置換ブロック法による被験者割付におけるブロックサイズ混合手法と固定手法の性能比較

○上原 歩夢[1]、室永 遼太郎[1]、菊池 宏和[1]

([1]シミック株式会社)

Performance Comparison of Mixed and Fixed Block Size Methods in Permuted Block Randomization: A SAS Simulation Study

Uehara Ayumu[1], Muronaga Ryotaro[1], Kikuchi Hirokazu[1]
[1]Data Science Division, CMIC Co., Ltd.

要旨

置換ブロック法による被験者割付において、予見可能性排除に推奨されるブロックサイズ混合手法と、ブロックサイズを固定した場合の性能をシミュレーションにより比較する。

キーワード：割付表、置換ブロック法、ブロックサイズ、無作為化、層別化、モンテカルロシミュレーション、PROC PLAN

1. はじめに

1.1 背景、目的

無作為化比較試験（Randomized Controlled Trial: RCT）では、治療群間の比較可能性（背景因子の等質性）を確保するため、ランダム化による被験者割付が実施される。特に被験者数が限られる治験や多施設共同試験では、群間の被験者数不均衡を抑制する目的で置換ブロック法（Permuted Block Randomization）が広く利用されている。置換ブロック法では、一定数の被験者が登録されるごとに治療群間の均衡が維持されるという利点がある一方で、ブロックサイズが固定されている場合には割付の予見可能性が高まり、選択バイアスの発生につながる可能性が指摘されている。この予見可能性の問題に対し、ICH E9 ガイドライン（2.3.2 節）では複数のブロックサイズをランダムに組み合わせる手法（ブロックサイズ混合手法）が紹介されている。一方で、ブロックサイズ混合手法は割付表の作成に手間を要するため、ブロックサイズを固定した場合と比べて、群間均衡性および検出力の観点でどの程度の性能差が生じるのかを把握しておくことは、実務上の手法選択において重要である。

本発表では、割付因子を設定した場合の割付手法である層別置換ブロック法を用いて層別無作為化を行う状

況を想定し、ブロックサイズが固定の場合および混合の場合について、割付比が計画から乖離する「可能性の大きさ」を定量化することを目的とする。

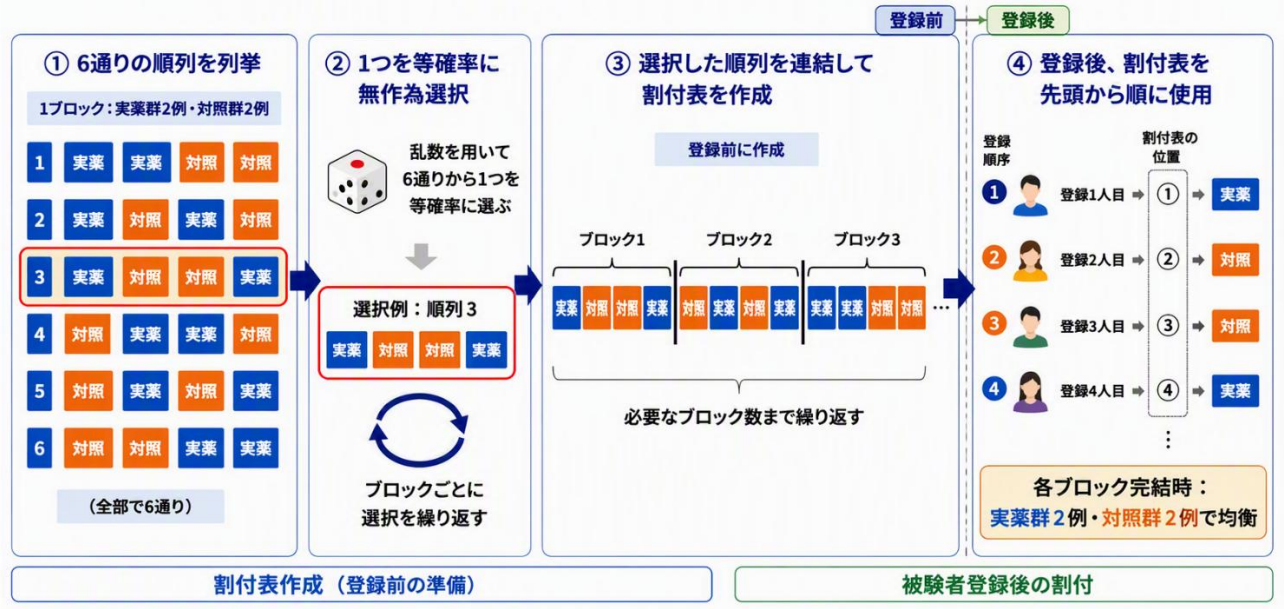
1.2 置換ブロック法の基礎

置換ブロック法はブロック内の割付順序をあらかじめランダムに並べ替えることで群間の被験者数偏りを一定範囲に抑える手法であり、ブロックサイズを固定すると割付の予見可能性が高まるという課題がある。また、割付因子（層別因子）を設定する場合には、層ごとに独立した割付表を作成する層別置換ブロック法が用いられる。層ごとに割付表を作成することにより、各層の内部で治療群間の被験者数の均衡が保たれるため、割付因子の分布が治療群間で偏ることを防ぐことができる。以下、ブロックサイズを固定・混合それぞれについて具体的な数値例で割付表作成の手順を示す。

1.2.1 ブロックサイズが固定の割付表

ブロックサイズ 4（実薬群 2 例・対照群 2 例）を例として、割付表作成および被験者割付の手順を示す。
まず、1 つのブロック内で実薬群 2 例および対照群 2 例を配置する順列をすべて列挙する。ブロックサイズ 4 では、実薬群 2 例・対照群 2 例の並び方は全部で 6 通り存在する。
次に、6 通りの順列の中から乱数を用いて等確率に 1 つを選択し、割付表に追加する。
この「順列の選択」と「割付表への追加」を、必要なブロック数に達するまで繰り返すことにより、あらかじめ治療群の並びが定められた 1 つの割付表が作成される。
被験者の登録後は、作成済みの割付表を先頭から順に使用する。すなわち、1 人目の登録被験者には割付表の 1 番目に記載された治療群を、2 人目には 2 番目の治療群を割り付ける。この操作を登録順に繰り返すことで、各ブロックが完結した時点で実薬群・対照群の被験者数が等しくなるように割り付けられる。

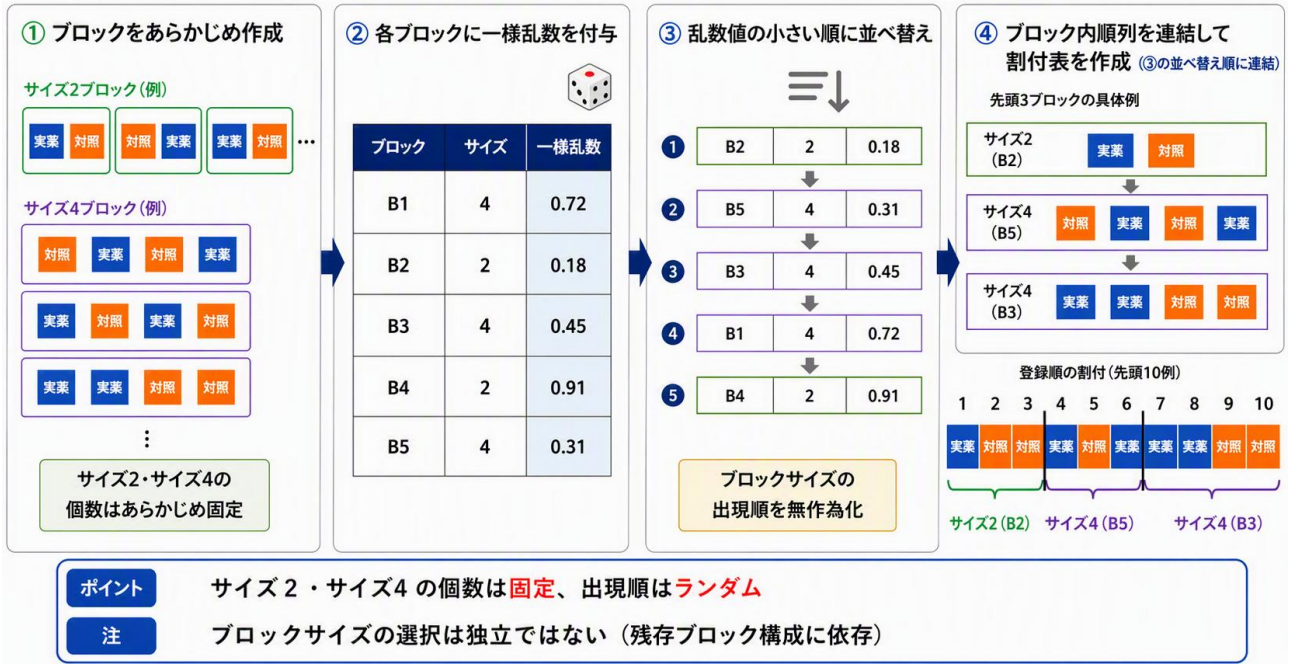
図 1. ブロックサイズが固定の場合の割付表作成



1.2.2 ブロックサイズが混合の割付表

本発表で用いるブロックサイズ混合手法では、ブロックサイズ 2 およびブロックサイズ 4 のブロックをあらかじめ所定数作成し、それらのブロックをランダムな順序で連結する。すなわち、あらかじめ各ブロックサイズの個数を固定したうえで、その出現順を無作為化する方式である。具体的には、作成した全ブロックに対して独立な一様乱数を付与し、乱数値の小さい順にブロックを並べ替える。並べ替え後の順に各ブロックを連結することにより、ブロックサイズ 2 および 4 がランダムな順序で出現する割付表を作成する。この方法を採用する場合のメリットとして、割付表の大きさをコントロールしやすいことがあげられる。
本方式では、サイズ 2 ブロックおよびサイズ 4 ブロックの個数をあらかじめ固定している一方、各ブロックの出現順はランダムである。したがって、ブロックサイズの出現順は無作為化されるが、各ブロックにおけるサイズ選択は独立ではない。例えば、すでにサイズ 2 ブロックが多く出現している場合には、残りの割付表におけるサイズ 4 ブロックの割合が相対的に高くなる。

図 2. ブロックサイズが混合の場合の割付表作成



2. 方法

2.1 割付表の作成

被験者割付には層別置換ブロック法を用いた。SAS の PROC PLAN を用いて、以下の 3 条件について割付表を作成した。SAS での実装は、SAS Version 9.4 を使用した。

- 固定ブロックサイズ 2（すべてのブロックのサイズを 2 とする）
- 固定ブロックサイズ 4（すべてのブロックのサイズを 4 とする）
- ブロックサイズ混合（サイズ 2 とサイズ 4 のブロックを混在させる）

PROC PLAN は実験計画および無作為割付表の作成に用いられるプロシジャであり、factors 文で割付の構造（割付因子の層数、ブロック数、ブロック内順列数など）を指定し、treatments 文で 1 ブロック内に並べる割付枠の数（key 値の水準数。ブロックサイズ 4 であれば key=4）を指定することで、ブロック内の並び順を無作為に入れ替えた割付表を生成する。被験者数 100 例に対応する割付表として、固定ブロックサイズ 2 では 50 ブロック、固定ブロックサイズ 4 では 25 ブロック、ブロックサイズ混合ではサイズ 2 を 20 ブロックおよびサイズ 4 を 15 ブロック生成した。ブロックサイズ混合では、作成した全ブロックに独立な一様乱数を付与し、乱数値の小さい順にブロックを並べ替えたうえで、その順に各ブロックを連結することにより、サイズ 2 および 4 のブロックがランダムな順序で出現する割付表を作成した。なお、PROC PLAN が出力する key 値はブロック内での通し番号であり、本発表では key 値 1 および 3 を実薬群、2 および 4 を対照群に対応させた（ブロックサイズ 2 のブロックでは、key 値 1 が実薬群、2 が対照群となる）。

SAS PROC PLAN のサンプルコード

例：ブロックサイズ 4 に固定した層別置換ブロック法の割付表

```
proc plan seed=20260918;
    factors strata=10 ordered block=25 ordered ban=4;
    treatments key=4;
    output out= plan4;
run;
```

各ステートメントの意味は以下である。

factors 文：割付表の構造を指定する。

- strata=10 ... 割付因子の層を 10 水準作成する。
- block =25 ... 1 層あたり 25 ブロックを配置する。
- ban = 4 ... 1 ブロックの割付枠(順列)を 4 つにする。
- ordered ... 層番号(strata)・ブロック番号(block)は昇順にする。

treatments 文：ブロック内の key 値 1〜4 を無作為な順序で割り当てる。

出力されたデータセット plan4 (図 3)を見ると、treatments ステートメントに指定した key という変数がランダムに並ぶ。この値に対して、群を設定する。

例： key = 1 or 3 → 実薬群、key = 2 or 4 → 対照群

図 3.データセット plan4 の一部抜粋

	strata	block	ban	key
1	1	1	2	4
2	1	1	3	1
3	1	1	1	2
4	1	1	4	3
5	1	2	3	2
6	1	2	1	4
7	1	2	2	3
8	1	2	4	1
9	1	3	1	3
10	1	3	4	2
11	1	3	3	4
12	1	3	2	1

2.2 シミュレーション条件

被験者数は 100 例とした。
割付の層数については 10・25・50・100・200 層の 5 条件を設定した。
層別置換ブロック法を用いるため、割付表は層ごとに 100 例分作成した。
各条件について 10,000 回のモンテカルロシミュレーションを実施し、各試行において割付表を独立に再生成した。
各被験者の割付は、(i)どの層に登録されるか、および(ii) 登録された層内でどの割付順序を引くか、という 2 種類の乱数によって決まる。本発表では(i)を各層へ等確率で登録されるものとして設定し、(ii)は前述の PROC PLAN およびブロックの並べ替えに用いた乱数により独立に生成されている。

2.3 評価指標

割付性能の評価指標として、全体割付差、割付比 1:1 からの乖離率(%)、および割付の偏りが実務上問題となる水準に達する確率（割付比が 2:1 以上となる確率、ならびに割付比 1:1 から 10%以上および 20%以上乖離する確率）を用いた。

・全体割付差

全体割付差は、

$$|N_{treat}\text{-}N_{control}|$$

により定義した。
ここで

- N_{treat}：実薬群の被験者数
- $N_{control}$：対照群の被験者数

を表す。

・割付比 1:1 からの乖離率および割付の偏りに関する確率指標

割付比 1:1 からの乖離率(%)は、全体割付差を総被験者数で除して百分率化した指標であり、（全体割付差／総被験者数×100）により定義した。また、実務上の意思決定に資する指標として、割付比が 2:1 以上に達する確率（いずれかの群の被験者数がもう一方の群の 2 倍以上となる試行の割合)、および割付比 1:1（各群 50 例）から 10%以上・20%以上乖離する確率に着目した。ここでいう「10%以上の乖離」とは、上記の乖離率が 10%以上となること、すなわち被験者数 100 例においては両群の被験者数の差が 10 例以上（55 例対 45 例以上の偏り）となることを指す。同様に「20%以上の乖離」は、両群の差が 20 例以上（60 例対 40 例以上の偏り）となることを指す。

・層ごとの平均割付差（層内不均衡の指標）

層別無作為化を行った試験では、割付で用いた層別因子を解析でも調整することが原則とされている（ICH E9, 5.7 節)。これは層別因子が真の治療効果の推定に影響し得る交絡因子となり得るためであり、層内で群間の比較可能性（均衡性）が保たれて初めて、この交絡の影響が適切に制御される。したがって、全体の被験者数比が 1:1 に近い場合であっても、個々の層内で実薬群・対照群の被験者数が大きく偏っていれば、割付因子が重要な予後因子である場合には交絡を十分に抑制できていない可能性がある。そこで本発表では、全体割付差とは別に、層ごとの割付差の絶対値を、実際に 1 例以上の被験者が登録された層についてのみ合計した「層別割付差の総和」を試行ごとに算出したうえで、これを登録があった層数（S^*）で除した「1 層あたりの平均割付差」を算出した。すなわち、1 層あたりの平均割付差は次式により定義される。

$$\frac{1}{S^*}\sum_{s:N_{treat,s}+N_{control,s}>0}\left|N_{treat,s}-N_{control,s}\right|$$

ここでS^*は実際に 1 例以上の被験者が登録された層の数（登録層数)、$N_{treat,s}$および$N_{control,s}$は第s層における実薬群および対照群の被験者数を表す。登録が 0 例の層は、分子・分母のいずれからも除外される。

2.4 全体割付差が検出力に与える影響の評価

割付比の偏り（全体割付差）が検定の検出力にどの程度影響するかを評価するため、各試行で得られた群別の被験者数から理論的な検出力を算出した。ここでの検出力は、2 群の平均値の差を対応のない t 検定で比較する状況を想定し、非心 t 分布を用いて算出したものである。

まず、被験者数 100 例が両群に均等（各群 50 例）に割り付けられた理想的な場合に検出力がちょうど 80%となる標準化効果量として、

$$d=0.5659$$

を設定した（有意水準は両側 5%とした)。

そのうえで、各試行で実際に得られた群別の被験者数を用いて非心パラメータを算出し、それに対応する検出力を求めた。これにより、割付が均等からずれた場合に検出力がどの程度低下するかを定量化することができる。

具体的には、2 群の t 検定において、両群の被験者数をそれぞれn_1、n_2、標準化効果量をdとすると、検定統計量は自由度n_1+n_2-2、非心パラメータλの非心 t 分布に従う。ここで非心パラメータは次式で与えられる。

$$\lambda = \frac{d}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

各群 50 例（全体割付差 0）の場合は$\lambda = 2.830$ となり、有意水準両側 5%における棄却限界値と比較することで、検出力 80.0%が得られる。一方、効果量 d の値を一定にしたうえで、割付が偏って全体割付差が 10（55 例対 45 例）となった場合など、群ごとの被験者数に差が生まれると検出力が低下する。割付差が大きくなるにつれて検出力がどのような推移になるかを確認するため、3.2 節で検出力曲線を作成した。

3. 結果

3.1 割付差の比較

全体割付差および 1 層あたりの平均割付差の平均値、標準偏差、最大値を表 1 に示す。1:1 からの平均乖離率は被験者数を 100 例に固定しているため、割付比 1:1 からの乖離率(%)は全体割付差を 100 で除して 100 倍した値、すなわち全体割付差の平均値と同じ数値になることに注意する。すべての指標について、固定ブロックサイズ 2、ブロックサイズ混合、固定ブロックサイズ 4 の順に大きくなる単調な傾向が全ての層数条件で一貫して観察された。また、全体割付差の平均値・標準偏差はいずれも層数が増えるほど増加した。1 層あたりの平均割付差についても、平均値は層数が増えるほど増加したが、標準偏差は逆に層数が増えるほど減少した。

表 1. 全体割付差および 1 層あたりの平均割付差の平均値・標準偏差・最大値

層数	条件	全体割付差の平均（=1:1 からの平均乖離率）(標準偏差)	全体割付差の最大	1 層あたりの平均割付差の平均 (標準偏差)	1 層あたりの平均割付差の最大
10	固定ブロックサイズ 2	1.69 (1.49)	8	0.50 (0.16)	1.00
10	ブロックサイズ混合	2.01 (1.71)	10	0.60 (0.19)	1.40
10	固定ブロックサイズ 4	2.17 (1.86)	12	0.66 (0.20)	1.40
25	固定ブロックサイズ 2	2.74 (2.20)	14	0.51 (0.10)	0.88
25	ブロックサイズ混合	3.27 (2.62)	16	0.61 (0.12)	1.04
25	固定ブロックサイズ 4	3.61 (2.84)	18	0.68 (0.13)	1.12
50	固定ブロックサイズ 2	3.98 (3.10)	18	0.57 (0.07)	0.83
50	ブロックサイズ混合	4.67 (3.63)	22	0.68 (0.09)	1.00
50	固定ブロックサイズ 4	5.27 (4.07)	26	0.79 (0.09)	1.13
100	固定ブロックサイズ 2	5.18 (3.98)	32	0.68 (0.06)	0.88
100	ブロックサイズ混合	5.80 (4.43)	30	0.77 (0.06)	1.02
100	固定ブロックサイズ 4	6.53 (4.94)	30	0.88 (0.07)	1.18
200	固定ブロックサイズ 2	6.34 (4.80)	30	0.80 (0.04)	0.96
200	ブロックサイズ混合	6.69 (5.18)	36	0.86 (0.05)	1.01

200	固定ブロックサイズ 4	7.30 (5.55)	40	0.93 (0.05)	1.13

続いて、1:1 の割付比率から 10%以上・20%以上乖離する確率および割付比が 2:1 以上に達する確率を表 2 に示す。2:1 以上の極端な不均衡は層数 200 の固定ブロックサイズ 4 条件でも最大 0.07%と、実務上ほぼ生じない水準にとどまった。一方、10%以上の乖離は層数が多い条件（100・200）ではいずれの条件でも 16〜33%程度の試行で発生しており、層数の設定が割付比の乖離抑制に関わる実務上の検討事項であることが示唆された。

表 2. 割付比 1:1 から 10%以上・20%以上乖離する確率および割付比が 2:1 以上に達する確率（%）

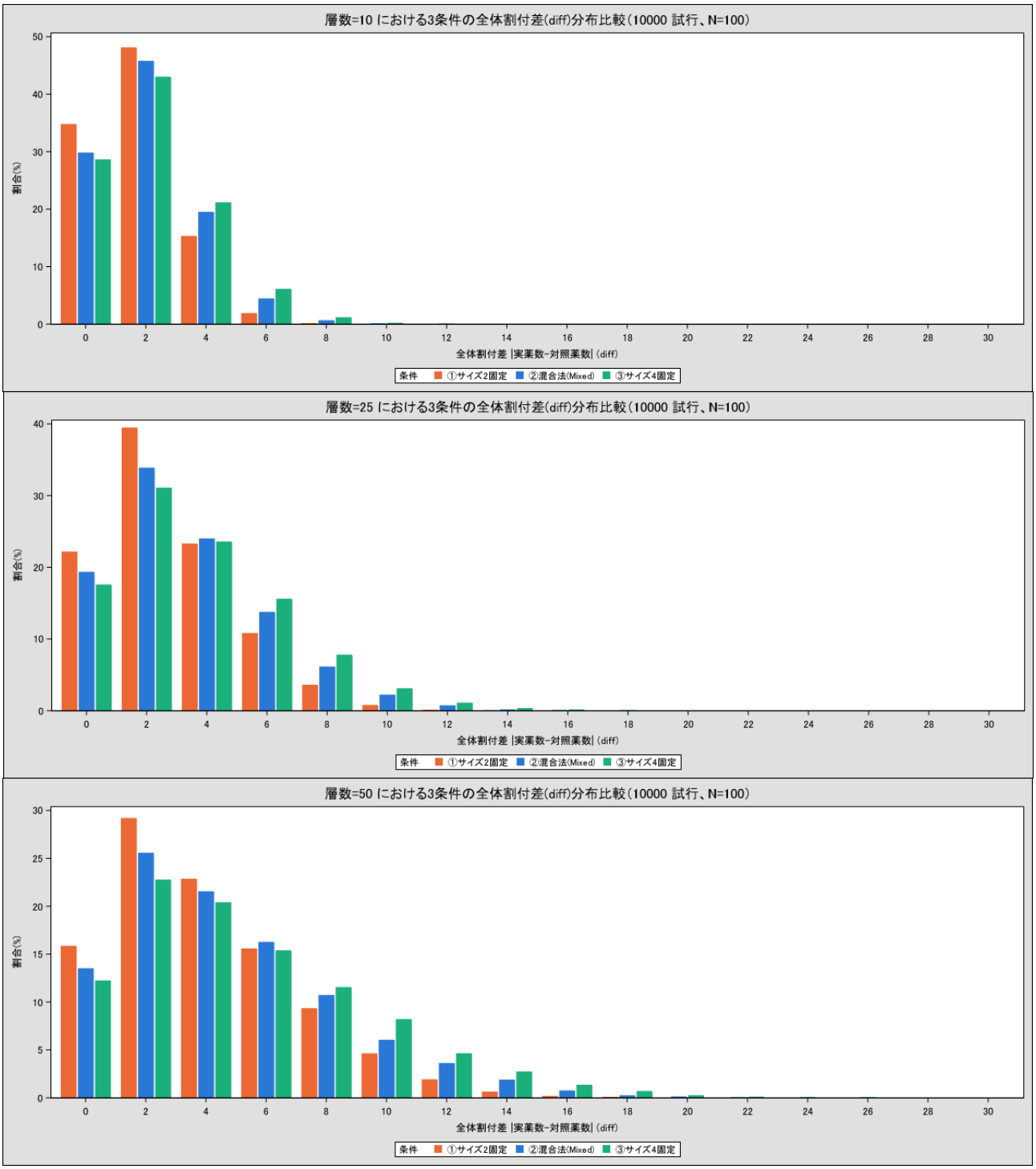
層数	条件	10%以上乖離確率(%)	20%以上乖離確率(%)	2:1 以上到達確率(%)
10	固定ブロックサイズ 2	0.00	0.00	0.00
10	ブロックサイズ混合	0.05	0.00	0.00
10	固定ブロックサイズ 4	0.17	0.00	0.00
25	固定ブロックサイズ 2	0.83	0.00	0.00
25	ブロックサイズ混合	3.04	0.00	0.00
25	固定ブロックサイズ 4	4.57	0.00	0.00
50	固定ブロックサイズ 2	7.32	0.00	0.00
50	ブロックサイズ混合	12.52	0.12	0.00
50	固定ブロックサイズ 4	17.83	0.35	0.00
100	固定ブロックサイズ 2	16.27	0.36	0.00
100	ブロックサイズ混合	21.48	0.93	0.00
100	固定ブロックサイズ 4	27.45	1.89	0.00
200	固定ブロックサイズ 2	26.06	1.65	0.00
200	ブロックサイズ混合	28.01	2.69	0.02
200	固定ブロックサイズ 4	32.59	3.70	0.07

注：確率は 10,000 回のモンテカルロシミュレーションにおける観測割合である。0.00%は 10,000 回中 0 回を表し、真の発生確率が 0 であることを意味しない

次に、5 つの層数条件それぞれについて、3 条件の全体割付差の分布を図 4 に示す。いずれの層数条件でも、固定ブロックサイズ 2 の分布は小さい割付差の領域に集中し、固定ブロックサイズ 4 は右方（大きな割付差側）に分布が広がり、ブロックサイズ混合は両者の中間的な分布を示した。

また、層数条件別・3 条件別の 1:1 の割付比率からの乖離率(%)の分布の箱ひげ図を作成した（図 5）。中央値・四分位範囲ともに固定ブロックサイズ 2 で最小、固定ブロックサイズ 4 で最大となり、層数が増えるほど外れ値の頻度も増加する傾向が認められた。

図 4. 層数別（10・25・50・100・200）× 3 条件の全体割付差（図中の diff）の分布比較



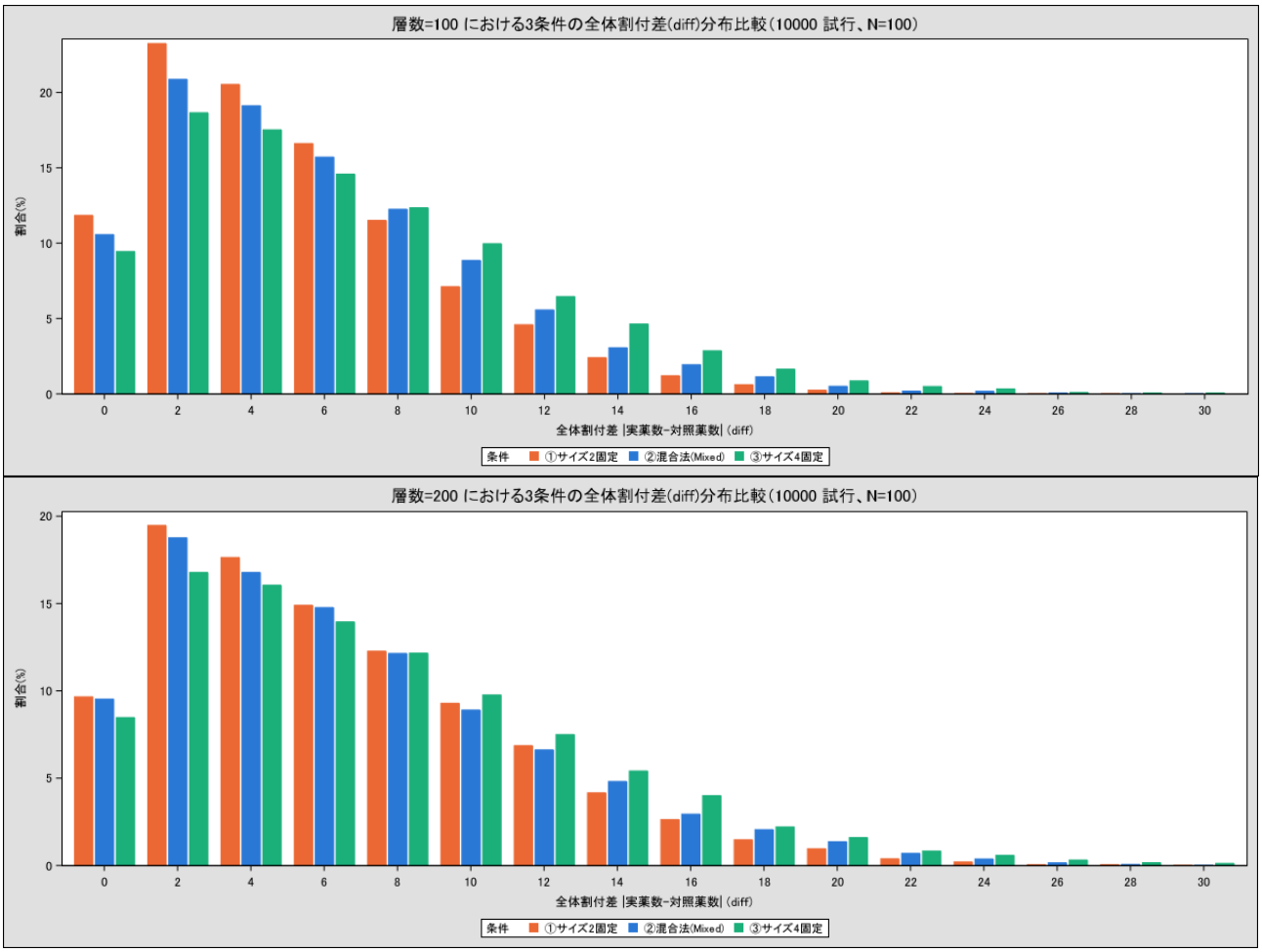
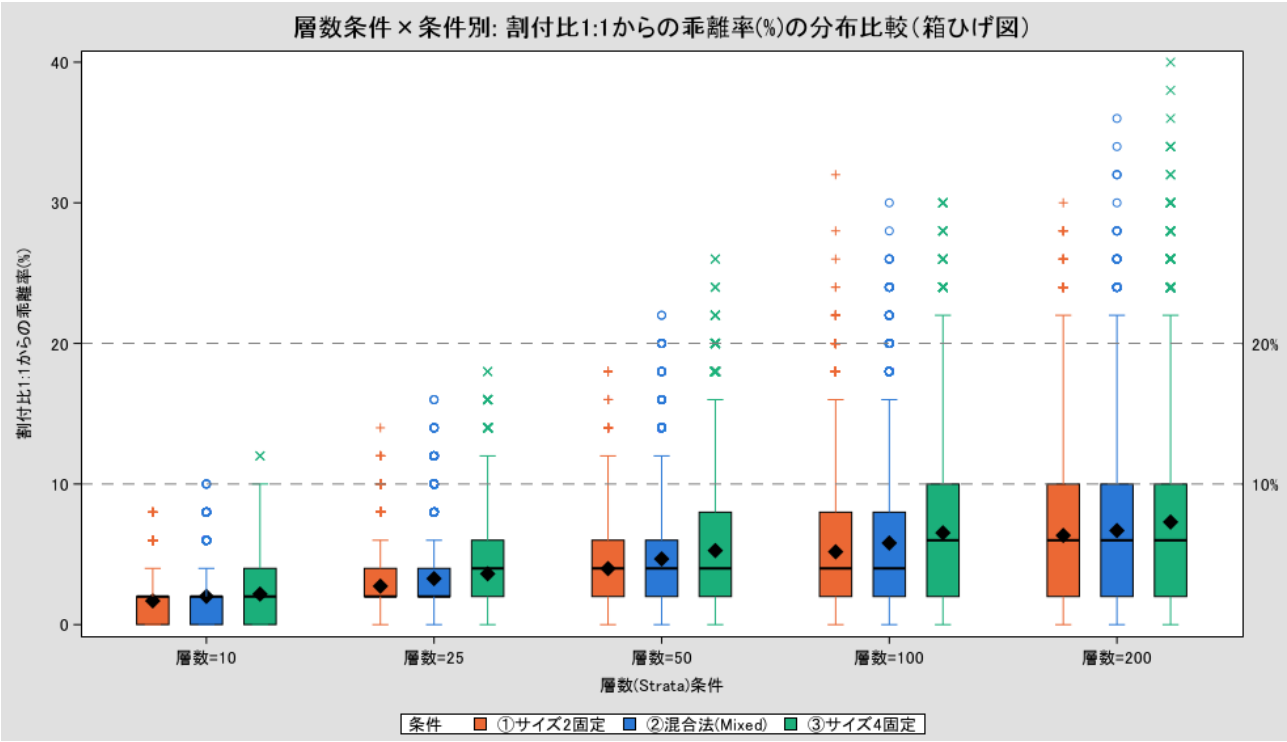


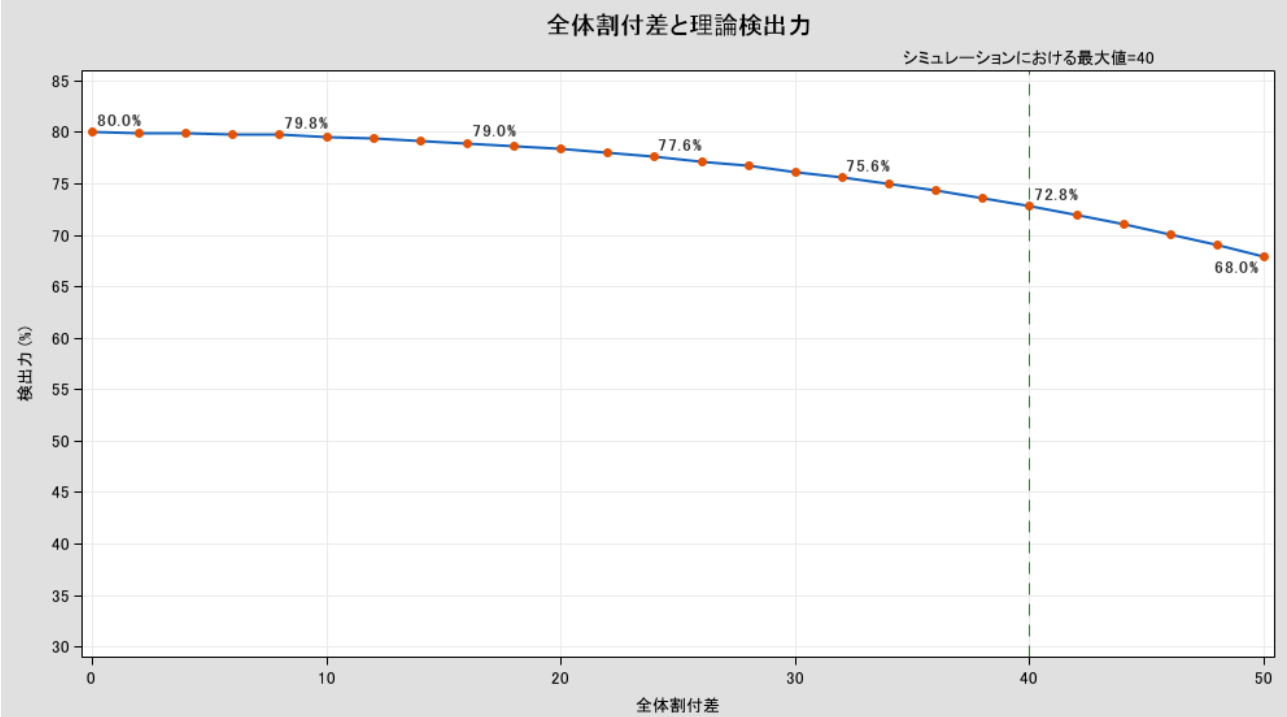
図 5. 層数別・3 条件ごとの割付比 1:1 からの乖離率(%)の箱ひげ図



3.2 全体割付差が検出力へ与える影響

図 6 に、全体割付差と理論検出力の関係を示す（理論検出力は 2.4 節で述べた非心 t 分布に基づき算出した）。
全体割付差 10 で理論検出力 79.6%、全体割付差 20 で 78.4%であった。
今回のシミュレーションで観測された最大の全体割付差（層数 200・固定ブロックサイズ 4、全体割付差 40）でも理論検出力は 72.8%であり、全体割付差 0（80.0%）からの低下幅は 7.2 ポイント程度に留まった。
一方、各条件の平均的な割付差（表 1 の全体割付差の平均、いずれも 10 未満）に基づく理論検出力は 79.8〜80.0%の範囲であり、条件間の差はほぼ認められなかった。

図 6. 全体割付差と理論検出力の関係（理論検出力曲線）



4. 考察

4.1 割付差の比較

3.1 節で示したとおり、全体割付差の平均値は固定ブロックサイズ 2、ブロックサイズ混合、固定ブロックサイズ 4 の順に単調に増加した。これは置換ブロック法の定義上、ブロックサイズが小さいほど群間均衡がより早期に達成されることから理論的に予想される結果と整合的である。ただし、最大値（10,000 試行中の極端な実現値）は必ずしも平均値と同じ順位になるとは限らない点に留意が必要である。全体割付差の標準偏差が層数の増加とともに拡大したことは、層数が多い試験ほど、割付を実施したときにどの程度の不均衡が生じるかが試験ごとにばらつきやすい、つまり割付差を事前に見通しにくいことを意味する。

また、層数が被験者数に対して相対的に多い条件（100・200）では 1 層あたりの平均登録被験者数が極めて少なくなり（層数 100 で 1 例、層数 200 で 0.5 例）、ブロックが完成しない層の割合が高くなる。この未完成ブロックの偏りによる影響が、1 層あたりの平均割付差、全体割付差および乖離率のばらつき（図 5 の外れ値増加）が層数の増加とともに大きくなったことの要因と考えられる。
一方、1 層あたりの平均割付差の標準偏差は層数が増えるほど縮小した。これは、平均を取る対象となる登録層数自体が層数条件の増加とともに大きくなるためである。

4.2 全体割付差と検出力

3.2 節の結果、全体割付差の平均値に基づく理論検出力は、いずれの層数条件においても 3 条件とも 79.8〜80.0%の範囲にとどまり、条件間の実質的な差はほとんど認められなかった。
一方、各条件で観測された最大割付差に基づく理論検出力は 72.8〜79.8%の範囲となり、層数が多い条件ほど低下する傾向が認められたが、これらは 10,000 試行中の極端な実現値である。この 80%からの低下幅（7.2 ポイント）は割付差が理想値（0）から最悪の実現値まで変動した場合の低下幅であり、全体割付差の平均同士の比較（79.8〜80.0%の範囲）では条件間の差はほぼ認められない。今回のシミュレーションの結果からは、割付差が検出力に与える影響は小さいと考えられる。
ただし、検出力への影響が小さいことをもって割付差を無視してよいわけではない。割付差が大きくなった場合には、実薬群に割り付けられる被験者数が計画より少なくなる可能性がある。実薬群の被験者数が治験実施計画書で予定していた数を大きく下回ると、有害事象の発現割合の推定精度が低下するなど、安全性プロファイルの確認に影響を及ぼす可能性がある。したがって、割付差のリスクは検出力の観点のみで評価するのではなく、安全性評価に必要な実薬群の被験者数が確保されるかという観点も含めたうえで、ブロックサイズや層数の設定、割付状況のモニタリングなど、適切な対応を検討すべきである。

4.3 層別不均衡の実務的留意点と本発表の限界

本発表では、層を解析で調整しない場合の層別不均衡が検出力へ及ぼす影響は検証していないが、層内の割付不均衡が大きい場合には層別因子による交絡を十分に抑制できていない可能性がある点に留意が必要である。また、表 2 のとおり層数が被験者数に対し相対的に多い条件では割付比 1:1 から 10%以上乖離する試行が 16〜33%発生しており、層数の設定（層あたりの登録数の確保）がブロックサイズの選択以上に割付比の乖離抑制に重要と考えられる。
本発表の主な限界として、被験者数 100 例・割付比 1:1 に固定しており他の条件での性能は未検証である点、およびブロックサイズ混合の本来の意義である割付の予見可能性の低減を直接測定しておらず、その定量化には選択バイアス指標（例：Blackwell-Hodges の予測可能性指標）などを用いた追加検討が必要である点が挙げられる。検討したブロックサイズの組み合わせや層への登録確率の仮定も本発表固有の設定であり、異なる設定下では結果が変わり得る。さらに、本発表のブロックサイズ混合手法は各ブロックサイズの個数をあらかじめ固定する方式であり、ブロックごとに独立にサイズを無作為選択する方式とは異なる点にも留意が必要である。

5. 結論

本発表では、置換ブロック法における固定ブロックサイズ 2、固定ブロックサイズ 4 およびブロックサイズ混合（サイズ 2・4）について、試行ごとに割付表を独立に再生成する 10,000 回のモンテカルロシミュレーションにより、被験者数 100 例・層数 10〜200 の 5 水準の条件下で比較した。

その結果、全体割付差および割付比 1:1 からの乖離率の平均値は、いずれの条件でも固定ブロックサイズ 2 で最小、固定ブロックサイズ 4 で最大となり、ブロックサイズ混合は両者の中間的な性能を示した。1 層あたりの平均割付差においても同様の順位が全条件で一貫して確認された。理論検出力に換算すると、平均ケース同士で比較した条件間の差はほぼ認められず（79.8〜80.0%の範囲）、割付差が最悪の実現値まで変動した場合の低下幅（同一条件内でのばらつき）も 7.2 ポイント程度に留まった。割付比が 2:1 以上に達する確率は全条件で 1%未満と低かった一方、割付比 1:1 から 10%以上乖離する確率は層数が多い条件で 16〜33%に達し、層数の設定が割付比の乖離抑制において重要であることが示された。

以上より、本発表で検討した条件下では均衡性・検出力の観点からブロックサイズ手法間の実質的な差は小さく、手法選択の判断基準は割付の予見可能性（選択バイアスのリスク）に置かれるべきである。ブロックサイズ混合手法は、この予見可能性を低減する手段として ICH E9（2.3.2 節）で紹介されており、本発表の結果は、均衡性を大きく損なうことなく予見可能性の低減を図れるという従来の理解と矛盾しないものであった。ただし、割付差が大きい試行では実薬群の被験者数が予定を下回り得ることから、安全性評価に必要な被験者数が確保されるかという観点も併せて検討することが望ましい。

6. 参考文献

1. ICH Harmonised Tripartite Guideline E9. Statistical Principles for Clinical Trials. 1998. Section 2.3.2 (Randomisation), Section 5.7 (Subgroups, Interactions and Covariates).

2. Blackwell D, Hodges JL Jr. Design for the control of selection bias. Annals of Mathematical Statistics. 1957;28(2):449-460.

3. SAS Institute Inc. SAS/STAT(R) User's Guide: The PLAN Procedure.