

医学論文統計学

1. よい研究に統計学は2の次である

True value = 観察+バイアス+誤差

このうち、誤差を削るために統計学を利用する。つまり、目的なく集められたデータから統計学を用いて何らかの結果が出ることはありえない。

2. 研究の流れ

仮説 デザイン データ収集 データクリーニング 研究論文

仮説を立てた時点で研究の4割が終了している。統計はデータクリーニングに使用するぐらいである。

一番肝心なのは仮説であり、それに見合ったデザインをつくり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そろって初めてよい論文が書ける。

まず、サンプルサイズを決定するがこのとき Effective size が大きいほど少ないサンプルでよい。データクリーニングとは、入力内容の確認、ロジックチェック(ありえない数字)、欠損や誤情報の問い合わせやデータの訂正である。

3. 研究成果を脅かすもの

選択バイアス

プロトコル違反や途中脱落や欠損値

情報バイアス

まとめ

データを集めてから論文を書くことはできない

解析をして意味のあるデータを集める

統計をまやかしに使わない