

# 希少疾患患者を対象とした臨床試験におけるAdaptive Designの提案

## A proposal of adaptive design of clinical trial for rare disease patients

17N3100038H 堀口 剛 (生物統計学研究室)

Go HORIGUCHI/ Biostatistics Lab.

**Key Words :** Adaptive design, Adaptive randomization, Rare disease, Clinical trials

### 1. 緒言

Adaptive Designとは、試験中に蓄積されていくデータに基づいて試験デザインの変更を行う臨床試験デザインであるため、一般的な固定デザインに比べより柔軟なアプローチを開発企業や研究者に提供することができる。こうしたAdaptive Designの目的の一つは、試験の効率性の向上である。効果が低いあるいは安全性に問題があるような試験治療は早めに開発を中止し、よい試験治療をできるだけ早く次のステージに移行させることが、臨床試験実施の効率性を考える上で重要である<sup>1)</sup>。こうした背景から、早期中止を前提としたAdaptive Designが注目されている。その中でもアウトカム適応的ランダム化は、途中の試験結果に適応しながら次の割付けを決める方法であり、より良い治療群へより多くの患者を割付けることを可能にする。しかし、臨床試験設定の多様性から、一般論としてアウトカム適応的ランダム化の有用性を議論することには限界があるため、より具体的な状況を想定した上での評価が必要であるが、試験が患者の相当な割合を含むことが前提となるような希少疾患に対する臨床試験では、アウトカム適応的ランダム化はより多くの試験参加者に恩恵を与えるため、より望ましいとする主張がある<sup>2)</sup>。また、希少疾患における臨床試験は、サンプルサイズ不足の問題から実施可能性の観点で多くの課題があり、試験デザインを工夫することが重要である<sup>3)</sup>。そこで本研究では、希少疾患患者を対象とした小規模な臨床試験におけるAdaptive Designの応用を提案し、その動作特性をシミュレーションによって評価することを目的とする。

具体的には原発性リンパ浮腫を対象疾患とし、アウトカム適応的ランダム化を含む、試験参加者にとってより有益な試験デザインを提案する。原発性リンパ浮腫は難治性かつ希少性の高い疾患であり、根本的治療法は存在しない。マッサージおよび運動・圧迫療法が標準治療となっているのが現状であり、遺伝子治療を含む新規療法の早急な開発が求められている。本研究の結果は、これらの開発研究において将来実施されるであろう臨床試験のデザインに関する重要な知見になると考えられる。

### 2. 研究方法

#### (1) 試験デザインの概要

試験デザインに影響を及ぼす原発性リンパ浮腫の疾患上の特徴は、既知で確立した治療法が存在せずプラセボの使用が可能であること、病態が安定しているため持ち越し効果が強くなければクロスオーバー試験も可能であること、比較的短期に治療効果が評価できることである。また、考えられる試験治療については、用量反応性はあるものの、効果が同じであれば副作用やコストの問題からできるだけ低用量を使いたいという背景があるため、複数の用量を検討可能な試験デザインとする。用量反応試験(第II相試験)と検証的な第III相試験を継続して行うことが困難であるのは希少疾患一般に言えることである。したがって、本研究の検討は、上記の特徴を満たす用量の検討を伴う希少疾患を対象とした臨床試験に一般的に当てはまると言える。なお、本研究では理論的な検討を深めるため、他の疾患での適用等の事前情報により、用量は2つに絞られていると想定する。

#### a) 試験デザイン①

試験開始時点で実薬群(高用量、低用量)とプラセボ群にランダム化し、効果的な用量を探索した後、新たな被験者を追加(ランダム化)した上でプラセボ群との比較により、安全性および有効性を検証する。最終解析は、プラセボ群とのカイ二乗検定(片側2.5%)により行う。なお、多重性の補正にはDunnnettタイプを適用する。その際、各群のサンプルサイズを考慮し棄却限界値をSASにより算出する。ランダム化は、被験者のアウトカムに応じて割付け確率を変更するアウトカム適応的ランダム化を行う。随時、無効性に関する決定ルール(事後確率に基づく)に関して評価を行い、無効と判断された治療群についてはその時点で中止する。

#### b) 試験デザイン②

試験前半にクロスオーバーデザインを採用することで効率的にプラセボに対する効果の検証を行い、後半で用量選択を行う。クロスオーバー部分の解析は残存効果および期間効果はないものと仮定した上でマクネマー検定(片側10%)により比較を行い、最終解析はカ

イ二乗検定（片側2.5%、多重性の補正：サンプルサイズを考慮したDunnettタイプ）によりプラセボ群との比較で行う。試験後半のランダム化には、アウトカム適応的ランダム化によって行う。ただし、前半のクロスオーバー部分で得られた結果は使用せず、後半から登録された被験者の結果のみを用いてランダム化確率を決定する。クロスオーバー終了後の解析において、有意差が得られなかった場合はその時点で試験を中止する。

## (2) 事前分布と事後分布

3群試験において、 $p_k$  ( $k = 1, \dots, 3$ )を治療  $k$  に関する反応率とする。各患者に関して、確率  $p_k$  で値1、確率  $1 - p_k$  で値0をとる2値アウトカムを観測する。ここでは各  $p_k$  に関して、 $\alpha_k, \beta_k$  を超パラメータとするベータ事前分布を仮定する。 $N_k$  を群  $k$  に関して計画された最大標本サイズとし、 $Y_k$  を  $n_k$  ( $1 \leq n_k \leq N_k$ ) 人の被験者のうちの反応数とすると、 $\text{Bin}(n_k, p_k)$  を試行数  $n$ 、注目する事象（今の場合、反応）を  $p$  とした2項分布として  $Y_k \sim \text{Bin}(n_k, p_k)$  となる。ベータ分布と2項分布との間の共役性に従い、 $p_k$  の事後分布は

$$p_k | y_k \sim \text{Beta}(\alpha_k + y_k, \beta_k + n_k - y_k)$$

となる。

## (3) 適応的ランダム化

適応的ランダム化とは、途中の割付け結果に適応しながら次の割付けを決める方法であり、各群に関する割付け確率が試験の進行につれて変化することを許容する。これは、共変量適応的とアウトカム適応的の2種類に分けられるが、本研究ではより良い治療群へより多くの患者を割付けることを目的とするため、アウトカム適応的ランダム化による割付けを採用した。

2群へのランダム化を考える場合、アウトカム適応的ランダム化では一般に各治療群の推定反応率に比例した割付け確率を用いることが多いが、このような割付け確率は反応率の点推定値のみに依存しており、これらの分散（バラツキ）を考慮していない。頻度流においてこうしたバラツキを考慮することは簡単ではないが、ベイズ流アプローチでは、患者を次の事後確率で治療群1に割付けることが可能となる<sup>4)</sup>。

$$\lambda = \Pr(p_1 > p_2 | D)$$

ここで、 $D$ は試験での蓄積データを表す。上式は、 $p_1$  および  $p_2$  の事後分布を比較することで、自動的に治療反応率の点および分散推定値の両方を考慮することになる。多群（本研究では3群）の場合は、各  $p_k$  をある固定された参照確率  $p$  と比較することでランダム化確率を決定する固定参照型適応的ランダム化と、比較集団における治療の反応率の平均を参照確率とする移動参照型適応的ランダム化の2通りの方法があるが、ここでは

より多くの患者を最良の治療に効率的に割付けることができる後者を採用した。

## (4) 無効性に関する決定ルール

群1の反応率が群2の反応率よりも大きい事後確率  $\Pr(p_2 > p_1 | y)$  が、

$$\Pr(p_2 > p_1 | y) < \theta$$

であれば、治療2は治療1よりも有望でないと主張できる。ここで、カットオフ確率  $\theta$  は0.05とした。プラセボ群に対する実薬群の劣性が上式によって示された場合、試験薬の無効性を根拠に試験を中止することとした。

なお、上式の計算には積分を要するため、モンテカルロ法を用いることで事後確率を算出した。本研究では、2つの確率変数の比較に関する計算のみなので、台形法による積分も使用可能であったが今後の拡張も考慮し一般的なモンテカルロ法を採用した。

## (5) シミュレーション設定とシナリオ

提案した臨床試験デザインの動作特性を評価するために、先述の2つの試験デザインおよび、3群へのランダム化確率を固定し単純ランダム化を行う（均等ランダム化）デザインに関して、後述の4つのシナリオを設定した上でシミュレーションを行った。考慮する群は3つ（プラセボ群、高用量群、低用量群）であり、アウトカムは治療効果の有無の2値変数とした。したがって、各群に対して治療の反応率の真値をシナリオとして与えることになる。全被験者数は、3つの試験デザインを通して30人とした。適応的ランダム化によって群間に偏りが生じる可能性があるため、試験デザイン①では、はじめの10人を均等ランダム化とした。試験デザイン②では、クロスオーバー部分に15人の被験者を割り振った。

シナリオについては、低用量群の効果が高用量群を上回ることはないという仮定の下、プラセボ群の反応率を0.2に固定し、シナリオ1では高用量群、低用量群ともに効果がない場合、シナリオ2では高用量群、低用量群ともに効果がある場合、シナリオ3では高用量群のみに効果がある場合、シナリオ4では高用量群、低用量群ともに効果はあるが高用量群の方が効果が大きい場合とした。なお、シミュレーションの反復回数は500とし、すべての解析にはSAS ver9.4を用いた。

## (6) 均等ランダム化デザインとの比較

各デザインに関してシナリオごとに、各群に割付けられた被験者数の平均、治療効果のあった被験者割合の平均、検出力あるいはタイプ I エラー率を算出した。この際、検出力は、対立仮説に対応するシナリオの下でシミュレートされた試験から評価でき、これらの成功試験の割合を求めることで算出した。一方、タイプ I エラー率は、帰無仮説に対応するシナリオの下でシ

ミュレートされた試験から評価でき、これらの成功試験の割合を求めることで算出した。成功試験の定義については、均等ランダム化デザインおよび試験デザイン①では、2つの用量群のうちどちらか一方あるいは両方がプラセボ群と比較して有意だった場合とし、試験デザイン②ではクロスオーバー終了時の検定が有意かつ試験後半においてどちらか一方あるいは両方の用量群がプラセボ群と比較して有意だった場合とした。

### 3. 結果

3つの試験デザインそれぞれに対して、各群の平均被験者数、治療効果のあった被験者の平均割合、検出力をシナリオごとに算出した結果を表-1に示した。各群の平均被験者数について、均等ランダム化デザインでは、当然ながら3つの群に均等に被験者が割付けられた。一方、試験デザイン①では、反応率が群間で差のないシナリオ1を除き、反応率の低い群に割付けられた被験者数は均等ランダム化に比べて減少し、逆に反応率の高い群に割付けられた被験者数は増加した。また、試験デザイン②の用量選択部分についても同様に、低用量群より高用量群の方が反応率が高いシナリオ3およびシ

ナリオ4において、高用量群に割付けられた被験者数が平均して多かった。

治療効果のあった被験者の平均割合について、3群とも反応率が0.2のシナリオ1では、均等ランダム化デザインと試験デザイン①が約20%であったのに対し、試験デザイン②は約36%と大きかった。シナリオ2,3,4では、共通して均等ランダム化デザイン、試験デザイン①、試験デザイン②の順に割合が大きくなるという結果であった。特にその差は、シナリオ3において顕著であり、試験デザイン②の治療効果のあった被験者割合は均等ランダム化デザインに比べ約2倍であった。

帰無仮説に適合するシナリオ1の下での成功試験の割合は、3つの試験デザインで0.025以下であり、各群の反応率が0.2というシナリオの下ではタイプ I エラー率が制御されていることが確認できた。対立仮説に適合するシナリオ2-4の下での成功試験の割合は、試験の検出力を意味するが、試験デザイン①の検出力は均等ランダム化デザインに比べ全体的に低かった。一方、試験デザイン②の検出力は均等ランダム化デザインとほぼ同等であった。

表-1 シミュレーション結果

シナリオ	デザイン	平均被験者数[人]			治療効果のあった被験者の平均割合	成功試験の割合	被験者数 [人]
		プラセボ	低用量	高用量			
1	反応率の真値	0.2	0.2	0.2			30
	ER <sup>*1</sup>	10.11	9.81	10.08	0.20	0.01	
	デザイン① <sup>*2</sup>	10.28	9.55	10.17	0.17	0.01	
	デザイン② <sup>*3</sup>	15	—	15	0.36	0.00	
	—	7.71	7.29				
2	反応率の真値	0.2	0.6	0.6			30
	ER	10.01	10.06	9.93	0.47	0.47	
	デザイン①	4.99	12.18	12.83	0.49	0.25	
	デザイン②	15	—	15	0.62	0.44	
	—	7.40	7.60				
3	反応率の真値	0.2	0.2	0.6			30
	ER	10.04	9.82	10.14	0.34	0.34	
	デザイン①	6.11	5.84	18.05	0.42	0.21	
	デザイン②	15	—	15	0.59	0.36	
	—	4.33	10.67				
4	反応率の真値	0.2	0.6	0.8			30
	ER	10.14	10.03	9.82	0.53	0.76	
	デザイン①	4.36	8.36	17.28	0.62	0.60	
	デザイン②	15	—	15	0.74	0.76	
	—	5.48	9.52				

\*1 ER：各群への割付確率が等しい均等ランダム化デザイン

\*2 デザイン①：アウトカム適応的ランダム化を採用した試験デザイン（用量選択→プラセボ検証）

\*3 デザイン②：前半にクロスオーバーを採用した試験デザイン（プラセボ検証→用量選択） なお、上段がクロスオーバー部分、下段が用量選択部分の結果を表す

## 4. 考察

原発性リンパ浮腫患者を対象とした臨床試験のデザインを2つ提案し、その動作特性をシミュレーションにより評価した。その結果、適応的ランダム化を導入したデザインでは、より効果のある治療群により多くの患者が割付けられた。さらに、クロスオーバーを採用したこと、および無効性による中止基準を設けたことにより、治療効果のあった患者の割合は大きく増加した。しかし、被験者数が群間でバラついたことにより、試験デザイン①の検出力は減少した。タイプ I エラーは、適切に制御されていることが確認できた。

提案した2つの試験デザインのシミュレーション結果から、患者への恩恵が大きくなることが示された。したがって、試験参加者に対する利益の大きさという観点では、適応的ランダム化を含む Adaptive Design の有用性は非常に大きいと言える。しかし、適応的ランダム化の是非については先行研究において広範囲な議論がなされており、サンプルサイズや検出力なども考慮に入れた総合的な評価が重要になる。ただし、希少疾患の患者を対象とした臨床試験という枠組みでは、サンプルサイズが制限されることに加え、試験が全患者の相当な割合を含むことになるため、適応的ランダム化の倫理的優位性は高い<sup>2)</sup>。本研究で提案した試験デザイン②においては、タイプ I エラーを制御し、均等ランダム化デザインと同等の検出力を維持した上で、試験参加者に対する恩恵を大きく増加させているため、限られたサンプルサイズの中でより効率的にかつ倫理的に試験を実施するという目的を達成する可能性が示された。試験デザイン②において検出力が改善したのは、クロスオーバーによって実質的なサンプルサイズが増加したためであり、被験者利益の観点からも可能であるならばクロスオーバーの採用を検討することの意義は大きいと言える。持ち越し効果の可能性がある場合は、最初の治療の効果を取り除くためのウォッシュアウト期間を含めるか、delayed design と呼ばれる方法を採用し、1期目の成績を主として解析し、1期においてプラセボが割付けられていた被験者の中で実薬とプラセボの比較を補助的に行うことが考えられる。

Adaptive Design を含むような複雑なデザインでは試験全体のタイプ I エラー率を解析的に推定することは困難となるため、シミュレーションによる評価が必要となる。したがって、本研究で提案した Adaptive Design を含む臨床試験では、複数のシナリオに対してシミュレーションを実施し、タイプ I エラーが制御されることを保証した上で、実際の臨床試験へ応用されることになる。今回は2つの試験デザインを提案したが、試験実

施に関する制約条件が変われば、それに基づく無数の試験デザインが考えられるため、あらゆる可能性を踏まえた上で事前のシミュレーションを実施し、あらかじめ設定した複数の基準に沿ってより適切なデザインを選択していくことを考えるべきである。

本研究はいくつかの限界点を含む。これらを以下に示す。1つはサンプルサイズが制限されていることによる検出力不足である。最終的な意思決定にもベイズ流の検定手順を取り入れ、事前分布として何らかのデータや情報を活用することができれば、検出力は改善するであろう。2つ目はデザイン設定の際の仮定についてである。本研究では、患者の登録速度とアウトカムの観測速度に関して特別な設定はせず、試験デザインをシミュレートしたが、実際の臨床試験では患者登録の速度は一定にはならない上、登録されてからアウトカムが測定されるまでにも時差があるため、本研究で実施したシミュレーション通りの結果が得られない可能性があることは考慮すべきである。

今後の臨床試験では、Adaptive Design を採用するかどうかの判断も含めて状況に応じた試験デザインを複数考え、シミュレーションによる評価を通してより適切なデザインを選択していくことが求められる。原発性リンパ浮腫については、現在（2019年1月）医師主導での遺伝子治療のパイロット試験が進行中であり、本研究で考察したような Adaptive Design を採用した試験が実施される可能性がある。

## 5. 結論

原発性リンパ浮腫の患者を対象とした臨床試験のデザインを2つ提案し、その動作特性をシミュレーションにより評価した結果、Adaptive Design を応用することによって試験参加者への恩恵が大きくなることが示された。アウトカム適応的ランダム化の実施に伴い検出力は減少するものの、クロスオーバーデザインを採用することで効率性が向上し検出力は改善した。試験全体のタイプ I エラーについては、どのデザインでも一定水準以下に制御できていることが確認できた。

## 文献

- 1) 森田智視. ベイズ流試験デザインの紹介. 臨床薬理. 2009;40:125-126.
- 2) Lee J, Chen N, Yin G. Worth adapting? Revisiting the usefulness of outcome-adaptive randomization. *Clin Cancer Res*. 2012;18:4498-4507.
- 3) Gupta S, Faughnan ME, Tomlinson GA, et al. A framework for applying unfamiliar trial designs in studies of rare diseases. *J Clin Epidemiol*. 2011;64:1085-1094.
- 4) Yin G, 手良向聡, 大門貴志訳. 臨床試験デザインベイズ流・頻度流の適応的方法. 株式会社メディカル・パブリケーションズ. 2014.