



プロトコルにサンプルサイズを記述する【第63回生物統計学】

1 概要

ヒト臨床試験（ヒト試験）を計画する際は、プロトコルを事前に作成する必要があります。特に機能性表示食品の届出を目指したヒト試験は、SPIRIT2013に基づく記述が推奨されています。しかし、プロトコルの作成は、医師をはじめとする多くの専門家の参画が必須であり、敷居が高いと感じる方も多いと思います。本稿は、プロトコルを作成するすべての方に向けてプロトコル作りのヒントをまとめます。

2 サンプルサイズ

本稿は、SPIRIT2013のチェックリストにある「サンプルサイズ」についてまとめます。

2.1 サンプルサイズとは

サンプルサイズとは、研究対象の大元の集団である「母集団」からランダムに抽出した「サンプル（標本）」の数のことを言います。

統計学的仮説検定において、検定統計量の算出における手続き上、サンプルサイズが大きくなると有意確率（ P 値）は小さくなります。従って、サンプルサイズが少なすぎると、効果があるのに効果を検出することができないことや、逆にサンプルサイズが大きくなりすぎると、意味のない差を検出することがあります。そこで、意味のある差をできるだけ少ない試験参加者で示すためにサンプルサイズの設計が重要となります。

サンプルサイズと似た用語に「サンプル数」があります。母集団からランダムに抽出した時、サンプル群の数が「サンプル数」、各サンプル群の個数が「サンプルサイズ」となるので注意しましょう。

2.2 サンプルサイズの設計

試験計画における試験参加者数は、試験デザイン、予算、実現可能性の重要な側面であり、通常、正式なサンプルサイズ計算を用いて決定されます。パイロット試験や希少疾患などの理由により、サンプルサイズが統計的に導き出されていない場合は、意図したサンプルサイズの理由とともに、その旨を明示する必要があります。サンプルサイズの設計は、一般的に1つの主要なアウトカムに基づいて行われ、そのアウトカムを高い確率で検出できるように、十分にサンプルサイズを大きくするべきという指針があります。

サンプルサイズ設計をプロトコルに記述する際には、一般的に、結果、各試験群における結果の測定値（例：表1）、アウトカムの検定方法、有意水準（ α ）、検出力（ $1-\beta$ ）、および算出された各試験群のサンプルサイズを含むべきであります。また、各試験群に想定される結果の根拠や参考資料を提示することも推奨されます。

サンプルサイズの設計は、「効果量（ d ）」、「有意水準（ α ）」、「検出力（ $1-\beta$ ）」を用いた方法がよく用いられます。

効果量とは、群間での平均値の差の程度、変数間の関連の強さなど、研究関心の程度を表す値をデータの単位に左右されないよう標準化したものを指します。平均値の場合、Cohen は経験的な解釈として、小さな効果量は $d = 0.2$ 、中程度の効果量は $d = 0.5$ 、大きな効果とは $d = 0.8$ と示しています。計算方法については、参考文献に詳しく記載があるのでそちらをご覧ください。

有意水準 (α) は、第 1 種の過誤 (帰無仮説を誤って棄却する過誤) を犯してしまう確率の水準を示し、検出力 ($1-\beta$) は、第 2 種の過誤 (帰無仮説を誤って採択する過誤) を犯してしまう確率の反対、つまり、正しく帰無仮説を棄却する確率を示します。一般に、 $\alpha = 0.05$ 、 $1-\beta = 0.8$ と設定されています。

効果量、有意水準、検出力、サンプルサイズは互いに影響し、例えば、効果量や有意水準が一定とすると、検出力を高くするとサンプルサイズは大きくなるといった関係があります。この関係を用いて、事前にサンプルサイズを決定する際に検定力分析が行われます。この分析は、手計算で行うと少々煩雑な計算が必要となるので、正確に分析するためにも統計ソフト R の pwr パッケージなどのソフトを用いて分析をすることを推奨します。

2.3 SPIRIT2013 声明

SPIRIT2013 声明にはサンプルサイズについて以下のような記述がなされています。

Sample size	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations
サンプルサイズ	研究目的を達成するために必要な推計参加者数とその決定方法。サンプルサイズのすべての算定のもととなる、臨床的・統計学的仮定を含む。

3 記載例

サンプルサイズの記載例は以下の通りです。

- サンプルサイズは、●●を X 週間摂取することにより▲▲が改善する仮説に基づいて算出された。予備実験において、●●を X 週間後摂取した後の▲▲の実測値の群間差は $\Delta\mu_{1-2}$ であった (被験食品群; $N = n_1, \mu_1 \pm \sigma_1$ 、プラセボ群; $N = n_2, \mu_2 \pm \sigma_2$) (▲▲は高いほうが望ましい)。これらを踏まえて、群間の効果量の Cohen の d は 0.7 と見積った。統計学的有意水準 (α) を両側 5%、統計学的検出力 ($1-\beta$) を 80% とし、 t 検定におけるサンプルサイズを計算すると、各群 Y 名であった。10%の脱力を考慮して、各群 Z 名とした。

サンプルサイズを記載する際は、サンプルサイズを設計するまでに至った、根拠や過程についてしっかりと触れましょう。

4 参考文献

- Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin JA, Doré CJ, Parulekar WR, Summerskill WS, Groves T, Schulz KF, Sox HC, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med.* 2013;158 (3): 200-7. (PMID: [23295957](https://pubmed.ncbi.nlm.nih.gov/23295957/))
- Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jeric K, Laupacis A, Moher D. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ.* 2013; 346: e7586. (PMID: [23303884](https://pubmed.ncbi.nlm.nih.gov/23303884/))
- 長島俊輔. 看護学分野での統計改革を目指して: t 検定における d 族効果量の報告状況とその普及に向けた課題. *日本看護研究学会雑誌*. 2018;41 (5): 1013-9. (DOI: <https://doi.org/10.15065/jjsnr.20180422032>)

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。