

15. 医

学会発表・学術論文発表一症例報告を含む医学論文及び学会研究会発表における患者プライバシー保護に関する指針

(平成16年4月6日、外科関連学会協議会 加盟学会)

- ① 患者個人の特定可能な氏名、入院番号、イニシャル又は「呼び名」は記載しない
- ② 患者の住所は記載しない。但し、疾患の発生場所が病態等に関与する場合は区域までに限定して記載することを可とする（神奈川県横浜市等）
- ③ 日付は、臨床経過を知る上で必要となることが多いので、個人が特定できないと判断される場合は年月までを記載してよい
- ④ 他の情報と診療科名を照合することにより患者が特定され得る場合、診療科名は記載しない
- ⑤ 既に他院等で診断・治療を受けている場合、その施設名並びに所在地を記載しない。但し、救急医療等で搬送元の記載が不可欠な場合はこの限りではない
- ⑥ 顔写真を提示する際には目を隠す。眼疾患の場合は、顔全体がわからないよう眼球のみの拡大写真とする
- ⑦ 症例を特定できる生検、剖検、画像処理に含まれる番号等は削除する
- ⑧ 以上を配慮しても個人が特定される可能性がある場合は、発表に関する同意を患者自身（または遺族か代理人、小児では保護者）から得るか、倫理委員会の承認を得る
- ⑨ 遺伝性疾患やヒトゲノム・遺伝子解析を伴う症例報告では、「ヒトゲノム・遺伝子解析研究に関する倫理指針」（平成13年3月29日）による規定を遵守する

→所々、改訂はされていると思いますが

16. 匿名と連結

識別性→匿名化

匿名化とは、個人の識別情報を取り除き、その人を識別できないようにすること

連結可能→Traceability（追跡可能）であること

連結可能匿名化とは、必要な場合に提供者を識別できるよう、当該提供者と新たに付された符号又は番号の対応表を残す方法による匿名化

連結不可能とは、その人と新たに付された符号又は番号の対応表を残さない（対応表を破棄する）こと

17. 個人情報保護の考え方と医療個人情報保護のずれ

連結・匿名化

匿名化されていないものや連結可能匿名化の場合で匿名化の対応表を保有していることにより個人の識別が可能な状態の情報については、個人情報保護の対象となる

例：例えば、A 病院が提供者から試料の提供を受け、連結可能匿名化した後、別法人の B 研究所に提供された情報は、B 研究所にとっては対応表を保有していない連結可能匿名化した情報となる。したがって、この場合、当該情報は、B 研究所においては「ヒトゲノム・遺伝子解析研究に関する倫理指針」で定義する個人情報に該当しないことになる。なお、A 病院と B 研究所が同一法人内にある場合は、対応表が A 病院で管理されている場合であっても、当該法人として対応表を保有していることになるので、A 病院から提供された連結可能匿名化された情報は、B 研究所においては個人情報となる

（「ヒトゲノム・遺伝子解析研究に関する倫理指針」についての Q&A、平成 17 年 3 月 18 日より）

「医学研究等における個人情報の取扱いの在り方等について」（平成 16 年 12 月 24 日、ライフサイエンス研究におけるヒト遺伝情報の取扱い等に関する小委員会、医学研究における個人情報の取扱いの在り方に関する専門委員会、個人遺伝情報保護小委員会）

I. ヒトゲノム・遺伝子解析研究における個人情報の取扱いの在り方及び研究の進展に対応した倫理指針の見直しについてで、「・・・研究の実施に支障を及ぼすことを危惧する意見が強く主張された。したがって、今後、同一法人内の連結可能匿名化情報の取扱いに関する問題点等について適宜実態把握を行うとともに、必要に応じて研究の実施に支障を及ぼさないような所要の措置が検討されるべきものとする。」と記載されている

「なお、ヒトゲノム・遺伝子解析研究に限った個別法ではなく、他の領域や分野も含んだ包括的な法律（4 つの研究指針を総括した研究一括法、医療分野も含んだ個人情報保護法の特別法、医療・医学研究や産業における利用も含めた個人遺伝情報の適正な取扱いの確保のための法律）又は基本法的な法律（生命倫理全般について基本的な原則等を示す生命倫理基本法）による個人遺伝情報の保護、又は保険や雇用等における個人

遺伝情報等に基づく差別の禁止を定めた法等の策定を検討すべきではないかという指摘があったが、いずれも研究にとどまる問題でないことはもとより、人権の尊重、インフォームド・コンセント、個人情報の保護等規定すべき項目が多岐にわたること、領域や分野によっては規制の程度や態様も一様ではないと考えられること等から慎重かつ十分な検討を要すべきであり、拙速な対応は適当ではないことを理由に。中長期的な課題と位置づけられるべきものと考えられる。今後の国際的な動向等も見据え。適切な時期にこの課題に対する検討が開始されるよう望む。」とも記載されている

以下を特に参考とさせていただきました。

- ・ 赤林朗編「入門・医療倫理Ⅰ」勁草書房 東京 2005年
- ・ 開原成充・樋口範雄編「医療の個人情報保護とセキュリティ 第2版」有斐閣 東京 2005年
- ・ 「個人情報の保護に関する法律」(平成15年5月30日、法律第57号)
- ・ 「ヒトゲノム・遺伝子解析研究に関する倫理指針」(平成13年3月29日(平成16年12月28日全部改正、平成17年6月29日一部改正)、文部科学省・厚生労働省・経済産業省)
- ・ 「遺伝子治療臨床研究に関する指針」(平成14年3月27日(平成16年12月28日全部改正)、文部科学省・厚生労働省)
- ・ 「疫学研究に関する倫理指針」(平成14年6月17日(平成16年12月28日全部改正、平成17年6月29日一部改正)、文部科学省・厚生労働省)
- ・ 「臨床研究に関する倫理指針」(平成15年7月30日(平成16年12月28日全部改正)、厚生労働省)
- ・ 「医療・介護関係事業者における個人情報の適切な取扱いのためのガイドライン」(平成16年12月24日、厚生労働省)
- ・ 「医学研究分野における当面の取組について」(平成17年1月27日、文部科学省研究開発振興局ライフサイエンス課生命倫理・安全対策室・厚生労働省大臣官房厚生下科学課・経済産業省製造産業局生物化学産業課事業環境整備室)
- ・ 「医学研究等における個人情報の取扱いの在り方等について」(平成16年12月24日、ライフサイエンス研究におけるヒト遺伝情報の取扱い等に関する小委員会、医学研究における個人情報の取扱いの在り方に関する専門委員会、個人遺伝情報保護小委員会)
- ・ 「医師の職業倫理指針」(平成16年2月、日本医師会)

測定方法を計画する(後半) Precision and Accuracy

Designing Clinical Research 2nd Chap4
Sep2006臨床研究センター勉強会
臨床研究フェロー 斉藤真梨

目次

- 測定のスケール
- 精度(precision)
- 正確性(accuracy)
- 測定方法が備えるべき他の条件
- 保存資料の測定
- 最後に
- まとめ

2

前回のおさらい

- 測定は興味ある現象を数値に変換するプロセス
- スケールの種類
 - 連続変数、カテゴリ変数
- 誤差の少なさの指標
 - 精度(precision)と正確性(accuracy)

3

測定値のスケール

- 連続変数
 - 量的に表現できる現象を表す
 - 身長、体重、血圧...
- カテゴリ変数
 - 量的に表現しにくい現象を表す
 - 生存/死亡、血液型、QOL...
 - 名義変数と順序変数
 - 2値変数と多値変数
- 情報量の多さからは連続変数が望ましい

4

備えるべき性質

- 精度(precision)
 - 測定値が安定している度合い
 - 偶然誤差(random error)が少ないこと
 - 測定を繰り返したときの結果の一致度で評価する
 - 標準偏差、変動計数、一致割合、 κ 計数、相関係数

FIGURE 4.1
The relationship between precision and accuracy

5

備えるべき性質

- 正確性(accuracy)
 - 測定値が真値に近い度合い
 - 系統誤差(systematic error)が小さいこと
 - ゴールドスタンダードとの一致度で評価
- 妥当性(validity)
 - ゴールドスタンダードがないときの指標
 - 内容妥当性、構成概念妥当性、基準関連妥当性

FIGURE 4.2
The difference between precision and accuracy

6

目次

- 測定のスケール
- 精度(precision)
- 正確性(accuracy)
- 測定方法が備えるべき他の条件
- 保存資料の測定
- 最後に
- まとめ

7

測定方法が備えるべき他の条件

- 感度(sensitivity)
 - 目的とする特性や状態の違いをよく捉えているか
 - 「禁煙した」よりも「喫煙本数が何本減った」
- 特異度(specificity)
 - 目的とする特性をどれだけ純粋に捉えられるか
 - 違うものをどれだけ排除できるか

8

測定の評価指標としての定義

- 検査のよさの指標
 - 感度: 陽性のものを正しく陽性と判定する確率
 - 特異度: 陰性のものを正しく陰性と判定する確率
- 似ている指標
 - 陽性的中率(Positive Predictive Value)
 - 陰性的中率(negative Predictive Value)

9

この2種類の違いは

- 感度と特異度
 - 真値が分母
 - スクリーニングの評価指標として使われる

		真の状態	
		疾患あり	疾患なし
検査結果	陽性	A	C
	陰性	B	D

10

この2種類の違いは

- 陽性的中率と陰性的中率
 - 検査結果が分母
 - 稀な疾患のときあまりよくない

		真の状態	
		疾患あり	疾患なし
検査結果	陽性	A	C
	陰性	B	D

11

Prevalent diseaseのとき

- 感度: 90%、特異度: 97.5%
- 陽性的中率: 95%、陰性的中率: 95%
- どちらの指標で評価しても変わりなさそう

		真の状態	
		疾患あり	疾患なし
検査結果	陽性	900	50
	陰性	100	1950
合計人数		1000	2000

11

Rare diseaseのとき

- 患者数がさっきの100分の1だったら
- 感度:90%、特異度:97.5%
- 陽性的中率:15%、陰性的中率:99.9%
- 陰/陽性的中率は潜在患者割合に依存する

		真の状態	
		疾患あり	疾患なし
検査結果	陽性	9	50
	陰性	1	1950
合計人数		10	2000 ¹³

測定方法が備えるべき他の条件

- データの適切な分布がえられること
 - データがとりうる広範囲をカバーできる指標を用いる
 - そのスケールにデータが満遍なく分布している
- 客観性
 - 測定者数を減らす
 - 測定機器を自動化する
 - ...
- 質的な測定
 - 定量的な測定だけでなく質的な測定項目も加えておく
 - 量的な測定だけだと視野が狭くなりがち

14

保存試料の測定

- 保存された試料やデータも使える
 - 臨床化学検査
 - 遺伝子解析
 - 放射線画像

15

保存資料の利点

- コストの削減
 - 追跡中の集団の中で、対象疾患を発症した人だけを検査できる
 - 例)コホート内ケースコントロール研究
- 精度の向上
 - 同じ試料について複数回測定できる
- 再利用、研究の向上
 - 新しい概念、測定技術ができたとき再評価できる

16

具体的に

- 遺伝疫学、分子疫学の進歩
 - 治療への反応性に関連する遺伝子型の解析
 - DNAを含む生体試料を利用
- 画像技術の進歩
 - 予防医学に役立つデータの検出
 - 大腿骨頭部の骨塩量、冠動脈壁のカルシウム量測定による骨折予防 など

17

最後に

- 測定の種類
 - 病歴、身体測定、...
 - 9章(交絡や因果関係の推論のための測定の組み合わせ)
 - 15章(質問票やその他の測定手段)
- 効率efficiency
 - 情報も大事だが時間・費用などのコストも大事
 - 測定項目の質を高める、項目数を減らす
 - 測定項目が多すぎると対象者・研究者にも負担になる

18

サンプルサイズを見積もるための準備

Designing Clinical Research 2nd Chap5
 Oct2006臨床研究センター勉強会
 国立成育医療センター臨床研究フェロー
 斉藤真梨

1

サンプルサイズとは

- ☞ サンプルサイズとは
 - ↳ 研究対象者の人数のこと
 - ↳ 少なければ偶然誤差の影響をうけやすい
 - ↳ 多ければ母集団の真値に近い結果が得られる
 - ↳ 多ければコストがかかる
- ☞ サンプルサイズ設計の必要性
 - ↳ 対象者数不足で結論が出せないかもしれない
 - ↳ 別の指標に変えなければならぬかもしれない
 - ↳ 事前に実施可能性を考えなければならない

2

サンプルサイズ設計に必要な情報

- ☞ 仮説
 - ↳ 帰無仮説と対立仮説
- ☞ 効果量
 - ↳ どれくらいの効果が期待できるか
 - ↳ 結果指標のバラツキ具合
- ☞ 許容するエラーの程度
 - ↳ α エラーと β エラー
 - ↳ 有意水準

3

研究テーマを研究仮説に

- ☞ 研究テーマ
 - ↳ 知りたいこと
 - ↳ 「閉経後女性におけるエストロゲンの心臓発作への影響」
- ☞ 研究仮説
 - ↳ 実際検証したいこと
 - ↳ 「閉経後の女性で、エストロゲン療法を受けた者は受けていない者に比べ心臓発作を起こす頻度が低い」

4

仮説が必要なものは

- ☞ 記述研究には不要
 - ↳ 分布の記述が目的のとき
- ☞ 比較研究では必要
 - ↳ 例) 患者群とコントロール群の間で...を比較
 日本では過去の米国での研究結果に比べ...
- ☞ 見分けるキーワード
 - ↳ 「よりも」「に比べ」「と同じ」
 - ↳ 「の原因である」「を生じる」「と関連がある」
- ☞ 示したいことがあれば仮説は必要

5

よい仮説の3条件

- ☞ 簡潔であること
 - ↳ 一つの予測因子と一つの結果
- ☞ 的確であること
 - ↳ 必要なことが具体的に書かれている
- ☞ 事前に立てていること
 - ↳ その仮説に基づいて研究を計画する

6

簡潔であること

- ④ 対象集団、予後因子、結果因子からなる
 - ✓ 例)糖尿病患者において、運動不足は、蛋白尿のリスクを上げる
- ④ 予後因子、結果因子は一つずつが基本
- ④ 複数の仮説では一度に検定できない
 - ✓ 糖尿病患者において、アルコール摂取が、蛋白尿と神経障害のリスクに影響する

7

複数の場合は

- ④ 別々に評価する
 - ✓ アルコール摂取は、蛋白尿のリスクに影響する
 - ✓ アルコール摂取は、神経障害のリスクに影響する
- ④ まとめてしまう
 - ✓ 「糖尿病患者において、ビール・ワイン・ウイスキーの摂取は、蛋白尿のリスクに関連がある」
- ④ 興味による
 - ✓ 別々に関連が知りたい
 - ✓ お酒の効果の違いに興味があるのではなく、飲酒の有無と蛋白尿のリスクの関連を知りたいとき

8

的確であること

- ④ はっきり表現されていること
 - ✓ 観察する因子
 - ✓ 用いる指標
 - ✓ 比較方法
- ④ 測定単位や基準値を明記しておく
 - ✓ ×「使用」 ○「服用量/日」「使用者割合」「服用頻度」
 - ✓ 「mg/day」、「>30mg/dl」
 - ✓ プロトコルに書いてもよい

9

事前に立てること

- ④ 研究を始める前に明文化する
- ④ 研究目的がはっきりする
- ④ 結果の解釈も容易になる
- ④ 事後仮説
 - ✓ 結果をみてから立てる仮説
 - ✓ 過剰解釈につながる
 - ✓ 多重性の問題が生じる(検定のしすぎ、後述)

10

研究仮説から検定仮説へ

- ④ 「運動不足は蛋白尿のリスクを上げる」
- ④ 帰無仮説(Null hypothesis)
 - ✓ H_0 : 運動不足と蛋白尿のリスクは関連がない
 - ✓ 「示したい効果がない」という仮説
- ④ 対立仮説(Alternative hypothesis)
 - ✓ H_1 : 運動不足と蛋白尿のリスクは関連がある
 - ✓ 研究者が研究を通じて示したいこと

11

片側仮説と両側仮説

- ④ 対立仮説にも2種類
- ④ 両側仮説
 - ✓ 関連・効果の向きを限定しない
 - ◇ A群とB群では胃潰瘍の頻度が異なる
- ④ 片側仮説
 - ✓ 関連・効果の向きを一方に限定
 - ◇ A薬群はプラセボ群より皮疹の発生頻度が高い
- ④ よほどの根拠がない限り両側
 - ✓ 当然と思われていた仮説が否定された例

12

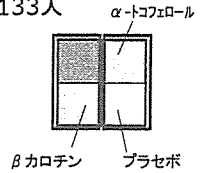
逆転したβカロチンの効果

- ☞ 1980年代の疫学的なエビデンス
 - ↳ 「βカロチンが癌(特に肺癌)の罹患率を下げる」
- ☞ ハイリスク集団を対象に二つの試験が行われた
- ☞ ATBC試験
 - ↳ フィンランド人29133人を5~8年追跡
- ☞ CARET試験
 - ↳ アメリカ人18314人を平均4年追跡

13

ATBC試験(1994)

- ☞ 対象: 50~69歳の喫煙男性29133人
- ☞ デザイン: 要因実験デザイン
 - ↳ βカロチン群
 - ↳ α-トコフェロール群
 - ↳ βカロチン+α-トコフェロール群
 - ↳ プラセボ群
- ☞ 結果
 - ↳ βカロチンを摂取した群で
 - ↳ 肺癌罹患率 1.18倍(95%CI: 1.03-1.36)
 - ↳ 総死亡率 1.08倍(95%CI: 1.01-1.13)



14

CARET試験(1996)

- ☞ 対象: 45~74歳の喫煙者・アスベスト曝露労働者
- ☞ デザイン: 多施設ランダム化二重盲検プラセボ対照試験
 - ↳ βカロチン+レチノール群
 - ↳ プラセボ群
- ☞ 結果
 - ↳ 肺癌発生率 1.28倍(95%CI: 1.04~1.57)
 - ↳ 総死亡率 1.17倍(95%CI: 1.03~1.33)
(脳溢血と心臓発作)

15

その後

- ☞ 当然試験は途中で中止
- ☞ ATBCの追跡調査
 - ↳ サプリメントを中止したところ死亡率は減少
 - ↳ それでもβカロチン群の全原因の死亡率は1.07倍
- ☞ CARETの追跡調査
 - ↳ 中止後6年後まで継続
 - ↳ サプリ群で肺癌発生率1.12倍、総死亡1.08倍
- ☞ βカロチンによっておきた細胞変化によると言われている
- ☞ 非喫煙者には有害ではない
- ☞ 野菜や果物に含まれるβカロチンではない
- ☞ 世界中の店頭からサプリメントが消える

16

研究仮説は片側

- ☞ 研究仮説(研究者がしめしたい仮説)は片側でよい
 - ↳ 対立仮説、帰無仮説は検定するためのツール
 - ↳ 対立仮説は両側がよい
- ☞ 両側検定は厳しい検定
 - ↳ 片側p値が0.04だったら両側p値は0.08
 - ↳ 片側検定なら有意水準5%で有意
 - ↳ 両側検定なら有意水準5%で有意でない
 - ↳ 詳細は後日
- ☞ 片側か両側か決めるのはデータを集める前に

17

サンプルサイズ設計に必要な情報

- ☞ 仮説
 - ↳ 帰無仮説と対立仮説
- ☞ 効果量
 - ↳ どれくらいの効果が期待できるか
 - ↳ 結果指標のバラツキ具合
- ☞ 許容するエラーの程度
 - ↳ αエラーとβエラー
 - ↳ 有意水準

18

効果量(effect size)

- ④ 研究者が示したい効果の大きさ
 - ✓ 例) 運動すると運動しないときより空腹時血糖が15mg/dlくらい下がるだろう
- ④ 研究者が決定する
 - ✓ 臨床的に意味のある値
 - ✓ 先行研究から
- ④ 効果量が小さいほど必要サンプルサイズは大きくなる
 - ✓ 大きな差のあるものなら少人数でも検出できる

19

効果量の指標

- ④ 差、比、オッズ比、率比…
 - ✓ 論文には大きく見える値が載っていることもある
 - ✓ 中年男性の難聴頻度30%、中年女性の難聴頻度20%
 - ◇ 絶対差 $30-20=10\%$
 - ◇ 比 $30/20=1.5$
 - ◇ 標準化差 $(30-20)/20=50\%$
- ④ サンプルサイズの計算方法
 - ✓ 指標によって異なる
 - ✓ 結果はほぼ同じ
 - ✓ 仮説が複数あるときは一番大事なものに注目

20

データの変動度(variability)

- ④ 効果量のバラツキ具合
 - ✓ バラツキやすい指標=サンプルがたくさん必要
- ④ P60の例
 - ✓ 測定値にバラツキが少ない(上図)
 - ◇ 違いがわかりやすい
 - ✓ 測定値のバラツキが大きい(下図)
 - ◇ 違いがわかりにくい
- ④ できるだけ安定した指標、安定した測定を選ぶ

21

サンプルサイズ設計に必要な情報

- ④ 仮説
 - ✓ 帰無仮説と対立仮説
- ④ 効果量
 - ✓ どれくらいの効果が期待できるか
 - ✓ 結果指標のバラツキ具合
- ④ 許容するエラーの程度
 - ✓ α エラーと β エラー
 - ✓ 有意水準

22

仮説検定とエラー

- ④ 研究仮説
「一日5分以上の運動は糖尿病をもつ中年女性の空腹時血糖を低下させる」
- ④ 母集団の調査は不可能
 - ✓ 対象者をサンプリング
 - ✓ サンプルで観察される出来事から、母集団における現象を推測
- ④ 仮説検定とエラーを裁判官の判決の例で説明

23

裁判官の頭の中

- ④ 被告が犯罪を犯したかどうか絶対的にはわからない
- ④ 「被告は無罪」を仮定する
- ④ 無罪を覆すのに十分な証拠があるか?
 - ✓ どれだけ十分な証拠なのか
 - ✓ 「疑わしきは罰せず」の原則
- ④ 判決「有罪！」／「無罪！」
- ④ 裁判官だって間違える可能性はある
 - ✓ 無実の人に有罪判決
 - ✓ 犯罪者に無罪判決

24

裁判官の頭の中

- ☞ 被告が犯罪に「**帰無仮説**」どうか絶対「**データ**」からない
- ☞ 「被告は無罪」を仮定する
- ☞ 無罪を覆すのに十分な証拠があるか？
 - ↳ どれだけ十分な証拠なのか
 - ↳ 「疑わしきは罰せず」の原則
- ☞ 判決「有罪！」／「無罪！」
- ☞ 裁判官だって間違える可能性はある
 - ↳ 無実の人に有罪判決
 - ↳ 犯罪者に無罪判決

有意水準

α エラー
 β エラー

25

結論を誤る確率

- ☞ 裁判よりは系統的なはず、でも..
 - ↳ 対象集団が母集団をうまく代表していないかもしれない
 - ↳ 測定にはランダム誤差がある
- ☞ エラーをいくつまで許容するかを事前に設定する

		母集団の状態	
		効果がない	効果がある
研究の結論	帰無仮説を採用	正しい	β エラー
	対立仮説を採用	α エラー	正しい

26

α エラーと β エラー

- ☞ α エラー
 - ↳ 母集団では(本当は)関連や効果がないのに、偶然データで効果が示してしまう確率
 - ↳ あわて者の α (効果があると早とちりした)
- ☞ β エラー
 - ↳ 母集団では(本当は)関連や効果があるのに、偶然データで効果が示せない確率
 - ↳ ぼんやり者の β (効果を見逃した)

27

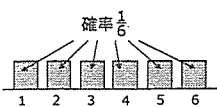
エラーをどれくらい許容するか

- ☞ エラーが小さい設定ほど厳密な研究といえる
 - ↳ α エラー: 5%、1%
 - ↳ β エラー: 20%、10%
- ☞ 検出力(power)
 - ↳ $100 - \beta$ エラー
 - ↳ 80%、90%
 - ↳ 本当に効果があり、それを研究で示せる確率
- ☞ サンプルを増やすと結果は真値に近づく
- ☞ エラーが減る

28

サンプルサイズとエラーの関係

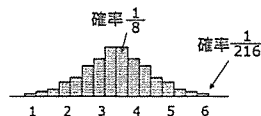
- ☞ 「さいころの目」の期待値
 - ↳ 無限回繰り返し投げる
 - ↳ 出る目の平均は3.5
- ☞ 「さいころを1回投げる」研究
 - ↳ 1~6が均等な確率 $1/6$ が出る



29

サンプルサイズとエラーの関係

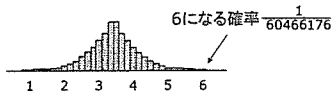
- ☞ 「さいころを3回投げて平均を調べる」研究
 - ↳ 3.5から遠い結果になる確率は低い
 - ↳ 3.5に近い結果になる確率は高い
- ☞ 1回投げる研究より真値3.5に近い値がやすい



30

サンプルサイズとエラーの関係

- ☞ 「さいころを10回投げて平均を調べる」研究
 - ✓ 3.5から遠い結果になる確率は低い
 - ✓ 3.5に近い結果になる確率は高い
- ☞ 3回投げる研究より真値3.5に近い値がやすい
- ☞ 回数を増やすほど真値に近づく



31

大きければいくらでもいい？

- ☞ サンプルサイズを増やせばエラーは減る
 - ✓ 極端な話、母集団全体を調べれば真値がわかる
- ☞ でもコストがかかる
- ☞ それだけの意味があるか
 - ✓ 「飲むと血圧が平均30mmHgさがる薬A」
少ない人数でも効果が検証できる
 - ✓ 「飲むと血圧が平均3mmHgさがる薬B」
偶然誤差に打ち消されかねない
たくさん的人数がいないと検証できない
 - ✓ そもそもB薬には臨床的な意味がある？

32

効果の大きさ ≠ p値の小ささ

- ☞ 同じ効果でもp値は変わる
 - ✓ 対象: 40~65歳の2型糖尿病を持つ女性100人
 - ✦ 運動群135mg/dl、非運動群140mg/dl
 - ✦ P値=0.07
 - ✓ 対象: 40~65歳の2型糖尿病を持つ女性500人
 - ✦ 運動群135mg/dl、非運動群140mg/dl
 - ✦ P値=0.01
- ☞ 両群の差はどちらも5mg/dl
- ☞ 効果が同じでもサンプルサイズが増えればp値は小さくなる

33

仮説の補足

- ☞ 多重仮説と多重性の問題
 - ✓ 一つの結論に対して複数の仮説をたてる
- ☞ 「運動は、糖尿病患者において、合併症を予防する」
 - ✓ 運動は、蛋白尿のリスクをへらす
 - ✓ 運動は、神経障害のリスクをへらす
 - ✓ 運動は、血糖値がさがる
- ☞ それぞれ検定する
- ☞ α エラー(母集団では効果がないのに、偶然データで効果が示してしまう確率)が増大

34

多重性の問題

- ☞ 一つの結論に対して複数回検定を行うと α エラーが大きくなるということ
- ☞ 対処
 - ✓ ボンフェローニ法(Bonferroni method)
 - ✦ それぞれの有意水準を「 α /検定回数」にする
 - ✦ さっきの例だとそれぞれ有意水準「0.05/3」で検定する
 - ✓ 検定が厳しくなる分サンプルはたくさん必要

35

調整の必要性

- ☞ ちなみに
- ☞ あくまでも一つの結論に対して複数の仮説があるときのみ必要
 - ✓ ゲノム解析で、疾患に関係する遺伝子を膨大な候補の中から探す場合など
- ☞ 一つの研究の中でも導く結論が異なる仮説なら不要
 - ✓ 「運動は、糖尿病患者の蛋白尿のリスクをへらす」
 - ✓ 「運動は、糖尿病患者のQOLをあげる」
- ☞ ただし、p値の小さい結果だけ報告するのはダメ

36

複数の仮説も

- ④ 欠点
 - ✓ 多重性の問題が...
 - ✓ 結論が不明瞭に...
- ④ 利点
 - ✓ 多くの研究テーマを扱える
 - ✓ 似た仮説が同時に示せればエビデンスが補強される
- ④ 主仮説primary hypothesisを決めておく
 - ✦ A薬はプラセボより入院・心臓死亡・総死亡率全てが低かった
 - ✦ A薬はプラセボより心臓死亡率は低かった
 - ✦ A薬はプラセボより入院・心臓死亡率は低く総死亡率は高かった

37

仮説を立てずに

- ④ 背景を整理し、結果をまとめた
- ④ 途中で予想外の変数間で相関があった
- ④ 検定したら有意だった
- ④ 新たな知見！
 - × こっちに重点をおいて報告しよう
 - 今後の研究に役立てよう
- ④ 悪いと言われる理由
 - ✓ その比較検討の妥当性が担保されていないから
 - ✓ 多重性の問題(偶然有意だっただけかも知れない)
- ④ Data mining自体は重要な作業

38

まとめ

- ④ サンプルサイズを見積もるために
 - ✓ 帰無仮説と対立仮説
 - ✓ 示したい効果量とバラツキ
 - ✓ 許容するエラーの程度
 - ✓ サンプルサイズとp値の関係
- ④ その他
 - ✓ 仮説が複数あるとき
 - ✓ 仮説をたてなかったとき

39

検定の概念と用語の整理

Designing Clinical Research 2nd Pre-Chap5
Sep2006臨床研究センター勉強会
臨床研究フェロー 齊藤真梨

論文の結果の解釈

- 具体的な指標の解釈は容易
 - 「実薬群の改善割合は90%であった」
 - 「血圧が〇〇mmHG減少した」
- 結果につきもののp値
 - 「 p 値は0.04で、統計的有意差が認められた」
 - 「 p 値は0.11で、有意差が認められなかった」
- P値とは？

「EBMについて」

- 「EBMが誤解されている！」と訴えるHPの記事
 - p 値とは、2つの治療効果が同じであるという確率を示したものです
 - p 値とは、帰無仮説が正しい可能性を示す
 - $p < 0.05$ であれば、2つの治療効果が同等である確率が5%未満であるということを示します
 - 95%以上の確率で片方の治療効果の劣が優れていると言えます
- それこそ誤解です

今回の目的

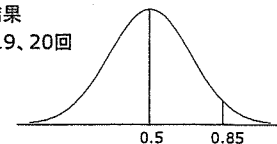
- 検定の考え方
 - 帰無仮説と対立仮説
 - P値
 - 有意水準
 - 統計学的有意
- エラーと検出力
 - α エラー
 - β エラー
 - 検出力(パワー)
- 用語のまとめ

検定の概念

- コイン投げの例で考える
- あるコインを20回投げる実験
 - 表が17回出た
- 仮説をたてる
 - 「このコインはゆがんでいる」
- 検定のための仮説
 - 帰無仮説「表の出る確率 $p=0.5$ 」
 - 対立仮説「表の出る確率 $p \neq 0.5$ 」

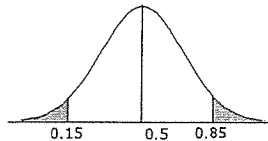
真値が0.5だったら

- 同じ実験をたくさん繰り返したら
- 結果は $P_0=0.5$ 周りにばらつくだろう
- この分布において17回($P=0.85$)という結果は普通？珍しい？
- 17回以上に珍しい結果
 - 帰無仮説を否定する結果
 - 0、1、2、3、17、18、19、20回



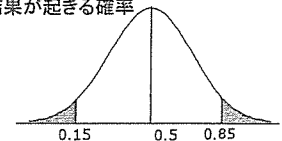
データ以上に稀な結果がでる確率

- これらの確率の合計を計算する
 - 色つきの部分の面積
- これがp値
 - 両側合計しても全体の0.26%
- 真値が0.5なら、得られたデータはかなり珍しい
 - ⇒ 真値の仮定がおかしい
 - ⇒ 帰無仮説を棄却する



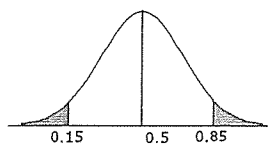
以上が検定の考え方

- 「コインはゆがんでいる」を示すために帰無仮説「ゆがんでいない」を仮定
- 仮定がデータと矛盾することを示す
- 矛盾の度合いがp値
 - 帰無仮説が正しいときに、得られたデータ以上に帰無仮説に矛盾する結果が起きる確率 (色つきの部分の面積)



P値がいくつなら棄却？

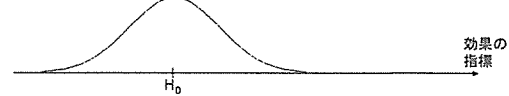
- 研究開始前に決めておく
 - 「p値が〇以下なら帰無仮説を棄却する」
- この〇の値が「有意水準」
 - p値 < 有意水準なら「統計的に有意」と結論する
 - 通常5%や10%
 - 今回は0.26% < 5%
 - 「有意水準」≠「P値」



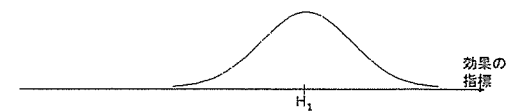
α エラーと β エラー

- 結果の分布

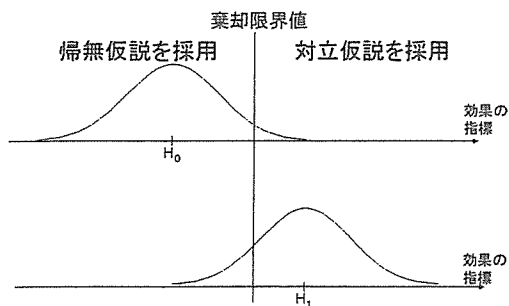
帰無仮説が正しかったら



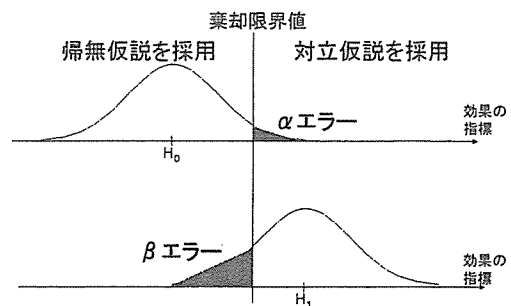
対立仮説が正しかったら



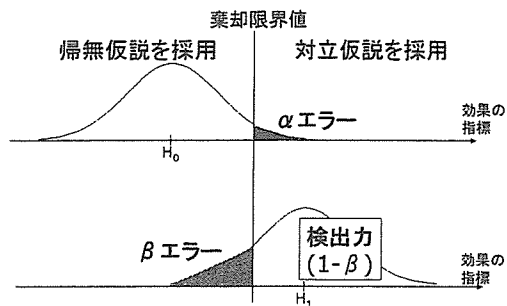
棄却限界値を境に



エラーがあって当たり前

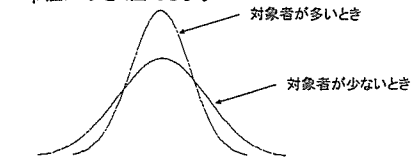


検出力とは



サンプルサイズとの関係

- 対象者数が多いと、結果の精度が上がる
- バラツキが小さくなる
- 分布がスリムになる
- p値も二つのエラーも小さくなる
- 対象者を増やしすぎると、医学的にどうでもいい差でもp値が小さく出てしまう



用語のまとめ①

- 帰無仮説 (Null hypothesis)
 - 否定したい仮説
- 対立仮説 (Alternative hypothesis)
 - 示したい仮説
- P値 (p value)
 - 帰無仮説が正しいときに、得られた結果以上に極端な結果が出る確率
- 有意水準 (Significant level)
 - P値がいくつ以下なら帰無仮説を棄却するかの基準

用語のまとめ②

- 棄却限界値 (Critical value)
 - これより帰無仮説から遠い値なら帰無仮説を棄却するという境目の値
 - 有意水準に対応する
- αエラー (α error)
 - 誤って帰無仮説が棄却されてしまう確率
- βエラー (β error)
 - 誤って対立仮説が棄却されてしまう確率
- 検出力 (Power)
 - 正しく対立仮説が採用される確率

用語のまとめ③

- 統計学的有意
 - 棄却限界値より極端な結果がでること
 - 臨床的に意味のある差ではない
 - P値が小さいほど効果が大きいわけではない

		真実	
		効果がない	効果がある
研究の結論	帰無仮説を採用	正しい	βエラー
	対立仮説を採用	αエラー	正しい (1-β)

観察研究をデザインする コホート研究

Designing Clinical Research 2nd Chap7
Oct2006臨床研究センター勉強会
国立成育医療センター臨床研究フェロー
斉藤真梨

1

実験研究が理想

- ☞ 喫煙と心疾患の関係を調べたい
- ☞ 実験(介入)研究なら因果関係がわかる
- ☞ 動物で
 - ↳ 動物に煙草の煙を強制的に吸わせる
 - ↳ すわせない動物と比較
 - ↳ 「すわせると～の発症が増える」
- ☞ 人間で
 - ↳ 人間に煙草の煙を強制的に吸わせる
 - ↳ すわせない人間と比較
- ☞ 倫理的に問題がある

2

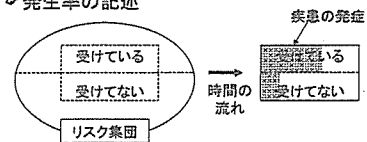
実験研究と観察研究

- ☞ 健康にいい影響があるかもしれない要因
 - ↳ 薬、治療法の候補
 - ↳ 実験的に調べられる
 - ↳ 臨床試験
- ☞ 健康に悪い影響があるかもしれない要因
 - ↳ 薬害、公害、生活習慣...
 - ↳ 実験的には調べられない
 - ↳ 観察研究

3

コホート研究とは

- ☞ 対象集団(コホート)一定期間追跡観察する
 - ↳ 興味ある曝露を受けている群
 - ↳ " " 受けていない群
- ☞ 目的
 - ↳ 因果関係の推測
 - ↳ 発生率の記述



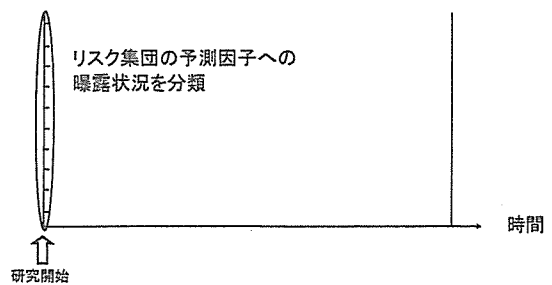
4

前向きコホート研究

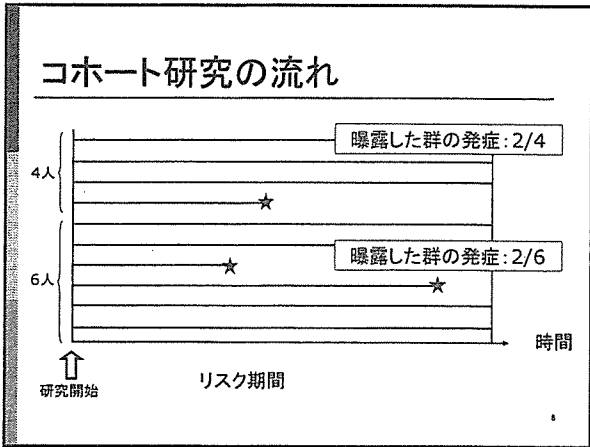
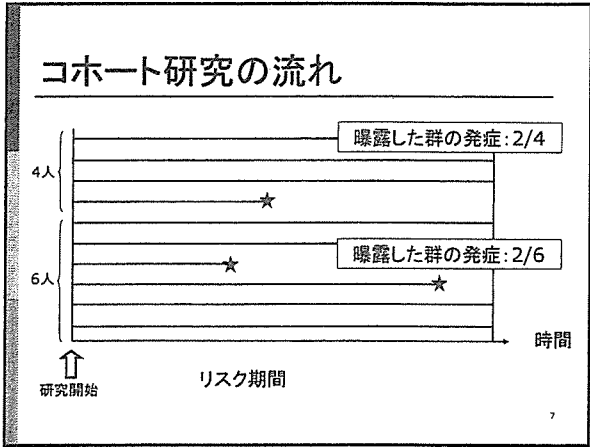
- ☞ リスク集団を設定
 - ↳ 興味ある疾患を発症していない人
 - ↳ 将来発症しうる人
 - 閉経後の女性
- ☞ 予測因子・背景・その他のリスク因子となりそうなものを測定
 - ↳ 要因に曝露した群
 - ↳ 曝露していない群
 - エストロゲン療法を受けている／受けていない群
- ☞ 観察、追跡
 - その後の心血管イベントの発症

5

コホート研究の流れ



6



- ### 利点と欠点
- ☞ 予測因子の測定が結果因子の測定より前
 - ✓ 結果に影響を受けない予測因子の測定ができる
 - ✓ 時間の前後関係が明確
 - 因果関係があるとは結論できない
 - 因果関係がいえるのは実験研究
 - ✓ 思い出しバイアスが入りにくい
 - ☞ コストがかかる
 - ✓ 膨大な人数を長期間追跡
 - ✓ 稀な疾患には向かない
 - 発症が少ない

- ### 行ううえでの注意点
- ☞ 基本的な前提
 - ✓ 研究開始時に測定する予測因子が、結果因子の影響を受けていない
 - ☞ 興味ある疾患の既往のある人は除く
 - ✓ Inception cohort
 - ☞ 研究開始前に一定の期間を設ける
 - ✓ リスク集団の候補をしばらく追跡
 - ✓ 興味ある疾患の発症がないことを確認
 - ✓ ベースラインデータ(予測因子など)を測定

- ### Nurses' Health Study Fuchs1999
- ☞ 癌と冠動脈性疾患に関する研究(1976~)
 - ✓ 米国在住30~55歳の女性看護師121700人
 - ✓ 質問紙を郵送
 - ✓ リスク因子、発症に関して2年おきに調査
 - ☞ サブスタディ
 - ✓ 仮説
「食物繊維摂取が大腸癌のリスクを低下させる」
 - ✓ 食歴調査票を用いて栄養調査(1980~1996)
 - ✓ 88757人から回答

食物繊維の摂取と大腸癌

- ☞ 対象集団を食物繊維摂取量を多い順に5分割
- ☞ 直腸結腸癌のリスク比
 - ✓ 摂取量が最も少ないGroup1を1として

	Group2	Group3	Group4	Group5
C+R	0.81 (0.64-1.04)	0.96 (0.75-1.21)	0.93 (0.72-1.19)	0.95 (0.73-1.25)
Colon	1.04 (0.78-1.39)	1.26 (0.94-1.68)	1.06 (0.77-1.45)	1.04 (0.74-1.46)
Rectal	0.54 (0.32-0.89)	0.47 (0.27-0.79)	0.68 (0.41-1.12)	0.63 (0.37-1.08)

どうやっても

- ☞ Trend (傾向) が見られない
 - ✓ 摂取量が増えるにつれてリスクが低下...しない
 - ✓ 食物繊維の摂取量が一定のサブグループ
 - ✓ 食品別食物繊維量と癌のリスク
- ☞ 結論「食物繊維には大腸がん予防効果はない」

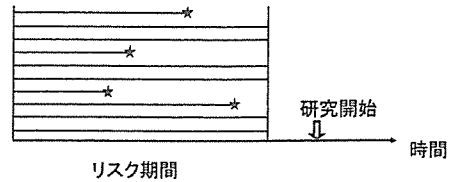
	Group2	Group3	Group4	Group5
Fruit	0.94 (0.74-1.18)	1.03 (0.82-1.29)	0.88 (0.69-1.12)	0.86 (0.67-1.10)
Vegetable	0.98 (0.78-1.23)	0.98 (0.78-1.25)	1.13 (0.89-1.43)	1.35 (1.05-1.72)

Group1より
有意に高い

13

後ろ向きコホート研究

- ☞ 本質的には前向きコホートと同様
- ☞ コホートの設定、予測因子の測定、結果因子の測定が過去
- ☞ 他の目的で設定されたコホート研究のデータを利用
- ☞ 研究の内容や質のコントロールが不可能



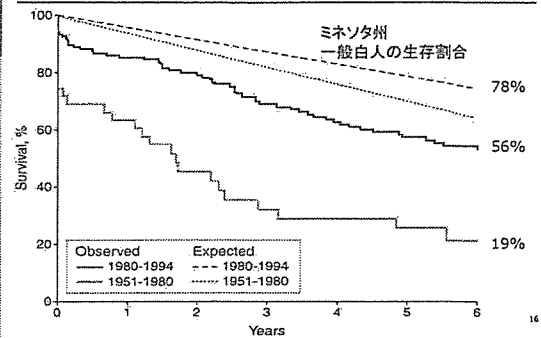
14

後ろ向きコホート研究 Clouse 1998

- ☞ ミネソタ州の住人の疾患登録データ
 - ✓ 背景、臨床症状、最大動脈瘤径、心血管へのリスク因子、手術適応、手術結果、生存情報...
- ☞ 仮説
「治療法の進歩が胸部動脈瘤の予後を改善」
- ☞ 1951~79年と1980~95年の比較

15

1980年の前後の比較



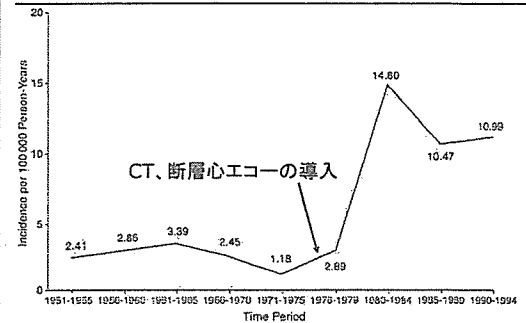
16

80~95年のデータで

- ☞ 大動脈瘤と診断された133人から
 - ✓ 動脈瘤の発症率は10.4人/10万人年
 - ✓ 5年間の累積破裂率は20%(95%CI:12-28%)
 - ✓ 女性のほうが破裂リスクが6.8倍(2.3-19.9)
 - ✓ 症状があるほうが // が7.0倍(2.6-19.3)
 - ✓ 動脈瘤の直径でサブグループ
 - 破裂割合は <4cm:0%
 - 6cm:16%(4-28%)
 - 6cm≦:31%(5-56%)
- ☞ 発生率の記述もできた

17

急に増えている



18