

サブグループ解析と層別解析【第25回生物統計学】

1 概要

サブグループ解析と層別解析は、解析段階で比較可能性を向上させるために用いられます。サブグループ解析を層別解析と呼んでいるものもありますが、この2つの解析手法は異なるものです。そこで、本稿ではサブグループ解析と層別解析の特徴を紹介します。

2 サブグループ解析と層別解析の違い

2.1 サブグループ解析 (Subgroup analysis)

サブセット解析や部分集団解析とも呼ばれ、サブグループごとに介入の効果を推定します。

例えば、実施したヒト臨床試験（ヒト試験）の対象者を「特定保健用食品の表示許可等について」に則り、「LDL コレステロール値（LDL-C）が境界域者及び軽症域者」とした場合は、LDL-C を要因として、「LDL-C 境界域グループ」と「LDL-C 軽症域グループ」といったサブグループを構築することができます。例として、12週間後の実測値をエンドポイントとして群間比較する際、標本集団全体への解析として群を要因としたモデルを適用する計画を立てたとすると、以下のようなモデルで解析します。

$$\text{観測値} = \text{切片} + \text{群} + \text{誤差}$$

サブグループ解析では、このモデルを「LDL-C 境界域グループ」と「LDL-C 軽症域グループ」にも適用して解析します。そのため、統計学的仮説が単純であり、結果の解釈は一般化可能性などについて追加するくらいであり、説明をしやすい方法となっています。上述の例は、LDL-C の測定値をカテゴリ化しているため、単純ですが、カテゴリ数が多いとサブグループの数が増えてしまい、グループあたりの症例数が減ってしまいます。また、ベースラインの LDL-C を交絡因子として取り扱っていますが、このほかにも交絡因子が考えられる場合はさらにグループ数は増えてしまいます。

2.2 層別解析 (Stratified analysis)

層別解析は、サブグループごとの結果を統合（重み付き平均など）して、1つの介入効果を推定します。上述の例の場合は、以下のようなモデルで解析します。

$$\text{観測値} = \text{切片} + \text{群} + \text{層別因子} + \text{誤差}$$

このモデルで誤差が小さくなるので、群の主効果における検出力が大きくなります。つまり、解析の切れ味が増大します。ここで、群と層別因子の交互作用をモデルに投入することもあります。他には、重みづけ平均

などでサブグループの結果を統合することもあります。このように、集団全体の結果を推定するための方法が層別解析です。しかし、サブグループ解析を層別解析としている場合もあるので文脈などから本意を理解できるようにしておく必要があります。

3 ガイドライン上での取り扱い

表 1. 特定保健用食品の表示許可等について（2020年4月1日付け消食表第109号）の一部

■別添 2 特定保健用食品申請に係る申請書作成上の留意事項，第 2 審査申請書の留意事項，2 審査申請書の添付資料，(4) 食品及び特定の保健の目的に資する栄養成分に係る保険の用途及び一日当たりの摂取目安量を医学的及び栄養学的に明らかにした資料，イ ヒトを対象とした試験，(ウ) 保険の用途に係る有効性等の判定方法
保険の用途に係る…(省略)
被験者が境界域者と軽症者のように複数の層で構成される場合は、原則として層別解析を行う。また、層別解析を行う場合は、各層で有意差検定に十分な被験者数（試験食摂取群、プラセボ食摂取群とともに）を確保する。ただし、被験者が境界域者の一層のみで構成されるなど、上記のおそれがない場合は、層別解析は必要ない。

層別解析という言葉は、サブグループ解析の意味で使用されていることがあります。特定保健用食品のガイドラインは、「層別解析を行う場合は、各層で有意差検定に十分な…(以下省略)」という文章があります。この文言から読み取れる解析方法はサブグループ解析のことであると推察されます。

4 まとめ

今回は、サブグループ解析と層別解析について紹介しました。どちらも比較可能性を高める便利な手法ですが、ヒト臨床試験（ヒト試験）を実施する前に解析計画に定めておくことが重要です。是非、ヒト試験でも活用してみてください。