

<実証主義的科学観>

True value=observed + bias + error

True value 神のみぞ知る唯一の絶対的存在

Bias 妥当性、疫学の問題

Error 偶然の誤差、確率論、統計の問題

*** 目的のないデータはゴミである**

<研究の流れ>

仮説 デザインデータ データ収集 データクリーニング
(これで論文の善し悪しが決まる)

<論文完成>

仮説が最も大切である！統計は確認のみ

<サンプルサイズの決定法>

effect size(どのくらいの効果が期待されるか)

Type I error level(P=0.05 で良いのが否か)

Type II error level(検出力)

<データクリーニング>

入力確認、ロジックチェック、欠損の問い合わせ、データ訂正、分布の異常、
再度の問い合わせ、データ固定

<中間解析>

- 1 ; プロトコール遵守
- 2 ; データ項目見直し、再定義
- 3 ; 研究継続の可否の検討

<研究成果を脅かすもの>

- 1 ; 選択バイアス (プロトコール違反、途中脱落、欠損値)
- 2 ; 情報バイアス (あり、なし分類など)

まとめ

- 1 ; データを集めてからではすでに手遅れである
- 2 ; 解析して意味のあるデータを集めること
- 3 ; 統計を数字のまやかしにしない

まず retrospective なデータであたりをつけてから prospective なプロトコールをたてる。

以下、余白