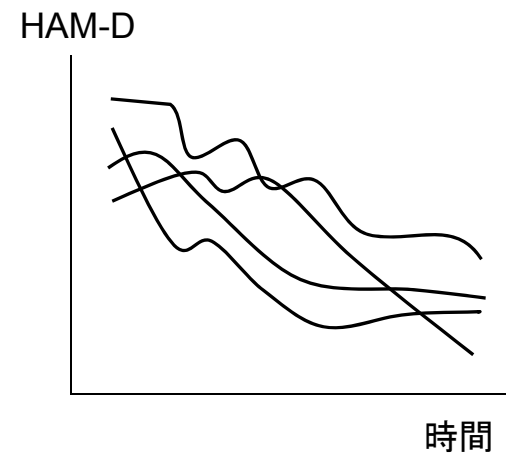


実薬 プラセボ

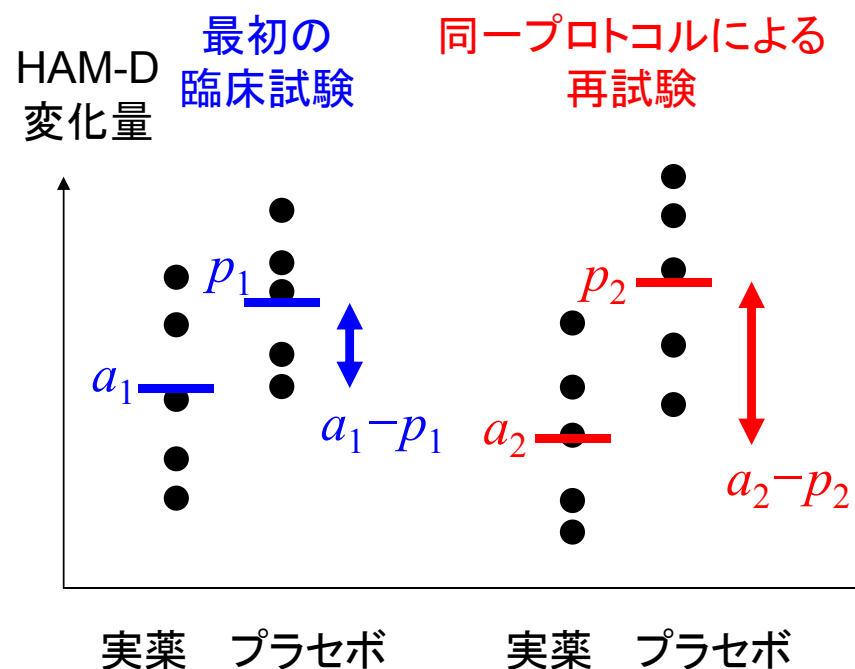
- 設定
  - 軽度から中等度の大うつ病患者
  - SSRIを6週間投与
  - ハミルトンうつ病評価尺度 (HAM-D)により評価
  - 投与前後の変化量を群間比較
- 結果
  - 実薬群の平均値:  $a_1$
  - プラセボ群の平均値:  $p_1$
  - 群間差:  $a_1 - p_1$

## データにはバラツキがある

- 個体間の差
  - HAM-Dは被験者によって異なる
- 個体内の差
  - HAM-Dは時間とともに変化する
- (誤差的)バラツキ random error
  - 同じものを測定しても値が異なること



## 2種類のバラツキ



1. データのバラツキ
  - 個体間差、個体内差
  - 標準偏差に対応
2. 推定値のバラツキ
  - 試験間のバラツキ
    - 群ごとの平均値
      - $a_1 \neq a_2$
      - $p_1 \neq p_2$
    - 平均値の差(群間差)
      - $a_1 - p_1 \neq a_2 - p_2$
  - 1のバラツキが原因
  - 標準誤差に対応

## 2種類のバラツキの特徴

1. データのバラツキ
    - データとして観測可能
    - 標準偏差は症例数を増やしても増加も減少もしない
      - ・ ヒストグラムの山が高くなるだけ
  2. 推定値のバラツキ
    - 同一の研究を繰り返さない限り観測不可能
    - 標準誤差は症例数を増やした分だけ減少する
      - ・ 症例数を増やせば、推定値の精度(精密度)が上がる  
(= 信頼区間幅が狭くなる)
- 臨床研究で本質的に制御したいバラツキは 2
    - そのために1も制御

## 症例数を増やせば答えが得られる？

- 症例数を限りなく増加すれば標準誤差は0
  - バラツキのない一意的な推定値 (信頼区間幅が0)
- その推定値がその臨床研究で観察したい曝露効果・治療効果の真値(true value)か？
  - Yes → **バイアス**がない
  - No → **バイアス**がある
- バイアス bias
  - 推定値からバラツキを除いても残る真値からの系統的なズレ

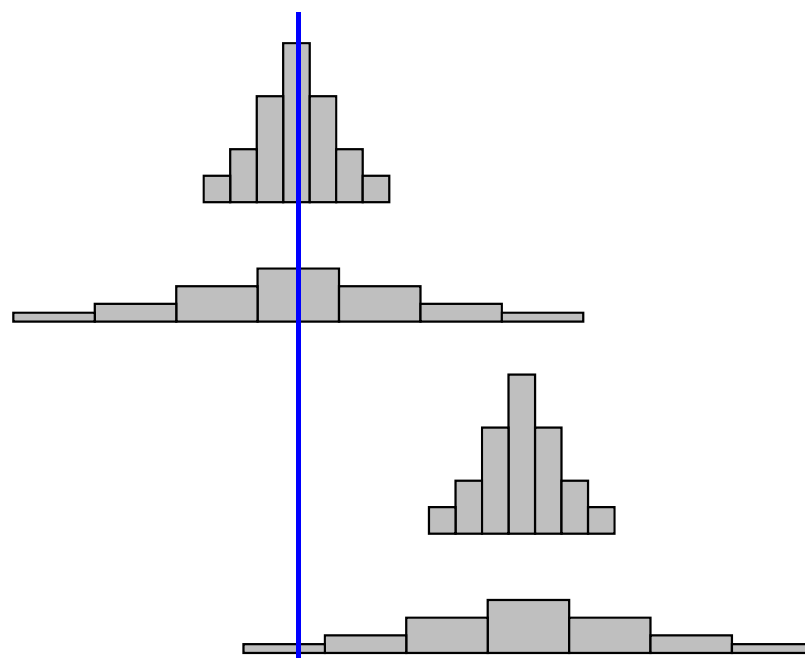
## 例: 日本人成人の平均身長を知りたい

- 全員の身長を測定すれば平均身長が分かる
- 全員は難しいので一部を抽出
  - デザイン1: 男性だけランダムサンプリング
  - デザイン2: 東京都からランダムサンプリング
  - デザイン3: 自分の友人全員の身長を調べる
  - いずれもどれだけ症例数を増やしても正しい値は得られない = **バイアス**がある
- 日本人成人全体からランダムサンプリング
  - **バイアス**のない値が得られる



## 精密度と正確度 precision and accuracy

観測値 = 真値 + バイアス + 誤差的バラツキ  
observed = true + bias + random error



accuracy      precision

○

○

○

×

×

○

×

×

真値 true value

## 医学研究は一度限り

---

- 得られる推定値を可能な限り真値に近づけたい
  - **バイアス**が小さい方がよい
  - **バラツキ**も小さい方がよい

## 医学研究の目標： 強いエビデンスを得ること

### エビデンスの強さ: 3つの評価基軸

- 誤差的**バラツキ**を小さくすること
  - (精度を高めること)
    - 精密度 precision; 信頼性 reliability; 再現性 reproducibility
- 偏り(**バイアス**)を小さくすること
  - 正確度 accuracy; 内的妥当性 internal validity
- 一般化可能性を高めること
  - 一般化可能性 generalizability;  
外的妥当性 external validity

## 補遺: バイアスの正確な定義

- $(\text{バイアス}) = (\text{推定値の期待値}) - (\text{真値})$
- 2種類のバイアス
  1. 症例数を増加すれば0に近づくバイアス
  2. 症例数を増加しても0に近づかないバイアス
- 本日の講義では1のバイアスは無視
  - 解析方法に起因するバイアス
  - 症例数がそれなりにあれば軽微
- 本質的なのは2のバイアス
  - 研究デザインに起因するバイアス
  - 深刻になりやすい

## バラツキへの対処法

---

- 症例数を増やすことが最も効果的
- 測定における工夫
  - バラツキの少ない評価尺度の選択
  - 複数回測定の平均値 (e.g. 血圧)
  - 測定方法の平準化 (e.g. 中央測定、構造化面接)

## コホート研究における内的妥当性

### 4つの妥当性が必要

- 比較の妥当性 Comparison Validity
  - そもそも比較群はよく似ている集団か？
- 追跡の妥当性 Follow-Up Validity
  - 対象者の選択的な脱落はないか？
- 測定の妥当性 Measurement Validity
  - 様々なデータの測定誤差(観測の不完全性)は大きくないか？
- 解析の妥当性 Specification Validity
  - 統計モデルの現実からの乖離は大きくないか？

## 主なバイアス

---

- 交絡 confounding
- 選択バイアス selection bias
- 情報バイアス information bias

## 比較の妥当性 comparison validity

- 曝露群と非曝露群は源泉集団から適切に選択されているか？
  - 選択バイアスがないか？
- 曝露群と非曝露群はよく似た集団か？
  - 交絡がないか？



## 選択バイアス selection bias

---

- 「研究に参加した人と参加しなかった人の間で、曝露-疾病間の関係が異なる」ことから生じるバイアス
  - 例: 健康な人ほど研究に参加しやすい
    - 健康労働者効果 healthy worker effect
    - 健康な人ほど健康に興味を持ち調査票研究に協力

## 外的妥当性としての「選択」ではない

- よくある誤解
  - 「標本が母集団を代表していなければ選択バイアスがある」
  - これは選択バイアスではなく外的妥当性(一般化可能性)の問題
- 内的妥当性としての「選択」
  - 研究計画上の選択基準に合致する集団に対して妥当な(バイアスのない)研究が行えているか
  - 例: 調査票研究
    - 選択基準に該当するのは調査票送付者全員(e.g. ある地域)
    - 回収率100%であれば選択バイアスはない
    - しかし外的妥当性の問題はあるかもしれない

## 選択バイアスの制御

- 対象者の選択基準を
  - コホート研究ではアウトカムと無関係に設定する
  - ケースコントロール研究では曝露と無関係に設定する
- 選択バイアスを引き起こす因子が測定できれば解析による調整が可能な場合もある
  - e.g. 調査票研究において調査に協力的な人の特徴
  - この場合の選択バイアスは交絡と同じこと
    - 選択バイアスと交絡の区別は時にあいまい

## 比較の妥当性 comparison validity

---

- 曝露群と非曝露群は源泉集団から適切に選択されているか？
  - 選択バイアスがないか？
- 曝露群と非曝露群はよく似た集団か？
  - 交絡がないか？

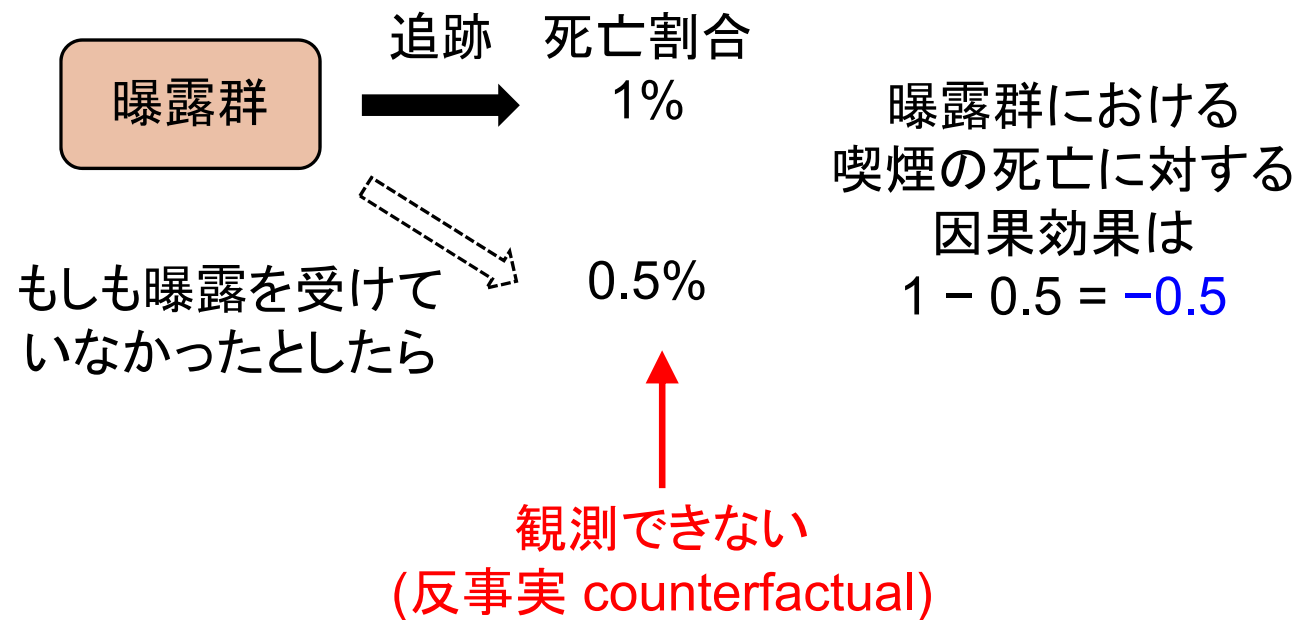
## 交絡を正しく理解するために

---

- そもそも医学研究で観察したい真値とは何か
  - 曝露がアウトカムに及ぼす効果の大きさ
- 理想の対照と因果効果を理解する必要

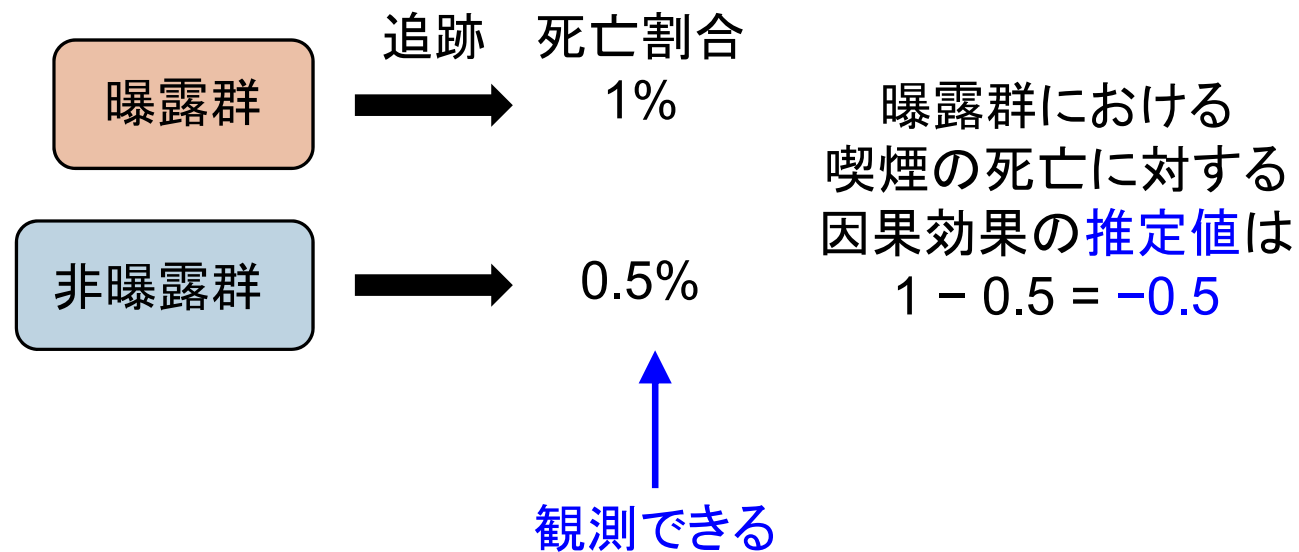
## 因果効果 causal effect

- 喫煙の全死亡に対するリスクの検討



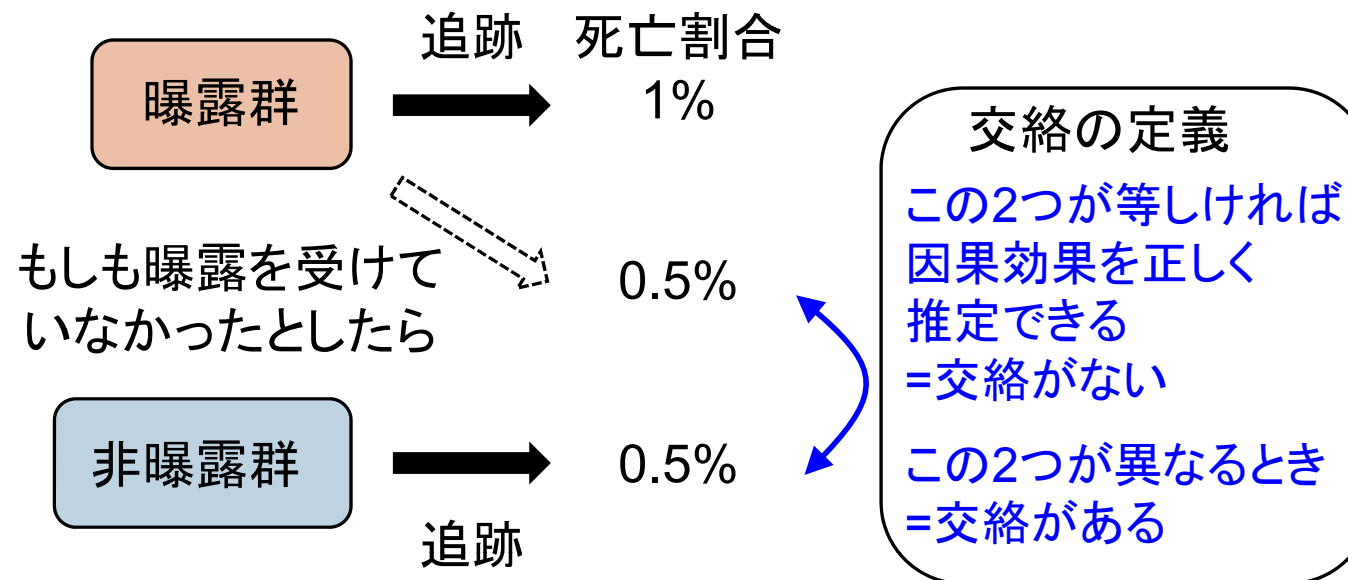
## 因果効果の推定

- 喫煙の全死亡に対するリスクの検討



# 交絡 confounding

- 喫煙の全死亡に対するリスクの検討





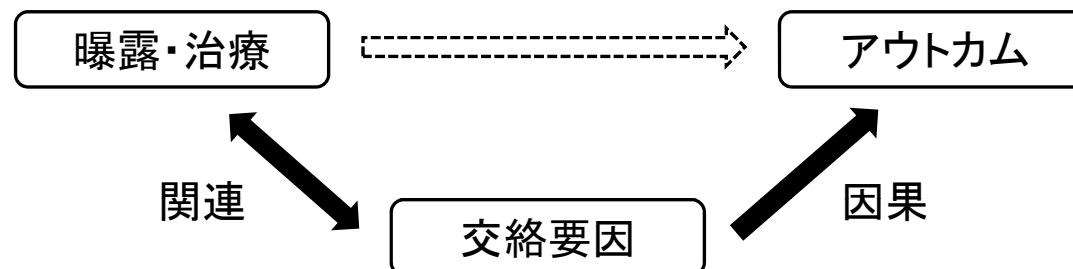
## 交絡の必要条件

---

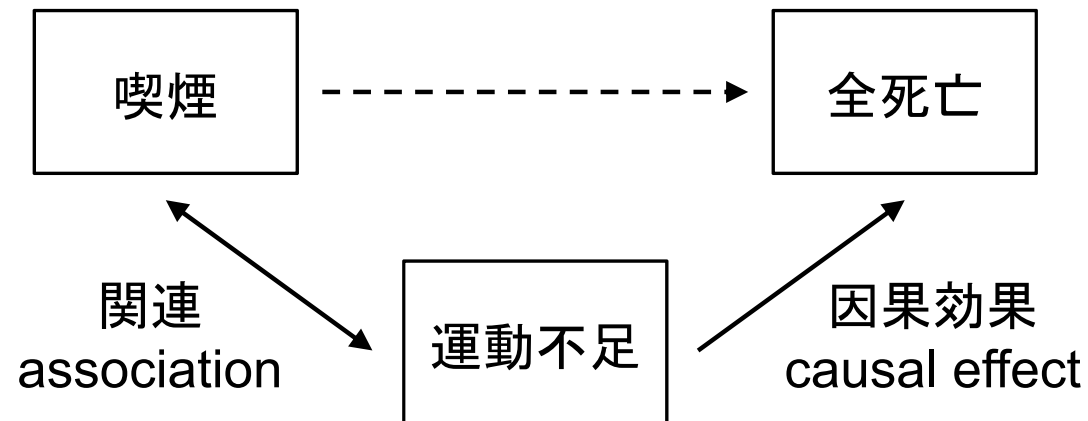
- どちらの集団でも曝露を受けていないにもかかわらず、結果が異なる
- 結果に影響を与える因子(リスク因子)が、曝露群と非曝露群で異なっている

## 交絡因子 confounding factor

- ある因子が交絡因子 (交絡を引き起こす変数) であるためには、
  1. 結果に因果的に影響している
  2. 比較する集団間で、その分布がことなっている (不均衡)
  3. 曝露-アウトカム間の因果連鎖の中間変数でない



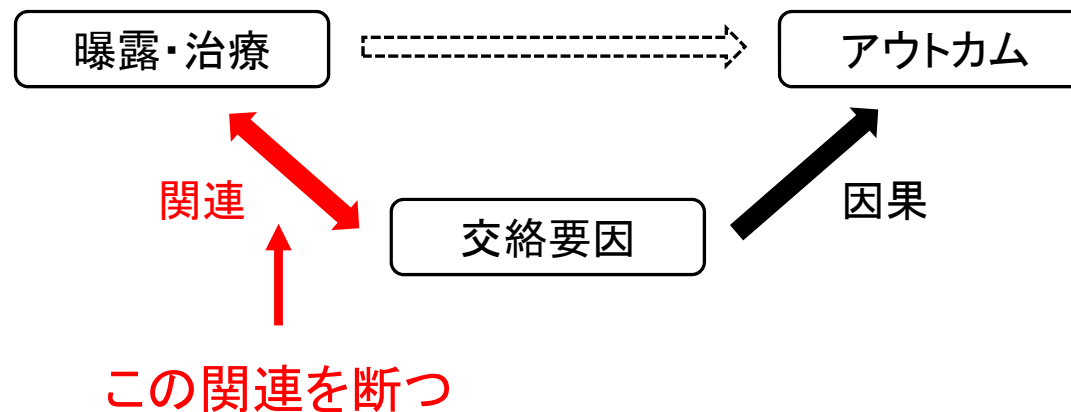
## 喫煙者ほど運動不足



- 交絡の影響
  - 喫煙に因果効果があっても、あるように見えてしまう
  - 喫煙の因果効果がある場合に、過大に見えてしまう

## 交絡を制御するためには

1. 結果に因果的に影響している
2. 比較する集団間で、その分布がことなっている  
(不均衡)
3. 曝露-アウトカム間の因果連鎖の中間変数でない



## 交絡の制御・調整

---

- デザインで
  - 集団の限定 restriction
  - マッチング matching
  - ランダム化 randomization
- 解析で
  - 層別解析 stratified analysis
  - 回帰モデル regression model

## 集団の限定 restriction

---

- 交絡因子が分かっている場合に、ある特定の水準の集団に対象を限定する
  - 例: 性別が交絡因子であることが分かっている場合に、男性だけを対象にする
- 限界
  - 多くの交絡因子がある場合
    - 対象者数が減る
    - 結果の一般化可能性がなくなる

## マッチング matching

---

- 交絡因子の分布が等しくなるように集団を選択する
  - 例: 男女の割合が等しくなるようにする  
(分布が等しいので交絡はない)
- 限界
  - 追跡で不均衡が生じるとせっかくのバランスが崩れる  
(選択バイアス)
  - 多数の交絡因子でマッチしようとする相手が見つからない

# ランダム化 randomization

---

- 交絡を制御する最も強力な手段
  - 測定された交絡因子だけでなく、未知・未測定の交絡因子についても対処できる
- 限界
  - 症例数が少ない研究では偶然による交絡因子の不均衡が生じる
  - 常に、ランダム化が可能とは限らない



## 解析による調整

---

- 回帰分析、ロジスティック回帰、Cox回帰...
  - 測定された交絡因子のみ調整可能
- 調整したとしても交絡を完全には取り除けない
  - 残差交絡 residual confounding

## 交絡の制御・調整

---

- 制御・調整できる交絡因子はランダム化を除けば測定された因子のみ
  - 未測定 of 交絡因子の影響は制御・調整できない
- 研究計画の重要性
  - 重要な交絡因子を見落とさないように
  - 過去の研究結果を参考

## 追跡の妥当性 follow-up validity

---

- 対象者の選択的な脱落はないか？
  - 選択バイアスがないか？

## 打ち切り censoring, 脱落 drop-out

- 対象者の一部が途中で追跡できなくなる
  - 打ち切り例のアウトカムは観察できない
- 打ち切りが
  - アウトカムと独立であればバイアスなし
    - 偶然の事故による打ち切り
    - 研究終了時の打ち切り
  - アウトカムと関連していたらバイアスあり (選択バイアスの1つ)
    - 喫煙と肺がんの関係に興味があるとして、脳卒中や心筋梗塞による死亡

## 追跡の妥当性を高めるために

- 対象者の追跡は最後まできちんと行う
  - 治療を中止してもイベントの発生だけは確認
- 打ち切りの理由を検討
- イベントの定義の変更
  - ランダムでない打ち切りと思われる対象者に関しては、イベントとみなして解析
  - 通常解析との結果の整合性の検討

## 測定の妥当性 measurement validity

- 曝露、アウトカム、交絡因子のそれぞれの測定誤差は大きくないか？
  - 情報バイアスがないか？

## 測定誤差 measurement error

- 曝露, アウトカム, 交絡因子それぞれに対する真値からのズレ
  - 誤差的なズレも測定誤差に含まれる
    - よって、あらゆるデータに測定誤差は生じる
  - 離散的な変数に対する測定誤差は誤分類 (misclassification)とも呼ぶ
    - e.g. 曝露群に分類されるべき対象者が非曝露群に分類
- ほとんどすべての研究で起こる
- ほとんどすべてバイアスを起こす

## 測定誤差を起こす原因

- 測定手法・手順の誤りに起因するもの
  - 測定の誤り、計測限界
    - ・ e.g. 測定機器の誤り、検出バイアス、思い出しバイアス
- 代替変数に起因するもの
  - 測りたい変数の代わりのものを測る
    - ・ e.g. 質問紙票
- 変数の定義に起因するもの
  - 変数の定義があいまい
    - ・ e.g. 「正常血圧」...正常とは?



## 測定の妥当性を高めるために

- 測定における誤差的バラツキは仕方ない
  - 機器の較正+測定者の訓練で最小に
- 測定の手順書の作成
- 定義があいまいな変数は使用しない
- 盲検化
  - 測定における系統的なズレを誤差的バラツキに変える

## 解析の妥当性 specification validity

- 例えば回帰分析
  - $y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \beta_3 \cdot x_3 + \dots$ 
    - y: アウトカム, x: 説明変数,  $\beta$ : 切片・回帰係数
  - 直線性の仮定が正しいか？
  - 因子の効果の加法性は正しいか？
- 解析上の仮定が真実から乖離するほどバイアスが生じる
- しかし、「適切な仮定」はデータからも検証不能

## まとめ

---

- バラツキとバイアスをできるだけ小さく
- バラツキを小さくするために
  - 十分な症例数, 測定方法の工夫
- バイアスを小さくするために
  - 比較の妥当性
  - 追跡の妥当性
  - 測定の妥当性
  - 解析の妥当性