

### 第三相マーカー層別試験の改良：サブグループ別逐次デザイン

がんの分子標的薬や免疫チェックポイント阻害薬等の中には作用機序が十分明白であり、治療標的等の状態を正確に高い再現性・頑健性をもって測定できる分子マーカー（予測マーカー）が利用可能なものがある。このようなケースでは予測マーカーによって規定される患者サブ集団（マーカー陽性集団）に限定したエンリッチメント試験を行うのが効率的である（ランダム化の患者数を少なくできるという意味で）。しかし、実際にはこれに該当しないケースも多い。例えば、マーカーの分析的妥当性や陽性判定のカットオフの妥当性がはっきりしないケース、別の作用経路による効果（オフターゲット効果）が十分想定されるケースなどである。これらのケースでは、マーカー陰性の患者でも一定の効果が出現する可能性を考える必要が生じる。このとき、検証的試験である第三相試験をどうデザインしたらよいだろうか？ 一つのアプローチは、マーカー陽性患者のエンリッチメント試験を先に実施し、そこで有望な結果が得られれば、マーカー陰性患者の（エンリッチメント）試験を後追いで実施する直列型エンリッチメントデザインである[1, 2]。別のアプローチは、マーカー陽性・陰性患者を同時に（concurrent に）評価する all-comers なデザイン、特に、マーカーを層別因子に含めたマーカー層別デザインである [3, 4]。

最近バスケット試験などのマスタープロトコール試験に代表されるように、エンリッチメントを基軸とした臨床試験の枠組みが注目され、その中で前者の直列型デザインが好まれる傾向があるようである。その際、concurrent にマーカー陰性患者も対象とする all-comers デザインは非効率的（さらに非倫理的）と、ひどく批判されることもある（この傾向は年々強くなっているように感じられる。ただし、従来型のマーカーを用いない all-comers 試験も併せて批判されていることもあるので注意が必要）。

そこで、本研究は後者の all-comers マーカー層別デザインに焦点をあて、その効率の改善を試みた。これまでのマーカー層別試験は、古典的な試験デザインのように（マーカー値とは関係なく）全患者集団を評価するという側面がそもそも強く、マーカー層別デザインを考える上でもこのことが色濃く反映されることが多い。このことは、例えば、試験対象集団におけるマーカーのプレバレンス（マーカー陽性の割合）が一般集団のそれとほぼ等しくなるようなデザイン（例えば、[3, 4]）にみることができる。これに対して、本研究では、マーカー層別試験を「サブグループの評価を中心としたもの」に組み直す。そのねらいは、1) 直列エンリッチメントデザインとの比較がやりやすくなること（そもそもこのことが本研究の動機）もあるが、より重要なことは、2) エンリッチメントを基軸とする新しい臨床試験の枠組みによりフィットしたソリューションを提示することである。

これより、提案するデザインはマーカー陽性・陰性集団に対して別々のデザインを考えるものとなる[5]。

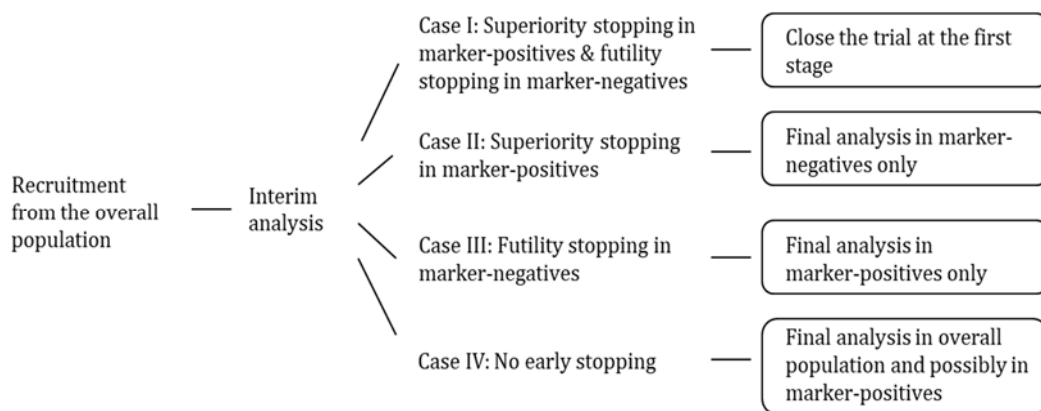


図 1. 提案するマーカー層別デザイン

マーカー陽性患者に対しては、マーカー陽性集団のエンリッチメントデザイン、マーカー陰性患者に対しては、(マーカー陽性で効果ありが示されたもとの) マーカー陰性集団のエンリッチメントデザインを考える。これは直列型デザインを **concurrent** に実施することと基本的には同じである。このとき、マーカー陽性集団に対して有効中止を意図した中間解析を入れても良い。

一方、マーカー陰性集団に対しては、無効中止を意図した中間解析を導入することが重要である。マーカー陰性患者に対しては、臨床的に意味の効果サイズを規定した無効中止基準を用意する(無効中止基準として一般に用いられるものよりもフォーマルな位置づけとなる)。

その一方で、**concurrent** な **all-comers** 試験のメリットは、全集団での治療効果を効率的に評価できることにあるので、マーカー陽性・陰性集団両方で中止がなかった場合には、全体をプールしての **overall** 検定(あるいは、**overall** 検定とサブ集団検定を組み合わせた多重検定)を実施することとする。

提案するデザインと直列型デザインを数値実験、多発性骨髄腫の臨床試験事例で比較したところ、提案デザインのパフォーマンスは様々なシナリオのもとで概ね良好であった。直列型デザインは、トータルの試験期間が長くなること、マーカーサブ集団間で治療効果に大きな差がない場合に効率が悪いこと、という欠点が認められた。直列型デザインを選択するにはこの点に注意が必要である。これに対して、提案デザインはマーカー層別試験における有力な選択肢になると考えられる。

文献：

1. Fridlyand J, Simon RM, Walrath JC, Roach N, Buller R, Schenkein DP, et al. Considerations for the successful co-development of targeted cancer therapies and companion diagnostics. *Nat Rev Drug Discov.* 2013; 12(10): 743-55.

2. Liu A, Liu C, Li Q, Yu KF, Yuan VW. A threshold sample-enrichment approach in a clinical trial with heterogeneous subpopulations. *Clin Trials*. 2010; 7(5): 537-45.
3. Mandrekar SJ, Sargent DJ. Clinical trial designs for predictive biomarker validation: theoretical considerations and practical challenges. *J Clin Oncol*. 2009; 27(24): 4027-34.
4. Simon R. Clinical trial designs for evaluating the medical utility of prognostic and predictive biomarkers in oncology. *Personalized Medicine*. 2010; 7(1): 33-47.
5. Matsui S, Crowley J. Biomarker-stratified phase III clinical trials: Enhancement with a subgroup-focused sequential design. *Clin Cancer Res*. 2018; 24: 994-1001.