

不均一性と統計的交互作用【第104回生物統計学】

1 概要

臨床試験は、治療効果を評価するための医学研究です。これらの試験実施には、労力と費用が掛かるので、1つの試験から可能な限りの情報を抽出するために、試験参加者のサブグループ解析が頻繁に行われます。サブグループ解析は、治療効果の不均一性を評価し、将来の研究に有用な情報を提供することができます。しかし、サブグループ解析は、解析上の課題をもたらす、過大評価や誤解を招く結果を招くことがあります。本稿は、2007年に公開された「Statistics in Medicine - Reporting of Subgroup Analyses in Clinical Trials」の内容に基づき、サブグループ解析の実施と報告に関連する課題をまとめます。

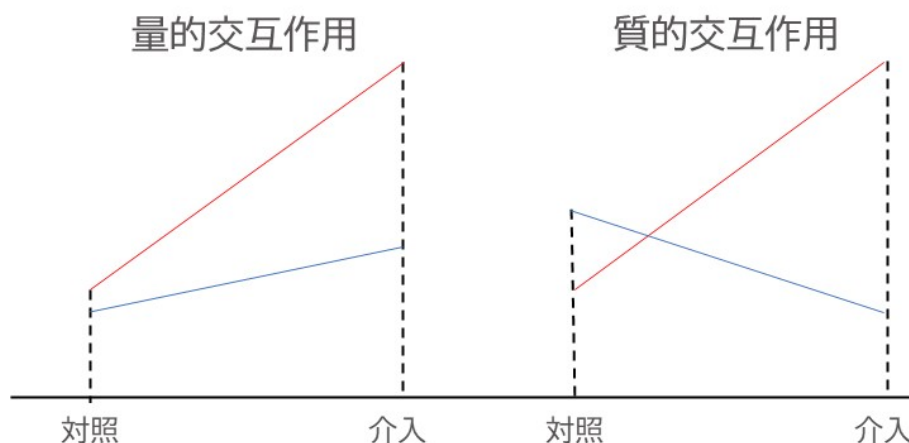
2 不均一性と統計的交互作用

本稿では、「Statistics in Medicine - Reporting of Subgroup Analyses in Clinical Trials」の「Heterogeneity and Statistical Interactions (不均一性と統計的交互作用)」についてまとめます。

2.1 不均一性が臨床試験にもたらす問題と交互作用の検定

試験参加者のベースライン時点の特性に対して介入が及ぼす効果は試験参加者ごとに異なります。臨床試験においてはこうした量的あるいは質的な不均一性が生じます。

臨床試験で生じる不均一性の問題の一つに「交互作用」があります。介入を受けた集団を二つのグループに分けた際、ある介入が、一方のグループに対しては望ましい効果を及ぼすにも関わらず、もう一方のグループに対して同等の効果を示さないことを「量的交互作用」、あるいは逆に望ましくない効果を及ぼすことを「質的交互作用」といいます。





臨床試験においては介入効果の非均一性の検出力が不足しているケースが多く、有意な交互作用の検出ができなければ、全ての試験参加者に対して介入に意義があると判断することができなくなってしまいます。

また解釈にも注意が必要です。犯しがちな間違いとして、サブグループ解析を行った結果推定された効果量から不均一性を判断してしまう場合です。例えば、性別のサブグループ解析を行って男性サブグループのみ介入群と対照群で有意差がついたとしても、介入効果の差が、性別が原因で生じたものなのかという答えにはなっておらず、「女性に対しては介入効果が無い」という仮説を採択することはできません。また、サブグループ解析で観測された介入効果の大きさは推定の精度が低いため、それを理由にして不均一性を判定することはできません。

3 交互作用の検定の例

交互作用の検定の例は以下の通りです。

・交互作用の評価

●●を摂取したことによる心血管疾患への影響を、プラセボ対照試験にて評価した。主要な解析の結果、心血管疾患発症のハザード比は 0.95 であった。

ベースライン時点で××の特性のあったサブグループと、解析対象者のうちそれに該当しないサブグループについてそれぞれハザード比を算出したところそれぞれ 0.80、1.10 であった。

交互作用の検定の結果、 $p = 0.045$ の有意な交互作用であることが示唆された。

4 参考文献

- Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in Medicine - Reporting of Subgroup Analyses in Clinical Trials. The New England Journal of Medicine. 2007; 357: 2189-21948. (PMID: [18032770](https://pubmed.ncbi.nlm.nih.gov/18032770/))

ヒト臨床試験（ヒト試験）で得られる結果は、様々な誤差を含んでいます。この誤差を小さくすることで介入効果を増大させることができます。オルトメディコは、多分野の専門家を有するため、様々なアプローチにより誤差を最小化する試験運営が可能です。引き続き、皆様にご満足いただけるような高品質なヒト試験を提供させていただきますので、今後ともどうぞ宜しくお願い申し上げます。