

❁ 医療職のための統計シリーズ

医療職のための学び直し—研究デザインから論文報告までの生物統計学の道標—
第16回 データ数に関する議論

ウヰムラ コウヘイ
上村 銅平*

I はじめに

試験治療の有効性を評価するための臨床試験を計画する際には、必要なサンプルサイズを何例に設定すればよいかという、データ数に関する議論を行うことは必要不可欠である。設定したサンプルサイズが不十分である場合、試験治療が真に有効であっても、統計解析の結果から試験としての明確な結論が示されない可能性が高くなる。一方、設定したサンプルサイズが過剰である場合、必要以上に多くの対象者が試験治療の潜在的リスクにさらされるだけでなく、試験の組み入れ期間や試験期間が長期間に及んだり、試験のコストや実施可能性にも問題が生じる。従って、試験の目的を達成するために必要最小限のサンプルサイズを見積もり設定することは、臨床試験を計画する上で最も重要な課題の1つとなっている。

臨床試験の必要サンプルサイズは、試験治療の有効性の大きさやデータのバラつきの大きさに依存する。試験治療の有効性が大きく、対照治療との差が明確である程、必要なサンプルサイズは少なくて済む。一方、試験治療と対照治療との有効性の差に比し、治療のアウトカムの個人差が大きい、つまりデータのバラつきが大きい場合、統計学的に有意な差を検出することは容易ではなく、必要なサンプルサイズも大きくなる。また、必要なサンプルサイズは、アウトカムをどのような指標で測定するか、すなわちエンドポイントを何に設定するか、試験デザインを比較試験とするか単群試験とするか等にも大きく依存する。連続型の変数として測定されるアウトカムよりも、二値アウトカムやTime-to-event型のアウトカムの方が必要なサンプルサイズが多く、長期間の追跡を必要とす

仮説検定の2種の過誤とサンプルサイズ設計

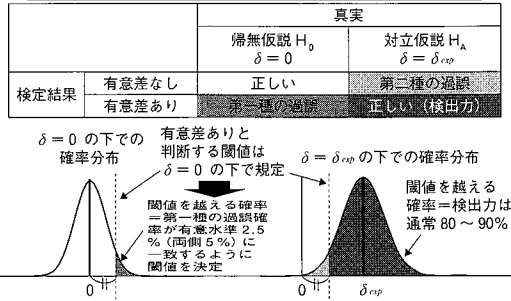


図1 仮説検定とサンプルサイズ設計の関係

る傾向がある。また、一般に単群試験の方が必要なサンプルサイズは少なくて済む。

今回の連載では、必要なサンプルサイズ設計の基礎的な方法論について解説する。また、臨床試験を計画する際に重要な要素となるデータ数の議論について、実際の臨床試験のサンプルサイズ設計の事例を挙げながら解説する。

II 仮説検定とサンプルサイズ設計の基礎

II章では、仮説検定とサンプルサイズ設計の基礎について解説する。図1は、仮説検定の2種類の過誤（確率）とサンプルサイズ設計の関係を表したものである。仮説検定では、実際のデータが期待する仮説の下で得られているかどうか、つまり期待する仮説が真に正しいか否かにかかわらず、比較する群間でアウトカムに差が無いという帰無仮説が正しいと仮定した確率分布をまず考える。そして、実際に観察されたデータ（正確には観察されたデータを含みそれ以上に差があるという方向のデータ）が観測される確率を帰無仮説の下で計算する。この確率がp値である。p値は仮想的な確率であることから、事象が起きている実際の確率というよ

* 東京大学大学院情報学環（生物統計情報学講座）特任講師