

医薬品開発の成功確率を高める 生物統計学の研究分野

上村 銅平

本稿では、著者の主な研究分野の1つである、臨床試験を効率的に実施し新薬の開発成功確率を高めることが可能な臨床試験方法論の開発に関し、研究分野の背景と方法論の概要について紹介する。

医薬品産業はハイリスク・ハイリターン型の産業であり、医薬品の開発には10年以上の時間と数百億～数千億円規模の費用が必要となる^[1, 2]。また、基礎研究の成果として世間の注目を賑わすことも多いリード化合物の発見に対し、最終的に薬事承認に至るのは、703,397化合物中のわずか28化合物(約1/25万, 0.0040%)しかなかったと報告されている^[3]。新薬の開発成功確率は年々低下しており、臨床試験を効率的に実施し成功確率を高めることが可能な臨床試験方法論の開発が生物統計学の分野に求められてきた。新薬の開発はいくつかの段階に分けられるが、どの段階にどの程度のコストがかかっており、どの段階の成功確率がどのような理由により低くなっているのだろうか。

基礎研究の段階で化合物の探索を終了すると、毒性試験、薬理試験を含む非臨床試験が実施され、ヒトへ最初に投与するために必要となる薬効薬理と潜在的なリスクに関する情報を動物実験等により集積する。この段階までのコストイメージは数億円程度である^[2]。非臨床試

験に関する一連の成績に問題が無いことが確認された後に、臨床試験のフェーズが開始する。臨床開発は薬物動態や小数例における安全性・忍容性を調べる臨床薬理試験などを行うフェーズⅠ、治療効果の探索的試験や用量設定試験などを行うフェーズⅡ、有効性と安全性を証明するための検証的試験などを行うフェーズⅢから構成される。フェーズⅢ試験が無事成功すると、非臨床試験成績及び全フェーズの臨床試験成績に対する規制当局の審査を経て、新薬として市場に流通することになる。しかし、上市前の最終段階であるフェーズⅢ試験は、成功確率が低い上、試験費用も開発費用全体の大部分を占めるほど高く、数百億円～1000億円にのぼることから、1製薬企業の命運をかけた大勝負になると言っても過言ではない。ではなぜ、フェーズⅢ試験にここまでコストがかかり、成功させることが難しいのか。

フェーズⅢ試験では、早期のフェーズよりも臨床現場へ近づけた幅広い患者集団において、新薬のベネフィットがリスクを上回ることを証明するための検証的試験の実施が求められる。動物実験により化合物の薬理作用が確認できていても、実際の患者集団へ安全性が担保可能な臨床用量にて投与してみると、想定よりも有効性が低かったということがしばしば起こり

得る。検証的試験のデザインは、プラセボと呼ばれる偽薬と試験薬とのランダム化比較試験であることが一般的であり、ランダムに割付けられた治療群間の治療成績に統計学的に有意な差があることをもって試験薬の有効性を検証することが必須の条件となる。統計学的な有意差とは、ランダム化試験で偶発的に観察され得る範囲に対し、それを超えるほどの治療成績の群間差が認められるか否か、即ち試験薬の有効性が検証されたかどうかを判定するための基準のことである。統計学的な判定が必要になるのは、薬剤に対する反応の個人差や各患者が背景として持っているリスク因子の程度の差によって、実際に観察される群間差が確率的に変動してしまうためである。

統計学的な有意差を無事示せるかどうかは、野球の打率のように確率的な勝負となる。試験薬に真に薬効が存在するのであれば、検証的試験に参加する被験者数を意味する「サンプルサイズ」を大きく設定すればするほど、統計学的な有意差を示せる確率（検出力と呼ぶ）を高めることが可能である。一方、サンプルサイズを大きく設定すればするほど、試験コストは膨らんでいき、また被験者のリクルート期間が延長してしまうことにより、上市のタイミングも遅れていく。上市のタイミングの遅れは、患者さんへ新薬を早く届けられないという社会的な問題だけに留まらない。リード化合物の発見時に取得された特許の期間には限りがある。20～25年の特許期間中は、独占的な製造販売が可能であることから、その期間内は薬価がある程度維持されることになる。しかし、実際には、研究開発期間に約9～17年もの長い年月を要

することから、上市後に一定の薬価にて販売可能な期間は10年程度とごく限られた期間となってしまう^[4]。この間に莫大な研究開発費用を回収できなければ、開発企業のビジネスモデルはたちまち立ち行かなり、新たな新薬を開発するための投資もできなくなってしまう。従って、いたずらにサンプルサイズを大きく設定することは、費用の面でも期間の面でも現実的ではない。そこで、検出力を80%～90%に設定し、その成功確率を達成するために必要最小限のサンプルサイズを設計することになる。

サンプルサイズ設計は、生物統計家が行う最重要業務の1つであるが、一筋縄ではいかない難しさを伴う。前述した試験の成功確率とサンプルサイズや試験期間との兼ね合いをさらに複雑にする要素が存在するからである。その要素とは、試験薬自体の実力の程度である。試験薬の薬効の大きさは、上記の兼ね合いを事前に決定するための最重要なパラメータであるのと同時に、今まさに検証しようとする対象そのものでもあり、試験成績が得られて初めて評価される量でもある。ゆえに、サンプルサイズ設計を間違えないためには試験薬の効果の大きさを正しく知っておくことが必須だが、試験を実施してみないことには、その実力の程度は正直わからないというジレンマが存在することになる。実際のサンプルサイズ設計は、試験薬の実力の程度を予想し、ある種の仮定の下での成功確率をシミュレーションするという作業にならざるを得ない。そこで、試験の途中に実施される中間解析結果の情報から、サンプルサイズを再推定する手法が提案されており^[5]、著者も新規手法の提案を行っている^[6,7]。

サンプルサイズ再推定は、サンプルサイズ不足による試験失敗のリスクを防ぎ、かつ無駄な投資や試験期間の延長を防ぐことで効率的に試験を実施するための臨床試験方法論として開発されてきた。これらのサンプルサイズ再推定の手法は、通常であれば事前に規定すべきパラメータに対し試験データの一部に依存させて試験途中にアダプテーションを行うことが、統計的推測へ与える負の影響を調整することを可能にする。具体的には、中間解析前後の各ステージの検定統計量を用いた重み付き検定統計量を構成したり、条件付き過誤関数による有意水準の調整を行ったりする。検証的試験の実施には、有効性と安全性のバランスが最も優れた臨床推奨用量の選択や適応すべき患者集団の定義 (e.g., ある遺伝子発現によるバイオマーカー陽性の集団へ限定すべきかどうか)、統計的な有意差の判定に用いる主要評価項目の決定など、サンプルサイズ以外にも事前に選択と決定を迫られるいくつかの重要なデザインパラメータが存在する。これらの選択を誤れば、試験の成功確率の低下に繋がるため、慎重な意思決定が求められる。サンプルサイズ再推定と同様に、試験開始前の情報のみで事前に選択または決定した上記の各デザインパラメータに対し、中間解析の結果から得られた情報を活用し、試験の途中にアダプテーションを行うこと自体を事前に計画する臨床試験方法論のことをアダプティブ・デザインと総称し、著者の主な研究分野の1つである。代表的なアダプティブ・デザイン

には、臨床推奨用量の選択を行うフェーズ II 試験と選択された用量による試験薬の有効性を検証するフェーズ III 試験を単一のフェーズにまとめて実施することにより、サンプルサイズの節約と開発期間の短縮を図るアダプティブ・シームレス・フェーズ II/III 試験、バイオマーカーに基づきターゲット集団を限定するか否かを決定するアダプティブ・エンリッチメント・デザイン等があり、適用事例が臨床のトップジャーナルにて紹介されている^[8]。コロナ禍において迅速な治療薬開発が全世界的に強く望まれている状況下で、複数の治療薬の候補を同時並行的にかつ継続的に検証し続けることが可能な、アダプティブ・プラットフォーム試験が大いなる成功を収めた^[9]。非常事態下における医療環境や流行状況の急激な変化に対する適応が随時求められる中、アダプティブ・デザインを応用した革新的な臨床試験デザインは、時代の要請に応える手段として活用された。今後のパンデミックを見据えた開発基盤の整備、国力を左右し得る医薬品産業を活性化させ、新薬の創出を加速させるための新しい技術の開発が求められる中、アダプティブ・デザインを応用した臨床試験方法論の開発が今後ますます重要視される研究分野であると著者は考えている。

以上、アダプティブ・デザインという生物統計学の研究分野が、実際の医薬品開発の効率化へ貢献し得る応用的な方法論であることを紹介した。読者各位に応用統計学の面白さと重要性について認知して貰えれば幸いである。

引用文献

- [1] 医薬産業政策研究所, 「創薬の成功確率分析 – 臨床試験に頂点を当てて –」, 政策研ニュース No.66 (2022 年 7 月)
- [2] 内閣府, 参考資料, 経済産業省.
<https://www8.cao.go.jp/cstp/kyogikai/life/4kai/siryos3-1-2.pdf>
- [3] 日本製薬工業協会調べ (2011~2015 年度).
<https://www8.cao.go.jp/cstp/gaiyo/yusikisha/20180405/siryos9.pdf>
- [4] 国家戦略特区ワーキンググループ, 特許庁ヒアリング配布資料.
- [5] Wang P, Chow SC. Sample size re-estimation in clinical trials. *Statistics in Medicine* 2021; 40(27):6133-6149.
- [6] Uemura K, Matsuyama Y, Ohashi Y. A Modification of the 50%-Conditional Power Approach for Increasing the Sample Size Based on an Interim Estimate of Treatment Difference. *Japanese Journal of Biometrics* 2008; 29(1):19-34.
- [7] Uemura K, Ando Y, Matsuyama Y. Utility of adaptive sample size designs and a review example. *Journal of Statistical Science and Application* 2017; 5:1-15.
- [8] Bhatt DL, Mehta C. Adaptive Designs for Clinical Trials. *The New England journal of medicine*, 2016; 375(1), 65–74.
- [9] Ross JS, Malani P. Learning From the Success of the ACTIV Platform. *JAMA*, 2023; 330(24), 2363–2364.



上村 鋼平 (うえむら こうへい)

[専門] 臨床試験方法論、アダプティブ・デザイン、プラットフォーム臨床試験

[主たる著書・論文]

- ・ Uemura K, Matsuyama Y, Ohashi Y. A Modification of the 50%-Conditional Power Approach for Increasing the Sample Size Based on an Interim Estimate of Treatment Difference. *Japanese Journal of Biometrics* 2008; 29 (1) :19-34.
- ・ Uemura K, Ando Y, Matsuyama Y. Utility of adaptive sample size designs and a review example. *Journal of Statistical Science and Application* 2017; 5:1-15.
- ・ Uemura K, Ono S, Michihata N, Yamana H, Yasunaga H. Duration of effectiveness of the COVID-19 vaccine in Japan: a retrospective cohort study using large-scale population-based registry data. *BMC infectious diseases*, 2024; 24 (1) , 648.

[現在の所属] 情報学環

[所属学会] 日本計量生物学会、日本感染症学会、ヘルスデータサイエンス学会