

- まず、「臨床研究をするのに統計なんて二の次で良い」、「目的なく集められたデータはゴミ」この理由について述べられていく。

- 実証主義的科学観

True Value = Observed + Bias + Error

True Value： 神のみぞ知る唯一絶対存在

神に近づくためにはどうするか？ 科学的論文を形成する

Bias： 人間がだまされる．デザインで決まる、妥当性、疫学の問題

Error： 偶然誤差、確率論、統計の問題

- 研究の流れ

仮説の検定 / 研究デザインと仮説検定法の選択 / データ収集 / データのクリーニング / 研究論文完成

このうち統計が関与するのはデータのクリーニングのあたりのみで、良い論文は統計の前に9割できている．統計を行ったら、「やっぱりね」というのがよい．

- 仮説

いわゆる統計解析自体は最後にさらっと行うもので、一番肝心なのは仮説である．

仮説に見合ったデザイン、そして仮説検証を可能にするように慎重に集められたデータが大切．

統計解析は magic bullet でも、玉手箱でもない．

サンプルサイズの決定、記述統計とスクリーニング、中間解析に統計が重要

- サンプルサイズ

effect size どれくらいの効果が期待されるのか

Type 1 error level： P=0.05 でよいのか

Type 2 error level： 検出力

まず事前データが必要である．いきなり RCT は無理．またあくまで偶然誤差だけの場合

- データのクリーニング

- 1．入力内容の確認（ダブル入力）
- 2．ロジックチェック（あり得ない数字に注意する）
- 3．欠損や誤情報の問い合わせ： 実はこれが重要
- 4．データ訂正
- 5．分布の異常
- 6．また問い合わせ、確認、訂正
- 7．データの固定

- 中間解析、評価

- 1．プロトコール遵守の状況
- 2．データ項目の見直し、再定義： 入力フォーマット改訂

3. 研究継続の可否： CAST、HRTT の例など

中間で使う統計は、平均をみたりするなど、簡単なものでよい

● 研究成果をおびやかすもの

1. 選択バイアス

前向き研究のプロトコル違反、途中脱落、欠損値をいかに減らすかが重要

プロトコル違反： INT (intention to treat)原則

Conservative

しかしバイアスであることに変わりなし、本来でべきものを出せない可能性

- 起こらないように予防するのが手堅い

欠損と脱落

参加者間でのプロトコル理解の徹底

データ収集後のフォーマットやレイアウト設計

データソースの明確化

センターごとの責任体制

プロジェクトマネージメントの一元化

2. 情報バイアス：特に差別的誤分類

例) ありなし分類の罫

1. あり 2. なし これでは足りない

1. あり 2. なし 3. 記載なし これが必要

曖昧な定義をマニュアル化する

3. 交絡

まとめ

データを集めてからでは手遅れ

解析して意味あるデータを集めること

統計を数字のまやかしにしないこと