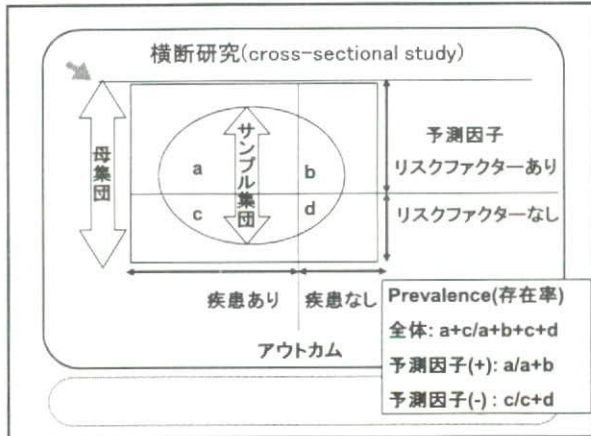


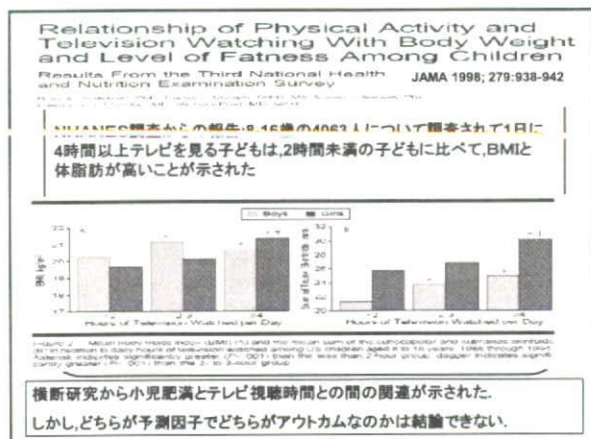
### 本日の内容

- 横断研究 cross-sectional study
- ケースコントロール研究 case-control study
- まとめ



### 横断研究の特徴

- 測定が一時点で行われ、経過を追う期間が無い  
⇒コホート研究との大きな違い
- 調査対象となる因子とその分布パターンを見る  
⇒アメリカにおけるNHANES(国民健康栄養調査) 各実施年の米国民の健康や習慣について重要な情報
- 種々の因子間の相互関係を調べることができる
- Prevalence(存在割合, 存在率)が得られる  
⇒一般的に、割合(proportion)は、単位がなく1を超えない(分子が分母に含まれる)が、率(rate)は単位量あたりの現象が起こる回数や度数で、単位があっても1を超えても良い。慣習上、存在率と記述される



### 横断研究の利点

- 最大の利点は、アウトカムの発生を待たなくてもよい  
⇒コストが安い理由が大きい  
⇒コホート研究や実験的研究の第1段階として行える
- 因子間の相互関係を調べることが出来る  
- 研究者が仮説を立てて、自由に予測因子とアウトカムを決めることで、その因子の間の相互関係を導くこともできる  
- どの因子を予測因子として、どの因子をアウトカムとするかは実際は難しい。他の影響を受けることがない体質的な因子(年齢や人種など)は予測因子によく用いられる

### 横断研究の欠点

- 因子間の因果関係を導き出すことは難しい
  - テレビを見ることにより子どもが肥満になっていくのか、肥満児はテレビを見るのが好きなのかはわからない

```

    graph TD
      A[テレビ] --> B[肥満]
      C[その他の因子] --> B
      A <--> C
  
```

- 一般集団における稀な疾患の存在率を示すには向いていない
- 存在率(prevalence)は分かるが、発生率(incidence)が分からない。疾患の予後、自然史、病因などの正確な情報が得られない。
- 存在率(存在割合)=発生率×罹病期間

### 連続横断研究 serial prevalence study

- 時代経過に伴う1つの集団の変化を知るのに良い方法である。初期の調査による学習効果が、後の調査に影響を及ぼすことの認識は必要。
- 例えば
  - 国勢調査を用いて、10年ごとの年齢別構成の変化を調べる
  - 保険などのデータを用いて、抗不安薬の処方される推移を調べる

などの研究がある。同一の集団を追跡調査しているわけではないので、コホート研究とは異なる

### ケースコントロール研究

過去 ← 現在

患者のサンプル (ケース群) ← 患者母集団

健康人のサンプル (コントロール群) ← 健康人母集団

患者での予測因子(+)  
健康人での予測因子(+)  
+は予測因子あり  
-は予測因子なし

オッズ比=ad/bc  
患者での予測因子(+)  
健康人での予測因子(+)  
オッズ比=a/c  
オッズ比=b/d

### ケースコントロール研究の特徴

- 過去にさかのぼって、色々なリスク因子の状況を比較する
- 費用と時間がかからない
- 対象疾患とある予測因子との関連の強さを、オッズ比という形で推定できるが、存在率(prevalence)や発生率(incidence)は不明である(オッズ比は、アウトカムが稀な場合にはリスク比の近似値となる)

表中の括弧内の数字は、例としての数字を示す

	ケース群	コントロール群(自由)	合計(コントロール群に依存)
予測因子なし	a (200)	b (100,1000と設定)	a+b (300,1200)
予測因子なし	c (800)	d (900,9000と設定)	c+d (1700,9800)
合計	a+c (1000)	b+d (1000, 10000と設定)	a+b+c+d (2000,11000)

存在率a+c/a+b+c+dはコントロールの数の決め方で変化する(1/2と1/11)  
オッズ比(ad/bc)は、コントロールの決め方によらない(2.25と2.25)

### サリドマイドのアザラシ奇形に関する有名な勘違い

- アザラシ奇形の子ども112人、奇形の無い子ども188人
- オッズ比は90/22 ÷ 2/188=380.5 (Lenz, 1962)
- 杉山(日本医事新報1969)の批判「非服用群の奇形発生率 22/208=0.106が高すぎるので Lenzのデータは信用できない」
- 吉村(1969)の指摘「コントロールの数を1000倍に増やせば、非服用群の奇形発生率 22/186022=0.00011, サリドマイドの服用率112/188112=0.00060となる。発生率はコントロールの設定の仕方により決定されるので、そのような議論は意味がない」

サリドマイド	アザラシ奇形		合計
	あり	なし	
服用	90	2	92
非服用	22	186	208
合計	112	188	300

サリドマイド	アザラシ奇形		合計
	あり	なし	
服用	90	2000	2090
非服用	22	186000	186022
合計	112	188000	188112

### ケースコントロール研究の利点(1)

- 稀なアウトカムに対して有効である
  - ⇒小児科には稀な疾患が多い!!(アスピリンとライ症候群、タミフルと異常行動など)
- 例)割礼と陰茎癌の発症に関する架空の研究を考える
  - 「陰茎癌の生涯累積発症率0.16%」,「リスク比が50と仮定」
  - 1) コホート研究⇒「両群の比が等しい」「80%の検出力」⇒6000人以上必要
  - 2) ランダム化比較試験⇒6000人以上必要以上に、陰茎癌の発症の平均年齢67歳のため時間がかかる
  - 3) ケースコントロール研究⇒両群16人でよい

時事通信12月16日のニュース 産業医研修会での矢野の指摘(2008)  
 2007/12/16-19:45 タミフル服用者6割=インフルエンザで異常行動  
 -今季新たに3例報告・厚生省部会

インフルエンザ治療薬「タミフル」と異常行動との関連を臨床面から検討している厚生労働省の専門家作業部会が16日開かれ、インフルエンザ罹患(りかん)時の異常行動137例の分析結果が報告された。うちタミフルを服用していた人は6割だった。

部会の内山真・日大教授は「インフルエンザで異常行動が起きることがはつきりした。タミフルが直接起こす可能性はかなり小さいが、リスクを上げるかどうかはまだわからない」としている。

調査はすべての医療機関を対象に、昨シーズンにインフルエンザと診断され、突然走りだしたり飛び降りたりした事例を収集。350例報告され、30歳未満の重度な異常行動は137例だった。

タミフル服用ありは82人(60%)、なしは52人(38%)、不明が3人(2%)。異常行動は寝ぼけの状態で起きることが多く、睡眠との関連を調べたところ、目覚めて直ちに起こった例が約半数で、タミフル服用の有無で差はなかった。

今シーズンについてもデータ収集・分析を続ける予定。

同日の部会ではまた、10月1日以降の異常行動が、タミフル使用者で3例、リレンザ使用者で2例起きていることも報告された。突然ベランダに向かって走りだしたり、飛び降りようとしたりした例で、いずれも回復した。

上記報道は本勉強会は院内のみの勉強会ということで時事ドットコムよりカットアンドペーストしました

リスクを上げるかどうかはまだわからない?  
 ケースコントロール研究はしていない?

タミフル服用歴 (危険因子)	異常行動 (アウトカム)	
	異常行動あり	異常行動なし
服用あり	82	?
服用なし	52	?
合計	134	?

実際には、別の分担研究の報告で、前年度のオッズ比が検討されており、「今年度の結果はまだわからない」という意味であろう。

アスピリンとライ症候群(Hurwitz ES, et al. NEJM 1985;313:849-57, Hurwitz ES, et al. JAMA 1987;257:1905-11)

- ケース群と4つのタイプのコントロール群(ケースと同じ救急室、ケースと同じ病院の入院中の患者、ケースと同じ学校または保育園、ケースと同じ居住地域)とを比較した

	危険因子曝露によるライ症候群が起こるオッズ比				
	救急室	入院	学校	地域	すべて
サリチル酸	(4.9) 39	(7.8) 66	(4.4) 33	(5.9) 44	(5.8) 40
アセトアミノフェン	0.04 [0.18]	0.13 [0.51]	0.04 [0.16]	0.08 [0.30]	0.06 [0.18]

表中の[]内は95%信頼区間の下限, []内は95%信頼区間の上限を示す  
 サンプリングバイアスの対策として複数のコントロール群が設定された

ケースコントロール研究の利点(2)

- 新しい仮説を立てるのが容易である
- 一度に数多くの予測因子と疾患の関連を調べることが出来る

(例)ハイチの子どもたちにおける急性腎不全の流行に関するケースコントロール研究  
 (O'Brien KL, et al. JAMA. 1998; 279: 1175-80)  
 地域のアセトアミノフェン使用に対する急性腎不全発症のオッズ比52.7 (95% CI, 15.2-197.2)  
 ⇒アセトアミノフェン製造時に使用するグリセリン溶液を汚染したジエチレングリコールの毒性によることがわかった

ケースコントロール研究の欠点

- 発生率,存在率,リスク差などを直接予測することが出来ない (アウトカムが稀な場合はオッズ比がリスク比の近似値になる)
- 危険因子ある時のリスクは,  $a/(a+b) \approx a/b$  (if  $a \ll b$ ),
- 危険因子ない時のリスクは,  $c/(c+d) \approx c/d$  (if  $c \ll d$ )
- リスク比  $\approx (a/b) \div (c/d) = ad/bc$  オッズ比  $= (a/c) \div (b/d) = ad/bc$

	疾患あり	疾患なし	合計
危険因子あり	a	b	a+b
危険因子なし	c	d	c+d

- 1度に扱えるアウトカムの数は1つに限られる
- 結果にバイアスが入り込みやすい
  - 選択バイアス (サンプリングバイアス)
  - 測定バイアス

選択バイアス

- 多くのケースコントロール研究はケースから始まる
- ツェーヘの選択
  - 理想的には, 目的母集団の, 対象疾患を持つ全員
  - 理想的には, 対象疾患を持つ全員からのランダムサンプル
  - 実際は, 研究者の施設を受診した患者さん
- コントロールの選択
  - 対象疾患に罹患していないという点以外ではケースと共通性の高い人, 同じ背景の集団
    - 1) 院内コントロール群
    - 2) マッチングを行う
    - 3) 集団サンプルを用いる
    - 4) 複数のコントロール群

### 測定バイアス

- 思い出しバイアス (recall bias)
  - ケースの方が曝露情報を詳細に記憶している
  - 対処法
    - ・疾患が生じる以前に記録されたデータを用いる
    - ・blinding  
調査者がケースかコントロールかを知らない  
対象者がどれがリスク因子かを知らない  
(ダミーの質問を入れるなど)  
実際は困難ことが多い
- 確認バイアス (ascertain bias)
  - 曝露している人の方が頻繁に検査を受ける
  - 発症の発見や研究への参加率が高い

### まとめ

- 横断研究
  - 全ての因子の観察が一時点
  - 疾患やリスク因子の分布の記述
  - 因果関係はとらえられない
  - 実験研究・コホート研究の事前調査にもなりうる
- ケースコントロール研究
  - 後ろ向き研究
  - 種な疾患に向いている
  - エビデンスはコホートよりも低い
- 結果の指標
  - コホート研究: リスク比, リスク差
  - 横断研究: 存在率
  - ケースコントロール研究: オッズ比

## 臨床研究センター勉強会 生物統計概論

2008/3/18  
国立成育医療センター 臨床研究センター  
斉藤真梨

## 解析作業

- データの要約(、母集団に関する推測)
  - 平均、標準偏差
  - 中央値、四分位範囲、範囲
  - 頻度、割合
- 群間の違いの定量化(推定)
  - 生の値の差、比
  - リスク差、リスク比、オッズ比
  - ハザード比
  - 他の要因を調整した解析(層別解析と回帰モデル)
- 群間の違いに関する決定(検定)
  - t検定、ウィルコクソン検定
  - $\chi^2$ 検定、Fisherの直接確率検定
  - ログランク検定、ウィルコクソン検定

2

## データの要約:具体例

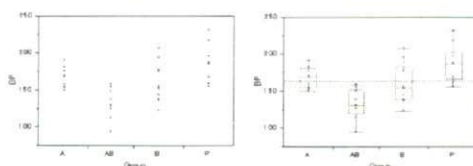
- 実験
  - 39匹のイヌを4群にわけ、降圧剤あるいはプラセボを与え、収縮期血圧を測定した
  - 降圧剤を与えた群で血圧が下がっているようにみえる
  - 個体差が大きい
  - このまま報告しても議論しづらい

対照	187	160	197	232	218	186	168	156	188	187
A薬	168	170	181	176	182	154	191	155	159	150
B薬	195	207	178	137	176	144	123	138	156	153
A+B併用	129	138	159	149	155	130	113	126	94	

3

## プロット

- 視覚的に4群の分布を確認
- 左図:横軸に群、縦軸に血圧、右図:箱ヒゲ図
  - 四分位範囲の箱と箱の高さ×1.5以内の最大値までのヒゲ
  - どの群も外れ値はなく、左右の非対称性もあまりない
  - ばらつき具合も4群で違いはなさそう(等分散性)



4

## 要約統計量

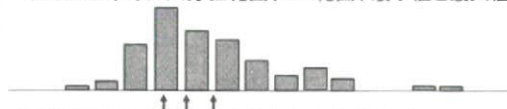
- 平均(Mean)と標準偏差(Standard Deviation)
  - SDのかわりに分散(Variance)でもいい
  - 39個のデータが8個に要約された
    - 対照群に比べA群で-19、B群は-27、併用群で-55
    - 効果が加法的

	平均	SD
対照	188	24
A薬	169	14
B薬	161	27
A+B併用	133	21

5

## 平均・SDが不適切な場合

- 左右に中心がずれている場合
  - 左にゆがんでいる、右に裾を引くなど
  - 平均が分布を適切に表現しているとは限らない
  - 例)「日本の平均貯蓄高1世帯当1309万円、景気低迷影響か」(1998年新聞見出し)
- 平均のかわりに中央値(Median)、最頻値(Mode)
  - 裾の長い方からmean, median, mode
- SDのかわりに四分位範囲、95%範囲、最小値と最大値



6

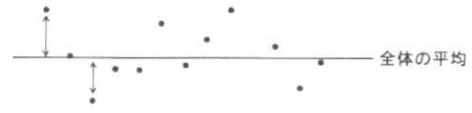
## 要約統計量の計算

- 分布の位置(中心)を示す3指標
  - 平均: 総和/サンプル数
  - 中央値: 小さい順に並べた真ん中(50%)の値
  - 最頻値: もっとも頻度が高かった値
- ばらつきの指標
  - 分散、標準偏差(次のスライド)
  - 四分位範囲(四分位偏差): 小さい順に並べたときの25%点~75%点に含まれる範囲
  - 95%範囲: 同じく2.5%~97.5%に含まれる区間 など
- 分布の形に合わせて選択
  - 左右対称なら平均とSD、ゆがんでいれば中央値と範囲
  - どちらも1峰性の場合のみ

7

## ばらつきの指標

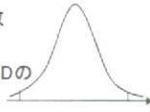
- 分散 (Variance) =  $\sum(\text{観測値} - \text{平均})^2 / (\text{サンプル数} - 1)$ 
  - 観測値の平均からのずれを2乗して、全サンプル分合計し、平均をとったもの
  - 2乗しないと全体で0になってしまう



8

## ばらつきの指標

- 標準偏差 (Standard Deviation) =  $\sqrt{\text{分散}}$ 
  - 平均と同じ次元の、ばらつきの指標
  - データが左右対称で1峰性なら平均 $\pm 2SD$ のなかにおよそ95%の観測値が含まれる
- 標準誤差 (Standard Error) =  $SD / \sqrt{\text{サンプル数}}$ 
  - 推定値(ここでは平均)のばらつき(推定精度)の指標
    - 10匹の犬の実験を10回繰り返した、10個の結果
    - 100匹の犬の実験を10回繰り返した、10個の結果
    - 10個の結果のばらつきは100匹の実験のほうが少ない
    - サンプル数が多いほど平均値の推定精度が増す



9

## SDとSEの使い分け

- SD
  - 生データのばらつきの指標
  - 研究対象集団について記述するとき
- SE
  - SEは推定値のばらつきの指標
  - 平均値を、源泉集団の平均の推定値として用いるとき
  - 群間の差、比やリスクの大きさを提示するとき



10

## 2変数以上の解析

11

## 統計手法の選択

- 目的
  - 要約
  - 群間比較
  - 他の因子の調整
- アウトカムの変数のタイプ
  - 連続変数、カテゴリ・2値変数、時間
- 研究の種類
  - ランダム化、非ランダム化
  - 介入、観察
  - 前向き、後ろ向き

12

## 目的×変数のタイプ×統計手法

アウトカムが	2値	連続変数	生存時間
要約 (分布の記述)	頻度集計 分割表	ヒストグラム 平均・標準偏差 相関	生存曲線 Kaplan-Meier法
単純な 群間比較	$\chi^2$ 検定 リスク推定	t検定 Wilcoxon検定 平均値の差の 推定	Log-rank検定 率比の推定
層別解析	Mantel- Haenszel法 標準化	分散分析	層別Log-rank
回帰モデル	Logistic回帰	分散分析 重回帰	Cox回帰

## 研究の種類×統計手法

- 研究デザインとともに例で紹介
  - 介入研究(臨床試験)
    - 単群試験
    - 複数群(ランダム化)比較試験
  - 観察研究
    - 追跡研究(コホート研究)
    - 断面研究
      - ケースコントロール研究
      - 断面研究(=横断研究、クロスセクショナル研究)

## 臨床試験

- 特定の集団にプロトコル(研究計画書)に定めた介入を行い、結果を観察する
- 目的
  - ある介入の成績(有効性・安全性・経済性)を評価する
    - 単群試験:1種類の介入を行い、先行研究と比較
    - 複数群の試験:複数の介入を行い、研究内で比較(群間比較)
- 統計的キーワード
  - ランダム化
  - 有意水準、p値
  - $\alpha$ エラーと $\beta$ エラー

## 具体例1

- デザイン
  - ランダム化比較試験
  - 重症喘息患者を100人ずつの2群に分け、それぞれ異なる吸入薬(A薬・B薬)を投与
  - 治療後の経過を比較する
- アウトカム
  - ①治療開始6時間後の呼吸機能スコア(0~100点)
  - ②酸素終了までの期間
  - ③有害事象発現
- 特徴
  - ランダム化しているので群間の背景因子がほぼ同じ
  - 異なるのは受けた治療だけと考える
- 結果の違いは治療の違い由来

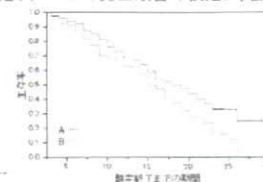
## 治療開始6時間後の呼吸機能スコア

- プロット
  - ばらつき幅
  - 左右の傾きの
  - 密度の濃い区間(ピーク)の位置と数
- 要約統計量
  - A群の平均 60.1 (95%信頼区間55.6-64.6)
  - B群の平均 46.1点 (95%CI: 41.6-50.7)
- 群間差とt検定
  - 群間差14.0点、 $P<0.0001$
  - 2群は有意に14点異なる
- 注意
  - 両群、あるいは一方が0や100に集まってしまったら(床効果/天井効果) Wilcoxon検定



## 酸素終了までの時間

- 生存曲線(Kaplan-Meier曲線)
- 要約統計量
  - A群:中央値(呼吸器科)15時間(四分位範囲 9~20時間)
  - B群:中央値15時間(四分位範囲 9~20時間)
- ハザード比(イベントの発生するスピードの比):B薬の方が1.45倍酸素終了が早い
- Log-Rank検定(イベントの発生順番の検定):p値<0.0001



## 有害事象発現

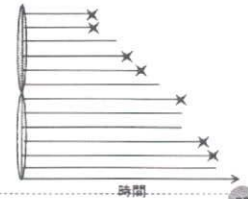
- 分割表
- 要約統計量
  - A群の発現人数15人(15%)、B群の発現人数8人(8.2%)
- リスク差7%、リスク比1.88倍
- $\chi^2$ 検定 p値=0.12
- 注意
  - 分割表の期待値の一つ以上が5以下になるほど小サンプルの場合はFisherの直接確率検定を用いる

	有害事象あり	有害事象なし	合計
A薬群	15 (15.0%)	85 (85.0%)	100
B薬群	8 (8.2%)	90 (91.8%)	98

19

## 追跡研究(コホート研究)

- 特定の集団を経時的に観察し、変化や疾患の発症を記録する
- 目的
  - 介入が倫理的、物理的に困難なリスク因子の評価
  - 状況が多様すぎて介入方法、リスク因子を統一的に扱えない場合
- キーワード
  - 前向き
  - 追跡
  - 交絡
  - 脱落・追跡不能



20

## 具体例2

- デザイン
  - コホート研究
  - 全国700の医療機関でインフルエンザと診断された18歳未満の患者10316人を診断後1週間まで観察
- アウトカム
  - ①薬Aの内服有無
  - ②異常行動・危険行動有無
  - ③発熱パターン
  - ④薬Aと回復までの期間
  - ⑤薬Aと異常行動の関連
- 特徴
  - 薬A有無はランダム化されていない
  - 他にも色々な介入が入る
  - 内服薬と異常行動の前後関係の情報は必須

21

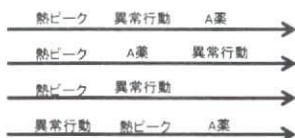
## A薬のリスクの解析

- A薬内服有無でピーク時体温が違うか
  - 群間の差とt検定
- A薬内服有無で回復までの期間が異なるか
  - 回復までの期間の差、ハザード比、Log-Rank検定
- A薬内服有無で異常行動出現頻度が異なるか
  - リスク差、リスク比と $\chi^2$ 検定(人数が少なければFisher検定)

22

## 実際データを集めたら

- リスク因子のタイミングとイベントのタイミングが色々
  - 診断、発熱、内服、異常行動、その他の治療が患者の受診タイミングや主治医の方針によりばらばら
  - 診断、発熱などは気づくタイミングが一様でない (detection bias)



23

## リアルな結果

- 研究対象10316人中
  - A薬服用者7870人
    - 服用後に幻覚、幻聴などの異常行動が700人に出現
      - 飛び降りなどの事故につながる危険行動は22人
    - 服用前に異常行動が出た人は285人、危険行動は9人
  - A薬を全く服用しなかった対象者
    - 異常行動が546人、危険行動が16人

	異常行動	危険行動	合計人数
A薬内服前	285 (3.8%)	9 (0.11%)	7870
A薬内服後	700 (8.9%)	22 (0.28%)	7870
内服なし	546 (22.3%)	16 (0.63%)	2446
	—	—	10316

<http://www.bm.mhlw.go.jp/shing/2008/01/dl/d0131-7g.pdf>

24



## デザインの情報が重要

- アウトカム情報のとり方
  - 危険行動有無、発熱からの日数、診断日からの日数…
  - 異常行動の出現時期との時間関係
    - 行動前、後
    - 発熱後XX日目に内服 など
  - 詳細がわからないと仮説に合う解析方法が分からない
- 他のリスク因子の情報
  - 年齢や熱形、A薬服用までの時期、A薬耐性の有無、予防摂取の有無、他の治療介入など

25

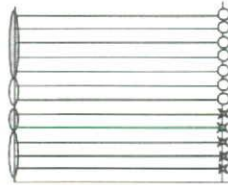
## 他のリスク因子の情報があれば

- 回帰分析
  - $\text{アウトカム(危険行動有無)} = a \times \text{薬剤有無} + b \times \text{年齢} + \dots$
  - 他の因子(年齢)も並列に評価し、互いの影響を除く
- 層別解析
  - 年齢で層別した解析
  - 層(年齢)ごとに評価した薬剤の影響を、最後に併合
- 多くの因子を調整した解析の難しさ
  - 因子間の影響についての仮説、時間関係、交互作用の有無
  - データ数、変数間の共線性など
- 研究の限界
  - 研究で把握できないリスク因子、交絡因子はたくさんある
  - 数多くの因子で調整して適切な解をえるためには多くのサンプル数と情報が必要

26

## 断面研究: ケースコントロール研究

- 疾患を発症した集団と、していない集団を集めて、リスク因子に対する過去の曝露歴を調べる
- 目的
  - 1つの疾患に対する過去の(複数の)リスクの評価
- キーワード
  - 稀な疾患の研究
  - オッズ比
  - 交絡
  - サンプルバイアス



27

## 具体例3

- デザイン
  - ケースコントロール研究
  - A病院に勤務する30~40代の、成人以降に花粉症を発症した人(ケース)と花粉症でない人(コントロール)
  - ケースとコントロールは交絡因子(年齢、性別)でマッチング
- リスク因子
  - ① 幼少期の居住環境
  - ② 学生時代の脂質摂取量(全摂取カロリーに対し0~100)
- 統計的手法
  - ① 分割表とオッズ比
  - ② ロジスティック回帰によるリスクの推定

28

## 居住環境

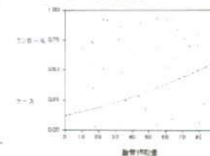
- 分割表
- 要約統計量
  - ケース(花粉症群)の都市部居住38人(38%)
  - コントロール(非花粉症群)の都市部居住155人(52%)
- オッズ比1.74(95%CI: 1.10~2.77)
  - 都市部の方が花粉症リスクが1.7倍高い
- $\chi^2$ 検定  $p=0.018$

	花粉症	花粉症なし
都市部	38	155
郊外	62	145
合計	100	300

29

## 脂質摂取量

- プロット
- 要約統計量
  - ケース(花粉症群)の摂取量の平均は30(SD: 15)
  - コントロール(非花粉症群)の平均は25(SD: 20)
- $\chi^2$ 検定  $p=0.022$
- ロジスティック回帰
  - $\text{Log(花粉症有/無)} = a \times \text{脂質摂取量}(\%) + b \times \text{居住地} + \dots$
  - $a=1.025$ と推定された
  - 脂質摂取量が5%増えると花粉症オッズが1.025<sup>5</sup>倍、約13%増加



30

## デザインがあいまいな報告を解釈する

### ●対象のサンプリング

- K大学で理系3年生の男女に就職したい企業を質問
    - 男子:トヨタ自動車、日立製作所、松下電気、サントリー、ソニー
    - 女子:資生堂、武田薬品、味の素、エーザイ、第一三共
    - ？女子には製菓企業が人気
  - 薬学部は他学部より女子が多い
- ### ●時間関係、他のリスク因子情報
- 会社の健康診断の結果
    - ボーナスが高い人ほど血圧が高かった
    - × 塩分をおおくとればボーナスが増える
  - 前後も不明だし、年齢という交絡因子が存在する

▶ 浜田知久馬著 1999

31

## シンプソンのパラドクス

	受験者	合格者	合格率
A予備校	260人	80人	30.8%
E予備校	70人	24人	34.3%

	理系の学部			文系の学部			
	受験者	合格者	合格率	受験者	合格者	合格率	
A校	200人	50人	25%	A校	60人	30人	50%
E校	20人	4人	20%	E校	50人	20人	40%

▶ 吉村功2006日本科学技術連盟セミナー講義

32

## 観察研究の解釈は要注意

- ランダム化も意図的な介入もしていない
  - リスク因子と疾患発症の間に関連があっても「因果関係がある」とは言いきれない
- 研究の質を高める
  - 研究対象集団の性質の把握
    - 研究対象は真に興味のある集団をよく代表しているか
    - 比較するときは群間で性質が類似しているか
  - リスクへの曝露と疾患発症の時間関係の把握
    - 前後関係は明確か、induction period(誘導期間)は十分か
  - 興味ある疾患とそのリスクに関連する因子の情報
    - リスクと疾患の関係を修飾する因子について調べたか
    - その影響は解析で補正できる程度か

▶

33

## まとめ

- 各統計手法
  - 要約とプロット
    - プロットと箱ヒゲ図、生存曲線(Kaplan-Meier法)、分割表
  - 検定
    - t検定、ログランク検定、 $\chi^2$ 検定
  - 推定
    - 値そのものの差、比
    - リスク比、リスク差、オッズ比
    - ハザード比
    - 回帰分析、層別解析
- 注意
  - 観察研究では交絡因子が多く存在
    - 調べたい仮説に忠実なサンプリングデザイン
    - 他のリスク因子を調べ、解析での調整が必要

▶

34

## 参考

- 浜田知久馬学会・論文発表のための統計学—統計パッケージを誤用しないために。真興交易医学書出版部、1999。
- 折笠秀樹監訳。はじめて学ぶ医療統計学。総合医学社、2003。
- 折笠秀樹著。臨床研究デザイン—医学研究における統計入門。真興交易医学書出版部、1996。
- Douglas G. Altman, Practical Statistics for Medical Research. Chapman&Hall, 1990
  - 木船 義久ら訳。医学研究における実用統計学
- Armitage P, Berry G, Matthews JNS. Statistical Methods in Medical Research 4 Sub. Wiley-Blackwell, 2001
  - 榎ら訳。医学研究のための統計的方法
- Rothman KJ. Epidemiology: An Introduction. Oxford University Press, 2002
  - 矢野栄二ら訳。ロスマンの疫学—科学的思考への誘い

35

## 成育医療臨床研究基盤整備勉強会(平成20(2008)年)

趣旨: 若い医師、生物統計家へ臨床研究の実際を教えるための勉強会である

(臨床研究の方法論、生物統計学、臨床疫学、メディカルライティング、倫理的配慮等)

初年度18年8月からはそれまでの月曜勉強会に金曜も追加された、これは火曜日が研修等と重なることが多くなったために他の時間の確保の必要なことと、フェローから現在の診療科での臨床研究状況を知りたいとの要望に応えるためである

平成19年度からは教科書に基づいた勉強会を、さらに薬物動態・小児薬理の抄読会も開始。臨床研究センター以外からの定期的な参加が増えた。

平成20年度からはそれまでの火曜日から月曜日に変更、外部向けの教科書勉強会と臨床研究センター内部向けのものと対象をわけた。

\*医学的研究のデザインの第11章は省略する

	日時	曜日	場所 会議室番号	演者	(所属)	題名
第1回	2008/6/16	月	12(1F)	前川貴伸	総合診療部	医学研究のデザイン 第9章 Enhancing Causal Inference in Observational Studies
第2回	2008/6/23	月	21(2F)	磯島 豪	内分泌代謝科	進捗状況: ×××による 成長促進治療についての検討
第3回	2008/7/7	月	11(1F)	磯島 豪	内分泌代謝科	医学研究のデザイン 第10章 Designing a Randomized Blinded Trial
第4回	2008/7/14	月	21(2F)	加藤忠明	成育政策科学 研究部長	小児慢性特定疾患のデータベースとその利用
第5回	2008/8/18	月	11(1F)	磯島 豪	内分泌代謝科	医学研究のデザイン 第12章 Designing Studies of Medical Tests
第6回	2008/8/25	月	11(1F)	前川貴伸	総合診療部	医学研究のデザイン 第13章 Utilizing Existing Databases
第7回	2008/9/22	月	12(1F)	佐古まゆみ	腎臓科	進捗状況: リツキシマブの医師主導治験について～計画から実施～進捗状況
検討会	2008/9/24	水	12(1F)	藤野明浩	第2専門診療部	検討会「リンパ×××の研究」
検討会	2008/11/10	月2	21(2F)	入江聖子	母性内科	検討会「×××妊婦の×××に関する検討」
検討会	2008/11/17	月3	12(1F)	津村由紀	アレルギー科	検討会「×××を用いた×××治療に関する研究」
第8回	2008/1/19	月3	21(2F)	斉藤真梨	治験管理室	統計勉強会① 臨床研究デザイン概論とJMPの紹介
第9回	2008/1/26	月4	21(2F)	坂本なほ子	疫学研究室室長	疫学概論
第10回	2008/2/2	月1	21(2F)	斉藤真梨	治験管理室	統計勉強会② 二変量の解析①連続のアウトカム
第11回	2008/2/16	月3	21(2F)	斉藤真梨	治験管理室	統計勉強会③ 二変量の解析②2値のアウトカム
講演	2009/3/9	月2	21(3F)	中川雅生	滋賀医大 治験管理セン ター病院	小児の臨床研究推進のための人材育成と環境整備に向けた取り組みの中でわかったこと
第12回	2008/3/16	月3	21(2F)	斉藤真梨	治験管理室	統計勉強会④ 生存時間解析
第13回	2008/3/23	月4	21(2F)	斉藤真梨	治験管理室	統計勉強会⑤ 層別解析・回帰モデル

## Chapter 9 : Enhancing Causal Interference in Observational Studies

Designing Clinical Research 4<sup>th</sup> Ed.

16<sup>th</sup> June, 2008

National Center for Child Health and Development  
Clinical Research Center  
Takanobu Maekawa

### Contents:

- 関連の種類
- 偽の関連を除外する方法
  - 偶然
  - バイアス
- 真の関連から、さらに因果関係を導く
  - 交絡の予防

### 関連

#### Associations

- 観察研究 (Chapter 8)の結果, AとBに関連が疑われる
- 知りたいこと
- AとBは本当に関連があるのか(関連の真偽)
- AはBの原因なのか(Cause-Effect; 因果関係)

### ある関連が疑われた時 5つの可能性を考える

			A $\nleftrightarrow$ B
			A $\nleftrightarrow$ B
			A $\leftarrow$ B
		Factor X	A $\swarrow$ B
			A $\rightarrow$ B

(TABLE 9.1, p128)

A: コーヒー摂取 B: 心筋梗塞

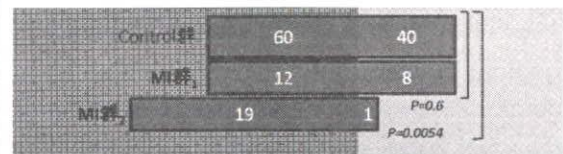
### ある関連が疑われた時 5つの可能性を考える

			A $\nleftrightarrow$ B
			A $\nleftrightarrow$ B
			A $\leftarrow$ B
		Factor X	A $\swarrow$ B
			A $\rightarrow$ B

(TABLE 9.1, p128)

A: コーヒー摂取 B: 心筋梗塞

### 「偶然」による偽の関連



コーヒー摂取習慣あり(60%) 習慣なし(40%)

- コーヒー摂取と心筋梗塞に関連がない場合を仮定すると...
- 心筋梗塞群におけるコーヒー摂取率は全体集団と同一(60%)であることが期待される(MI群)
- だが、偶然この期待通りにサンプリングされないこともある(MI群)
- 偶然により生じる期待値とのズレを偶然誤差(Random error)という。
- 偶然誤差が統計的に有意である場合、Type 1 error ( $\alpha$ -error)という。

## 「バイアス」による偽の関連

- Research Question
  - 研究者が知りたいと思う問題
  - 理想
- Study Plan
  - 実施可能性を考えて妥協して作った研究計画
  - 作り出した現実
- バイアス
  - 理想と現実のズレ
  - 研究者の恣意や問題のある測定方法、予想しない原因により、特定の方向に結果が歪められる。

## Research Question & Study plan

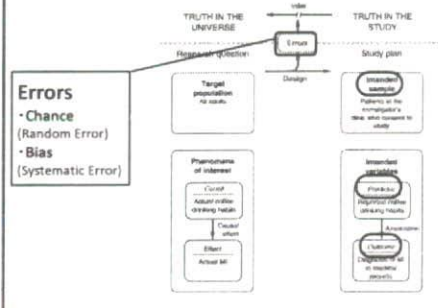


FIGURE 8.1. Misleading flow by comparing the research question and the study plan.

## バイアスの原因

- 主な3つ
  - 対象集団の代表性
  - 測定する予測因子は真に興味あるリスクを捉えているか
  - 測定する結果因子は真に興味ある状態を捉えているか
- コーヒーの例では
  - 対象が心筋梗塞以外の入院患者
    - コーヒー摂取量が一般成人より少ないかも知れない
    - 他の病因による心筋梗塞のリスク
  - 測定方法
    - 時期、回数、尺度、...

## 偽の関連の除去には

- デザイン段階での対処が重要
- 偶然による場合
  - サンプルサイズを大きくする
  - 測定精度を高める (Chapter4, Table 4.2)
- バイアスによる場合
  - 3点(対象, 予測因子, 説明因子)を検討する
  - 他の研究との比較
  - 感度解析(解析の段階での工夫)

## ある関連が疑われた時 5つの可能性を考える

			A <del>→</del> B
			A <del>←</del> B
			A ← B
		Factor X	A → B
			A → B

TABLE 9.1, p128

A: コーヒー摂取 B: 心筋梗塞

## 真の関連

### 因果関係を示唆するエビデンス

- 関連の普遍性 (consistency)
- 関連の強さ (strength)
- 量-反応関係 (dose-response)
- 生物学的妥当性 (biologic plausibility)
- 時間的妥当性 (temporal)

### 注意点)

- 交絡因子と結果の関係もまた、一貫性や量反応関係を示す可能性がある。
- 一般的には、交絡因子によって生じた関連は、真の因果関係より関連が弱いのが普通である。

## 結果-原因

原因因子と結果因子が逆

- CRP高値と心筋梗塞
- 携帯電話の使用と交通事故

- 原因と結果の時間的前後関係を考慮する。
- 時間的前後関係が比較的明確なコホート研究では問題となることはすくない。

注意)

- 結果が生じる(発病)までに潜伏期間の長い原因の場合、コホート調査でも前後関係の解釈が難しい場合がある。例えば2型糖尿病と膵臓癌発症の関連。

## 交絡(Confounding)

- 交絡因子(Confounding variable)

A confounding variable is one that is associated with the predictor variable, and a cause of outcome variable.

- 交絡がおきている

- 交絡因子によって関係がゆがめられている状態



## 例えば

心筋梗塞のリスクを調べる観察研究

- コーヒー習慣と心筋梗塞発症に関連が疑われた  
コーヒーは真の原因か？
- 真の原因がないか？
- コーヒーは真の原因と関連のある事象にすぎない  
無実の傍観者(innocent bystander)ではないか？



## 例えば

心筋梗塞のリスクを調べる観察研究

- コーヒー習慣と心筋梗塞発症に関連が疑われた  
コーヒーは真の原因か？
- 喫煙は心筋梗塞のリスクである
- コーヒーと喫煙には高い相関がある
- 喫煙が交絡因子、コーヒーは無実の傍観者

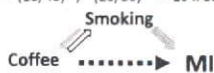


## 例題(Appendix 9A)

	喫煙 あり			
	MIあり		MIあり	
	100	50	50	100

MIに対するコーヒー摂取のオッズ比

- 喫煙者 =  $(80/20) / (40/10) = 80 \times 10 / 20 \times 40 = 1$
- 非喫煙者 =  $(10/40) / (20/80) = 10 \times 80 / 40 \times 20 = 1$

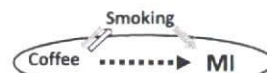


## 例題(Appendix 9A)

	喫煙 あり+			
	MIあり			

MIに対するコーヒー摂取のオッズ比

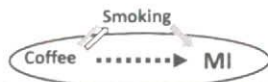
- 喫煙者 + 非喫煙者 =  $(A / B) / (C / D)$



### 例題(Appendix 9A)

	喫煙 あり+	
	MIあり	
	150	150

MIに対するコーヒー摂取のオッズ比  
 ・ 喫煙者+非喫煙者 =  $(90/60) / (60/90) = 90 \times 90 / 60 \times 60 = 2.25$



### 例題(Appendix 9A)

	喫煙 あり			
	MIあり		MIあり	
	A	B	C	D

MIに対する喫煙のオッズ比  
 ・  $(A/B) / (C/D)$



	MI(+)	
喫煙(+)		

### 例題(Appendix 9A)

	喫煙 あり			
	MIあり		MIあり	
	100	50	50	100

MIに対する喫煙のオッズ比  
 ・  $(100/50) / (50/100) = 100 \times 100 / 50 \times 50 = 4$



	MI(+)	
喫煙(+)		

### 例題(Appendix 9A)

	喫煙 あり			
	MIあり		MIあり	
	150		150	

コーヒー習慣の喫煙のオッズ比  
 ・  $(120/30) / (30/120) = 120 \times 120 / 30 \times 30 = 16$



	喫煙(+)	
Coffee(+)		

### 交絡への対処

- 研究計画(デザイン)
  - 対象者の限定
  - マッチング
- 解析段階
  - 層別解析
  - 回帰分析

### デザインによる対処

- 対象者の限定
  - 均一な背景の集団にだけ焦点をしぼる
  - わかりやすい
  - 研究テーマに的が絞られる
- 一般性は損なわれる
- 場合によっては十分なサンプルの確保が困難

## デザインによる対処

- マッチング
  - ある因子の分布を比較群間でそろえること  
予測因子 + 関連が予測される因子
- ペアマッチング(pairwise matching)
  - コーヒー + 非喫煙 → 非喫煙 Control
  - コーヒー + 喫煙10~20本 → 喫煙10~20本 Control
- 頻度マッチング(frequency matching)
  - (喫煙10~20本) 群20人 → 喫煙10~20本 Control 40人

## 特徴

- マッチング因子による交絡が制御される
- コントロールが選択しやすい
- 未測定因子の影響も制御できる
  - 例: 兄弟マッチングで遺伝要因・環境要因を制御
- マッチング因子の影響を調べられない
- 特殊な解析が必要
- オーバーマッチングによる検出力の低下
  - 群間で、マッチング因子の分布を揃えたために予測因子の分布も似てしまうこと
  - バイアスにはならない

## 解析による対処

- デザイン段階での制御が重要
  - たくさんの因子での限定やマッチングは不可能
  - 研究前にわからないこともある
  - 交絡の強さの程度がわからない
- 主な二つ
  - 層別解析
  - 回帰分析

## 層別解析

- 方法
  - 対象者を属性(層別因子)で分類
  - 層ごとに解析
  - 結果を併合する
    - 人数に応じて重み付けすることが多い
- ≠サブグループ解析
  - 一部の対象者についてのみ解析・結果の提示

## 層別解析の特徴

- わかりやすい
- 層の作り方は任意
- 層別因子はカテゴリ変数として扱う
- 層別因子が多いと層内に対象者が少なくなり、結果が不安定になる
  - 人種・性別・年齢・体格・喫煙・居住環境...
- 測定してある因子しか使えない

## 対処方法の選択

- 「対象者の限定」
  - 興味ある集団が決まっていたら
- 「マッチング」
  - 年齢・人種・性別など他の因子に影響をうけない属性因子について
- 「層別解析」
  - 予定しているときはサンプルは多めに
- 「回帰分析」
  - 補足的に
  - 色々やってみる



## まとめ

- ある事象AとBに関連が疑われた場合、5つの可能性を考える
- 一番知りたいのは因果関係である
- そのために「偶然」「バイアス」「交絡」を除去
  - 偶然:  $\alpha$ エラー、精度とサンプルサイズ
  - バイアス: 研究計画の吟味
  - 交絡: 対象選択と評価方法を吟味する

## まとめ 真の関連

因果関係を示唆するエビデンス

- 関連の普遍性 (consistency)
- 関連の強さ (strength)
- 量-反応関係 (dose-response)
- 生物学的妥当性 (biologic plausibility)
- 時間的妥当性 (temporal)

注意点)

- 交絡因子と結果の関係もまた、一貫性や量反応関係を示す可能性がある。
- 一般的には、交絡因子によって生じた関連は、真の因果関係より関連が弱いのが普通である。

2008年7月7日 臨床研究センター勉強会  
 医学研究のデザイン 第10章  
 実験的研究をデザインする

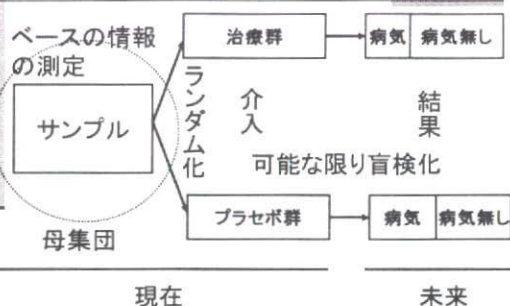
国立成育医療センター  
 内分泌代謝科 臨床研究フェロー  
 磯島 豪

実験的研究では

- 研究者は対象者に介入(治療)を行い,その結果を観察する
- 因果関係の推論が可能である
- 特に,ランダム化と盲検化が大切である
- 費用と時間がかかる
- 一度の研究で証明できる仮説が少ない
- 対象者に,重篤な副作用が生じることがある
- 研究テーマによっては,適切で無いものもある

研究テーマに関する知見が十分成熟した段階になって初めて行う

ランダム化臨床試験の手順(Figure 10.1)



本日の内容

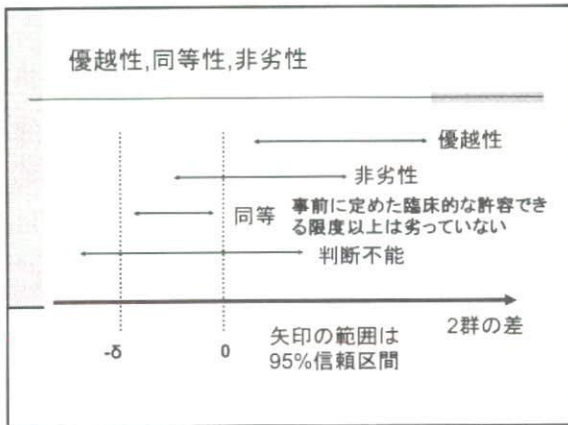
- 介入とコントロール群の選択
- 測定する結果の選択
- 対象者の選択
- ベースラインの変数の測定
- ランダム化と盲検化

介入とコントロール群の選択(1)

- 介入の選択
  - 介入の選択することが,臨床試験計画の第一段階である
  - 効果と安全のバランスは,研究の状況に依存する  
介入を考える時,重症の病気に対しては,効果の方が,軽症の病気に対しては,安全の方が大きな問題となる
  - 一つの介入の方が,複雑な介入よりも結論が明確になる。しかし,医療の現場では単純化が難しい場合もある
  - 単純な介入の方が,実施可能性が高まる

介入とコントロール群の選択(2)

- コントロール群の選択
  - 介入を全く受けず,盲検化されているのが最高のコントロール群である。介入と区別出来ないプラセボを用いる事がそれに相当する。
  - 試験治療の結果のリスクを下げる共介入の存在について注意する必要がある。
  - 標準治療と新しい治療を比較する際には,同等性試験が行われる
  - 対象者がランダム化されることが難しいと予想される有望な外科治療などでは,wait-list controlが使用されることもある。これは,介入の結果がすぐに出る場合に限られる



- 測定する結果の選択
- 測定する結果の選択は,試験デザイン,費用など多くの事に影響する
  - 臨床試験には,種々の結果測定が含まれるが,主要評価項目を1つ決めるべきである,主要評価項目を決めることでサンプルサイズが計算できる
  - 試験の結果に影響する代替指標を用いることも有用である,代替指標には,疑問となった結果を良く表しているものもあれば(例:HIVウイルス量),そうでないものもある(例:骨密度)
  - 介入による予想される副作用は,必ず測定されなければならない。

- 対象者の選択 (第3章参照)
- 選択基準の設定
    - inclusion criteria (取り込み基準)
      - 研究にふさわしく,かつ統計学的に有意な介入効果を得やすい集団を定義する
    - exclusion criteria (除外基準)
      - むやみに設けすぎない,設定される場合,理由は主に5つに分類される。
        - 1) 試験治療が有害, 2) 試験治療が効果がない可能性が高い, 3) アドヒアランスが悪い可能性が高い, 4) 経過観察が困難, 5) 研究への参加に現実的な問題がある

- ベースラインの変数の測定 (1)
- 対象者の特性を記述する
    - 研究結果の一般化可能性を他の人が判断できる材料の測定
  - 予測因子やサブグループ化に役立つ因子を測定する
    - 主要評価項目とは別の予測因子と結果の関連を研究できる
    - 群間で予測因子の分布が偶然異なった場合に,統計学的に補正して比較できる
    - サブグループにより,介入の効果が検討できる

- ベースラインの変数の測定 (2)
- 結果変数のベースラインを測定する
    - 2値変数の場合は,研究開始時点での状態と研究後の結果を比較,連続変数の場合は,研究前後での変化を比較することが多い
  - 無駄な測定はしない
    - 測定を最小限にして大規模な臨床試験を行う
    - ランダム化により交絡因子が除去されるはず
  - 検体バンクを確立する
    - 結果に影響する生物学的予測因子が後から予測できる
    - 後から別のテーマでの研究ができる

ランダム化によるベースライン因子の交絡の除去と盲検化による共介入の交絡の除去(表10.2)

関連の生じた原因	結論の確実性を高めるための方策
1. 偶然	観察研究の場合と同じ
2. バイアス	
3. 結果 - 原因関係	実験的研究ではなし
4. 交絡	ランダム化前の交絡因子 → ランダム化 ランダム化後の交絡因子 → 盲検化
5. 因果関係	

## ランダム化とは

- 予見不可能に群を決定すること
  - 選択バイアスを防ぐ
- 群間の比較可能性を確率的に保証する
  - 未知の因子についても保証する
  - 偶然偏ることもある
- 第三者が割り付けるのが理想
- ランダム化の遵守

## ランダム化の前に

- ベースラインデータの測定
- 選択・除外基準を満たすことの確認
- インフォームドコンセントの確認

## 小規模研究におけるランダム化

- コイントス
- 乱数の利用
- 予見不可能性の保証が大切

## 大規模研究におけるランダム化

- 第三者機関で割付
- 研究者の意図が入らないことが大切

## ランダム化の方法

- 割付比
- ランダム化法
  - 通常のランダム化  
単純ランダム化, ブロックランダム化, 層別ランダム化
  - 動的割付(割付結果を見ながら次の人を割り付ける)
  - 最小化法, つぼモデル

## 単純ランダム化

- 割り付ける群をA群とB群とすると
  - ランダムにAとBを発生させる
  - 数が少ないとAとBの数が違う可能性がある
  - 性別, 年齢といった治療に影響する因子が片方に偏る可能性がある