

シンポジウム2 ●がんゲノム情報をもとに行う治験・臨床研究——実臨床での活用を見据えて

ゲノム情報を活用する臨床試験のデザインと統計解析

国立研究開発法人国立がん研究センター研究支援センター生物統計部

野村 尚吾

がん領域では、遺伝子変異情報に代表されるバイオマーカーに基づく薬剤開発が積極的に進められている。この種の薬剤開発を進めるうえでは、当該バイオマーカーが当該薬剤の有効性が予測できるかどうかに関する事前情報の確信度が重要である。この確信度が高い場合は当該薬剤の効果が認められる部分集団だけを登録する enrichment design を採用すればよいが、確信度が低い場合は全集団を登録する all comers design を採用し、全集団・部分集団に対する検証的な解析計画を組込むなどの工夫が必要になる。実際の臨床試験では、上記に加え、基礎研究の成果や対象の希少性等も考慮し、適切なデザイン・解析手法を選択しなければならない。

以上は医薬品とコンパニオン診断薬を1:1に対応させて開発を進めることを前提としたアプローチ（註：詳細は文献¹⁻⁹⁾を参照）だが、腫瘍検体や血漿から複数の遺伝子異常を一度に検査可能な遺伝子パネル検査が普及したいま、従来のアプローチは効率の観点で必ずしも適切とは限らない。2016年末の21st century cures actの影響を受け、米国では医薬品開発のコスト増を食い止めるための、臨床試験の効率化に向けたさまざまな取組みが始まっている¹⁰⁾。このうち、ゲノム情報を活用した臨床試験への影響が大きいトピックスにマスタープロトコルとリアルワールドデータの利活用がある。本稿の目的はこれらトピックスの概説である。

1 マスタープロトコル

マスタープロトコルにはアンブレラ試験・バスケット試験・プラットフォーム試験がある¹¹⁻¹⁵⁾。以降、この3つの試験デザインを順に説明する。

アンブレラ試験は遺伝子パネル検査を用いた遺伝子スクリーニング研究を入り口とし、同定された遺伝子異常

を対象とするサブ試験に登録する枠組みを組込んだ試験である。コンパニオン診断薬の複数回検査による負担軽減や共通インフラを利用することによる効率化がメリットである。米国 SWOG (Southwest Oncology Group) で実施される SWOG1400 試験 (LUNG-MAP)、英国 MRC (Medical Research Council) で実施される FOCUS-4 試験、わが国の SCRUM-Japan (産学連携全国がんゲノムスクリーニング) プロジェクトなどの例がある。デザイン・統計解析に関しては従来の考え方を踏襲すればよく、アンブレラ試験特有の課題として特筆すべきものはない。アンブレラ試験の課題は、その運営の難しさである。契約・有害事象報告・プロトコル改訂・EDC (electronic data capturing system) 管理・独立データモニタリング委員会の運営など、研究組織は多くの業務を抱え、試験着手の遅れに繋がる。先の LUNG-MAP では、登録中に nivolumab が扁平上皮がんに適応拡大され、試験デザイン的大幅な変更が余儀なくされた。近年、抗体・薬物複合体や免疫チェックポイント阻害剤など、開発の動向がめまぐるしく変化している。アンブレラ試験は単一試験から複数の知見が得られるため魅力的だが、短期間で実施・完遂できるよう工夫しなければならない。

バスケット試験は特定薬剤の効果をがん種横断的に評価する枠組みを指す。複数がん種を単一試験で一まとめに評価する場合はインフラの重複を避けることができ、効率的である。どのがん種でも効果が一樣に認められる場合は個別がん種で独立に評価するより効率的だが、効果の乏しいがん種が含まれる可能性がある。KIT 変異に対する imatinib、MSI-High に対する pembrolizumab、TRK 融合遺伝子に対する larotrectinib など、いずれも単剤の奏効割合を単群で評価する探索的試験が複数実施されている（併用療法やランダム化比較試験の事例は筆者

の知るかぎりない)。デザイン・統計解析の観点では、「似通った効果を示す集団と効果の乏しい集団が混在している」という特徴をうまく考慮することが望ましい。ベイズ流階層モデルは似通った効果を示す集団だけで構成される場合には効果的な手法である^{14,15)}。しかし、*BRAF V600E* 変異を標的とする vemurafenib のバスケット試験のように肺がんで高い奏効割合が確認されたものの、大腸がんでは1例も奏効を認めないような事例¹⁶⁾もある。このような効果の異質性が認められる場合、ベイズ流階層モデルは個別がん種で独立に解析する場合と比べて検出力は同等ないし劣るとする数値実験の結果¹⁷⁾が報告されている。効果の異質性が早期試験で判明することはバスケット試験のメリットである。実際に大腸がんではバスケット試験の治験をもとに *BRAF* 阻害剤・*EGFR* 阻害剤・*PI3K* 阻害剤の併用療法が高い治療効果を発揮している。バスケット試験はアンブレラ試験と比較して、試験運営の負担はそれほど大きくない。参加施設を広く募らないのであれば、試験運営の負担は第Ⅰ相試験と大きく変わらない。限られた誌面から詳細を割愛したが、階層ベイズモデルについては文献^{18~20)}を参照いただきたい。

プラットフォーム試験は複数の新規薬剤を同一試験で一度に多群試験として評価する枠組みである^{15,21)}。ベイズ流解析手法を用い、試験途中の複数回の試験治療群の削除・追加を許容した米国の I-SPY2 試験が有名である。また、英国 MRC でも類似した多群試験 (STAMPEDE 試験) が実施されている。共通の対照群を設けた多群試験デザインは群の数があれば増えるほど効率が上がる²¹⁾。複数薬剤の優劣が1試験で評価できる点もメリットである。デザイン・統計解析に関しては、ベイズ流解析手法を中心に複数の提案がある^{15,22~24)}。アンブレラ試験と同様、試験運営の面で負担が強いられる点が課題である。

なお、アンブレラ試験・バスケット試験・プラットフォーム試験は相互排他な概念ではなく、それぞれを組合わせた試験デザインを採用してもよい。たとえば、アンブレラ試験の遺伝子変異 X に対するサブ試験で他がん種を含めたバスケット試験を採用する (例: SCRUM-Japan プロジェクト傘下の臨床試験) ことも可能である。

2 リアルワールドデータの利活用

希少がんに対する薬剤開発では、ときにランダム化比較試験の実施が困難である。単群試験を実施し、奏効割合を過去対照データから設定した閾値と比較することが多いが、一般に信頼できる過去対照データが存在するのは稀で、閾値の妥当性について疑問が残る場面が少なく

ない。もし、単群試験と似通ったリアルワールドデータ (以降、RWD) が存在すれば、閾値の妥当性の精査や多変量解析等による交絡調整が可能かもしれない。近年、がん領域では、RWD をこのように対照群として利活用できないか、議論が進められている。

RWD を対照群データとして利活用するうえでは、①治験以外のデータソースの信頼性をどのように担保すべきか、②種々の偏りを補正するためにどのようなデザイン・解析手法を採用するか、という2つのポイントが極めて重要である。RWD は一般に日常診療や前向き観察研究として収集されるため、資金や人員の観点から治験と同じ信頼性基準を要求するのは非現実的である。RWD 保有者や申請者が原資料との整合性やデータの入力ミスやロジカルチェックに関して一定の手順を定め、品質を担保することが望ましいが、現状で参照すべきガイドラインはない。承認前の臨床試験で RWD を利活用する際のデータ信頼性に関しては AMED 林班 柴田分担班の成果物「患者レジストリデータを医薬品等の承認申請資料等として活用する場合におけるデータの信頼性担保に関する提言」を、製造販売後臨床試験に関しては 2017 年に改正された GPSP 省令や関連通知^{25,26)}が参考になる。活用に際しては規制当局への相談が必須と考えられる。2点目の統計解析手法に関しては、統計的因果推論の枠組みで提案された傾向スコアが有力な解析手法であるが、適用に際してはデータ収集が治験のように精密でないことやサンプルサイズの制約・欠測データ・選択バイアスなどの影響を評価しなければならない。また、傾向スコア構築を担当する解析担当者の独立性^{27~30)}など、実施体制の観点でも工夫が必要かもしれない。傾向スコア以外に操作変数法^{31,32)}や単一標本に基づく仮説検定を応用した手法^{33,34)}、情報のある事前分布を用いたベイズ流アプローチ^{35,36)}などがある。

ま と め

臨床試験を取り巻く最近のトピックスのうち、マスタープロトコルと RWD 利活用について概説した。いずれも運営上の負担が発生し、効率向上の代償になっている点に留意が必要である。この種の新しい概念は数年内に規制当局のガイダンスや ICH ガイドラインに組込まれる可能性が高い。本稿が今後の動向を注視するきっかけになれば幸いである。

文 献

- 1) Sargent DJ, Conley BA, Allegra C, Collette L. Clinical trial designs for predictive marker validation in cancer treatment trials. *J Clin Oncol* 2005; 23 (9): 2020-7.

- 2) Mandrekar SJ, Sargent DJ. Clinical trial designs for predictive biomarker validation: theoretical considerations and practical challenges. *J Clin Oncol* 2009; 27 (24): 4027-34.
- 3) Freidlin B, McShane LM, Korn EL. Randomized clinical trials with biomarkers: design issues. *J Natl Cancer Inst* 2010; 102(3): 152-60.
- 4) Simon R. Clinical trials for predictive medicine. *Stat Med* 2012; 31 (25): 3031-40.
- 5) Freidlin B, Sun Z, Gray R, Korn EL. Phase III clinical trials that integrate treatment and biomarker evaluation. *J Clin Oncol* 2013; 31 (25): 3158-61.
- 6) Matsui S, Choai Y, Nonaka T. Comparison of statistical analysis plans in randomize-all phase III trials with a predictive biomarker. *Clin Cancer Res* 2014; 20 (11): 2820-30.
- 7) Freidlin B, Korn EL. Biomarker enrichment strategies: matching trial design to biomarker credentials. *Nature reviews Clinical oncology* 2014; 11 (2): 81.
- 8) Simon RM. Genomic clinical trials and predictive medicine: Cambridge University Press; 2013.
- 9) Manji A, Brana I, Amir E, et al. Evolution of clinical trial design in early drug development: systematic review of expansion cohort use in single-agent phase I cancer trials. *J Clin Oncol* 2013; 31 (33): 4260-7.
- 10) Administration USFaD. Submission to Congress: Food & Drug Administration Work Plan and Proposed Funding Allocations of FDA Innovation Account. U. S. Food and Drug Administration; 2017.
- 11) Woodcock J, LaVange LM. Master Protocols to Study Multiple Therapies, Multiple Diseases, or Both. *N Engl J Med* 2017; 377 (1): 62-70.
- 12) Renfro LA, An M-W, Mandrekar SJ. Precision oncology: A new era of cancer clinical trials. *Cancer Lett* 2017; 387: 121-6.
- 13) Renfro LA, Mandrekar SJ. Definitions and statistical properties of master protocols for personalized medicine in oncology. *J Biopharm Stat* 2017; 1-12.
- 14) Thall PF, Wathen JK, Bekele BN, et al. Hierarchical Bayesian approaches to phase II trials in diseases with multiple subtypes. *Stat Med* 2003; 22 (5): 763-80.
- 15) Berry D. Emerging innovations in clinical trial design. *Clin Pharmacol Ther* 2016; 99 (1): 82-91.
- 16) Hyman DM, Puzanov I, Subbiah V, et al. Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. *N Engl J Med* 2015; 373 (8): 726-36.
- 17) Freidlin B, Korn EL. Borrowing information across subgroups in phase II trials: is it useful? *Clin Cancer Res* 2013.
- 18) LeBlanc M, Rankin C, Crowley J. Multiple histology phase II trials. *Clin Cancer Res* 2009; 15 (13): 4256-62.
- 19) Simon R, Geyer S, Subramanian J, Roychowdhury S, editors. The Bayesian basket design for genomic variant-driven phase II trials. *Semin Oncol*; 2016: Elsevier.
- 20) Cunanen KM, Iasonos A, Shen R, et al. An efficient basket trial design. *Stat Med* 2017; 36 (10): 1568-79.
- 21) Freidlin B, Korn EL, Gray R, Martin A. Multi-arm clinical trials of new agents: some design considerations. *Clin Cancer Res* 2008; 14 (14): 4368-71.
- 22) Berry DA. The Brave New World of clinical cancer research: adaptive biomarker-driven trials integrating clinical practice with clinical research. *Mol Oncol* 2015; 9 (5): 951-9.
- 23) Harrington D, Parmigiani G, Equipoise FB, et al. Adaptive randomization of neratinib in early breast cancer. *N Engl J Med* 2016; 2016 (375): 1591-4.
- 24) 野村尚吾, 柴田大朗. I-SPY2 試験: デザイン・解析手法の解説. *腫瘍内科 (Clinical oncology)* 2017; 20 (3): 239-44.
- 25) Zelen M. A new design for randomized clinical trials. *N Engl J Med* 1979; 300 (22): 1242-5.
- 26) Altman D, Whitehead J, Parmar M, et al. Randomised consent designs in cancer clinical trials. *Eur J Cancer* 1995; 31 (12): 1934-44.
- 27) Li H, Yue LQ. Statistical and regulatory issues in nonrandomized medical device clinical studies. *J Biopharm Stat* 2007; 18(1): 20-30.
- 28) Yue LQ. Regulatory considerations in the design of comparative observational studies using propensity scores. *J Biopharm Stat* 2012; 22 (6): 1272-9.
- 29) Yue LQ, Campbell G, Lu N, et al. Utilizing national and international registries to enhance pre-market medical device regulatory evaluation. *J Biopharm Stat* 2016; 26 (6): 1136-45.
- 30) Yue LQ, Lu N, Xu Y. Designing premarket observational comparative studies using existing data as controls: challenges and opportunities. *J Biopharm Stat* 2014; 24 (5): 994-1010.
- 31) Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. *Pharmacoepidemiol Drug Saf* 2010; 19 (6): 537-54.
- 32) Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Stat Med* 2014; 33 (13): 2297-340.
- 33) Makuch RW, Simon RM. Sample size considerations for non-randomized comparative studies. *J Chronic Dis* 1980; 33 (3): 175-81.
- 34) Dixon DO, Simon R. Sample size considerations for studies comparing survival curves using historical controls. *J Clin Epidemiol* 1988; 41 (12): 1209-13.
- 35) Hobbs BP, Carlin BP, Mandrekar SJ, Sargent DJ. Hierarchical commensurate and power prior models for adaptive incorporation of historical information in clinical trials. *Biometrics* 2011; 67 (3): 1047-56.
- 36) Ibrahim JG, Chen MH, Gwon Y, Chen F. The power prior: theory and applications. *Stat Med* 2015; 34 (28): 3724-49.