



# 臨床試験における 施設効果とリスク管理

関西 SAS ユーザー会

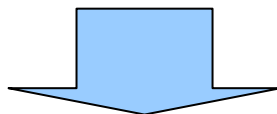
塩野義製薬株式会社 解析センター

小笠原博幸・長谷川貴大・田崎武信

## 臨床試験の共通の目的

- 薬剤の有効性を正確に評価したい
- 試験が成功する確率を高めたい

試験の成功確率に影響を及ぼす要因はなにか

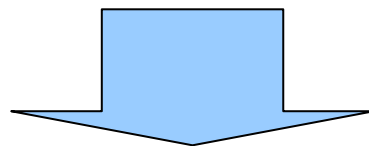


試験の計画段階で考慮すべき要因はなにか

# 検討事項

臨床試験の成功確率に及ぼす影響を以下の2点から検討した

1. **施設当たりの被験者数**が試験の成功確率に及ぼす影響



**施設を考慮した解析**により影響を調整して評価できるか

2. 同一デザインの試験を2つ同時に実施したときの試験全体の成功確率に及ぼす**評価方法**および**試験間変動**の影響

# 施設あたりの被験者数が試験の成功確率に及ぼす影響

## 多施設共同臨床試験

### 利点

- 臨床試験に組み込む被験者をたくさん、早く集積できる
- 試験結果をより広く一般化できる

現実には

施設あたりに組み込む被験者数が少ない

施設あたりの被験者数が不均等になる

Ruvuna(2004): 施設あたりの被験者数が不均等なとき, 試験の検出力は低下する

臨床試験の成功確率にどのような影響を及ぼすのか

# 施設あたりの被験者数が試験の成功確率に及ぼす影響

施設あたりの被験者数  
および施設間差の影響

施設あたりの被験者数が  
インバランスであることの影響

施設あたりの被験者数

インバランス  
なし

インバランス  
あり

施設間差/誤差

インバランス

0.5倍

2倍

1倍

小さい

大きい

施設あたりの被験者数の違い

10施設  
× 20例

100施設  
× 2例

	0.5倍	2倍	1倍
10施設 × 20例			
100施設 × 2例			

5施設 × 10例	8施設 × 10例
5施設 × 30例	2施設 × 60例

施設間の被験者数での  
インバランスの大きさの違い

施設間差の大きさの違い

# 施設あたりの被験者数が試験の成功確率に及ぼす影響

被験薬(100例)と対照薬(100例)を比較する並行群間比較試験

## シミュレーションデータ

施設  $j$  で薬剤  $i$  による治療を受けた  
 $k$  番目の被験者の応答

$$Y_{ijk} = \text{全平均} + \text{薬剤 } i \text{ の効果} + \text{施設 } j \text{ の効果} + \text{誤差}$$

薬剤群間差  
= 被験薬の効果 - 対照薬の効果  
0.0~1.0

施設  $j$  の効果  $\sim N(0, 0.5)$   
 $N(0, 1)$   
 $N(0, 2)$

誤差  $\sim N(0, 1)$

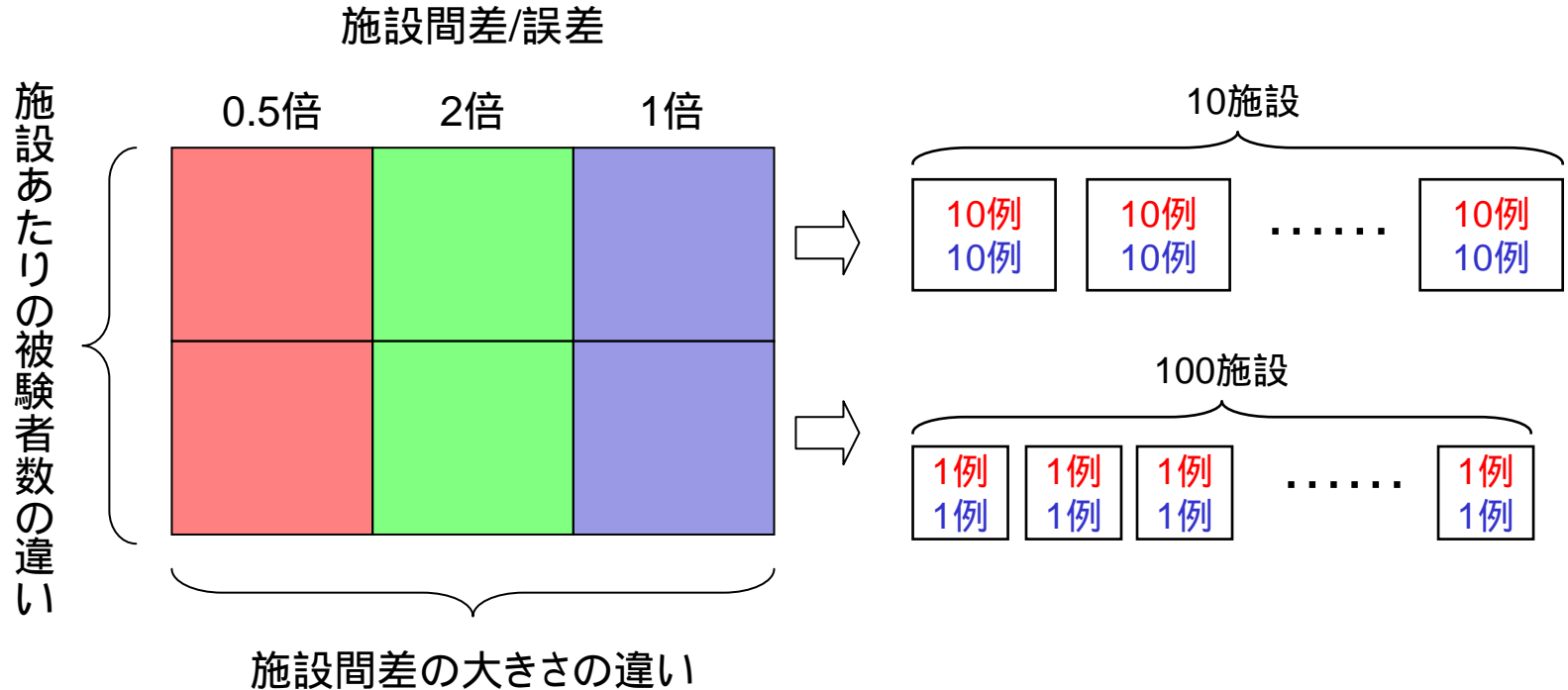
各条件で薬剤群間差を0.0~1.0と変化させ、それぞれ10000回の反復を行なった

**成功確率**: t 検定によって薬剤群間差を片側有意水準0.025で検出した割合(検出力)  
(施設効果を考慮しない解析)

# 施設あたりの被験者数が試験の成功確率に及ぼす影響

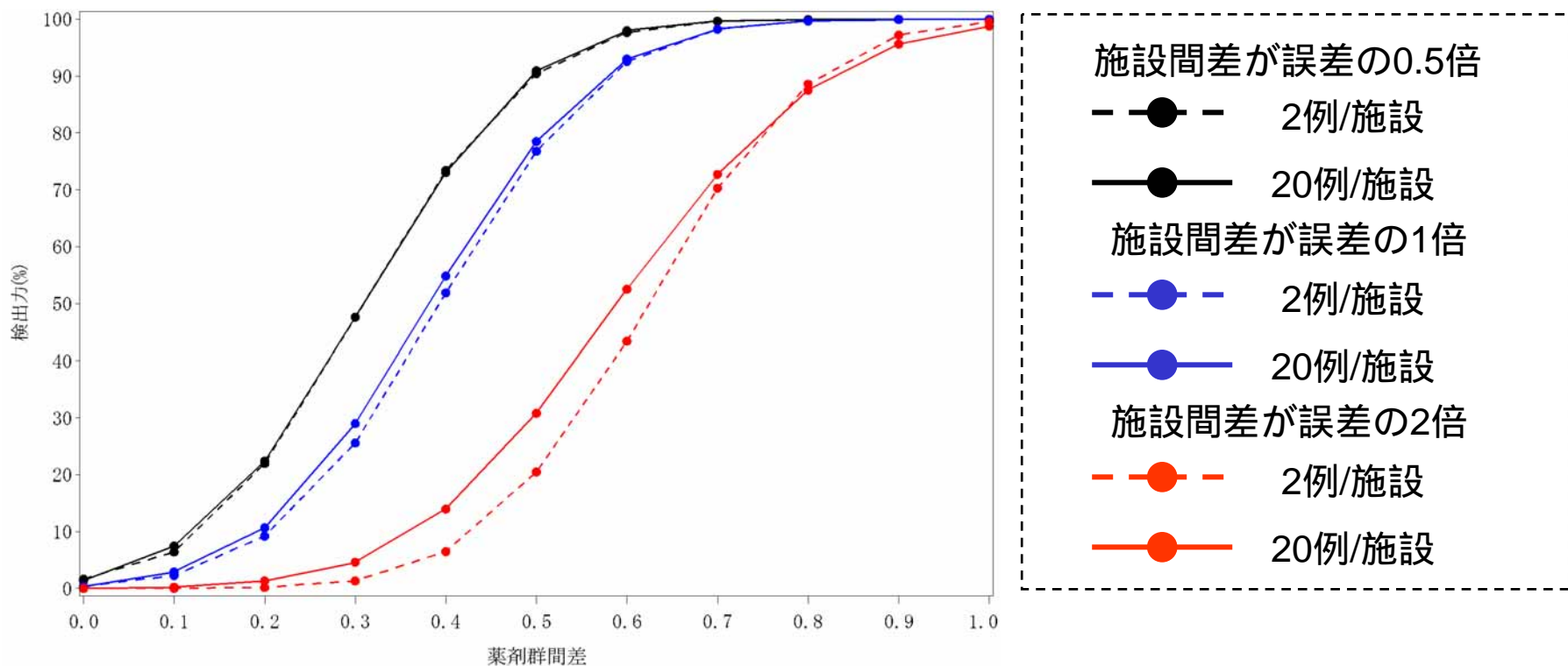
## 施設あたりの被験者数および施設間差の影響

### シミュレーション状況



# 施設あたりの被験者数が試験の成功確率に及ぼす影響

結果：施設あたりの被験者数および施設間差の影響



施設間差を小さくすることによって試験の成功確率が高まった。

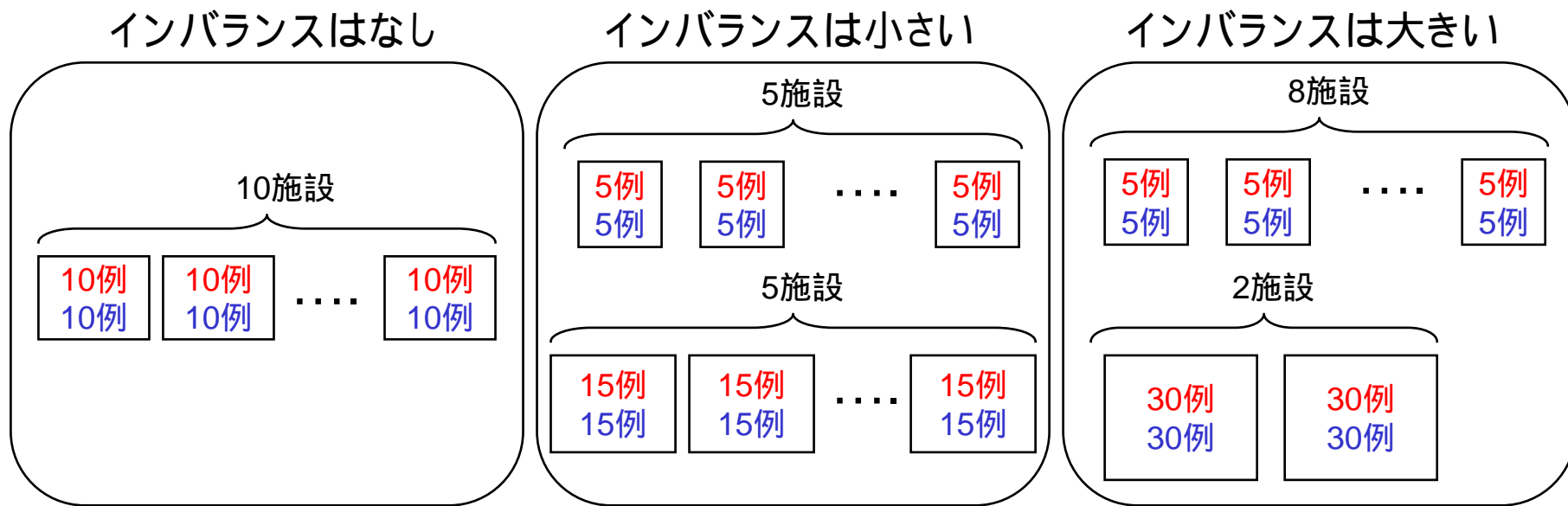
施設間差が大きいときほど施設あたりの被験者数の影響が大きかった。

# 施設あたりの被験者数が試験の成功確率に及ぼす影響

## 施設あたりの被験者数がインバランスであることの影響

### シミュレーション状況

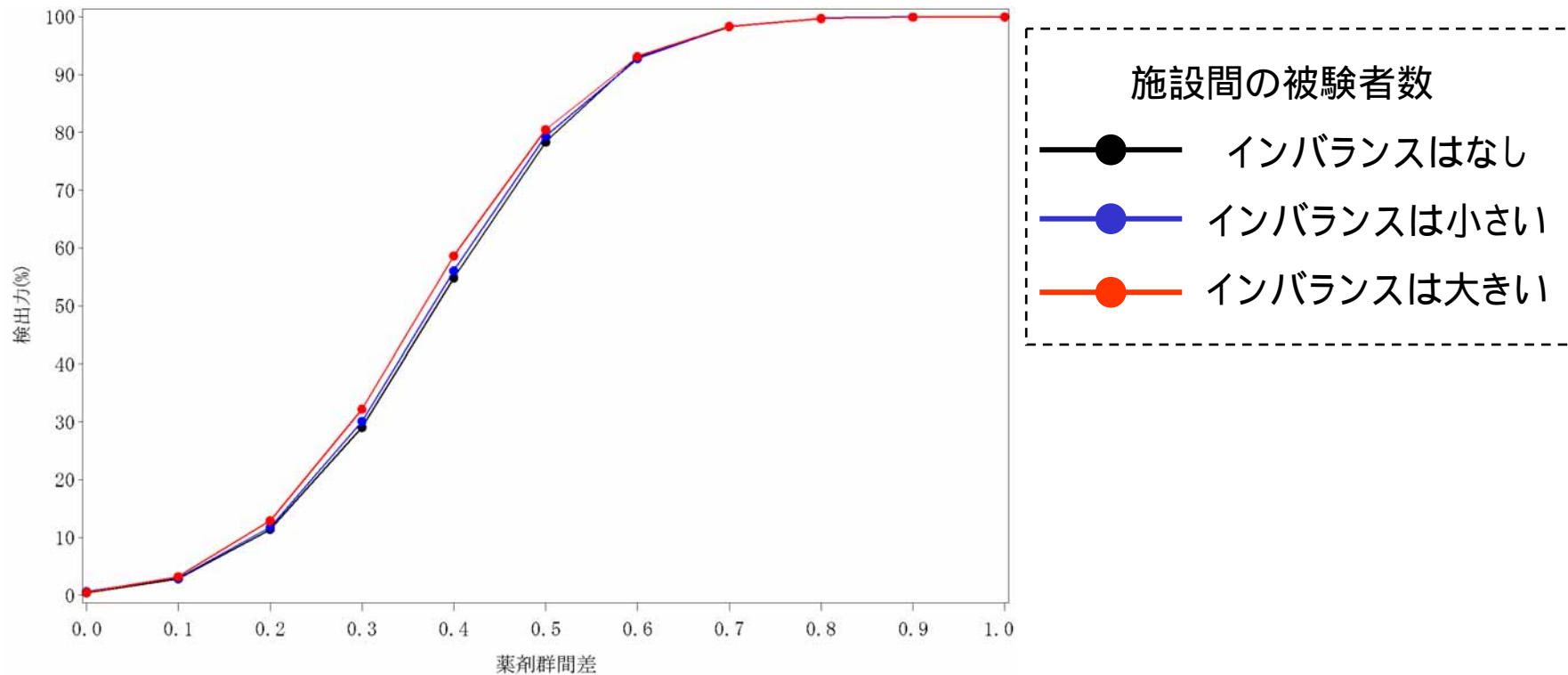
10施設を組み込んだ試験を想定



Ruvuna(2004)の報告では施設あたりの被験者数がインバランスであることによって試験の成功確率が低下することが示唆されていた

# 施設あたりの被験者数が試験の成功確率に及ぼす影響

結果：施設あたりの被験者数がインバランスであることの影響

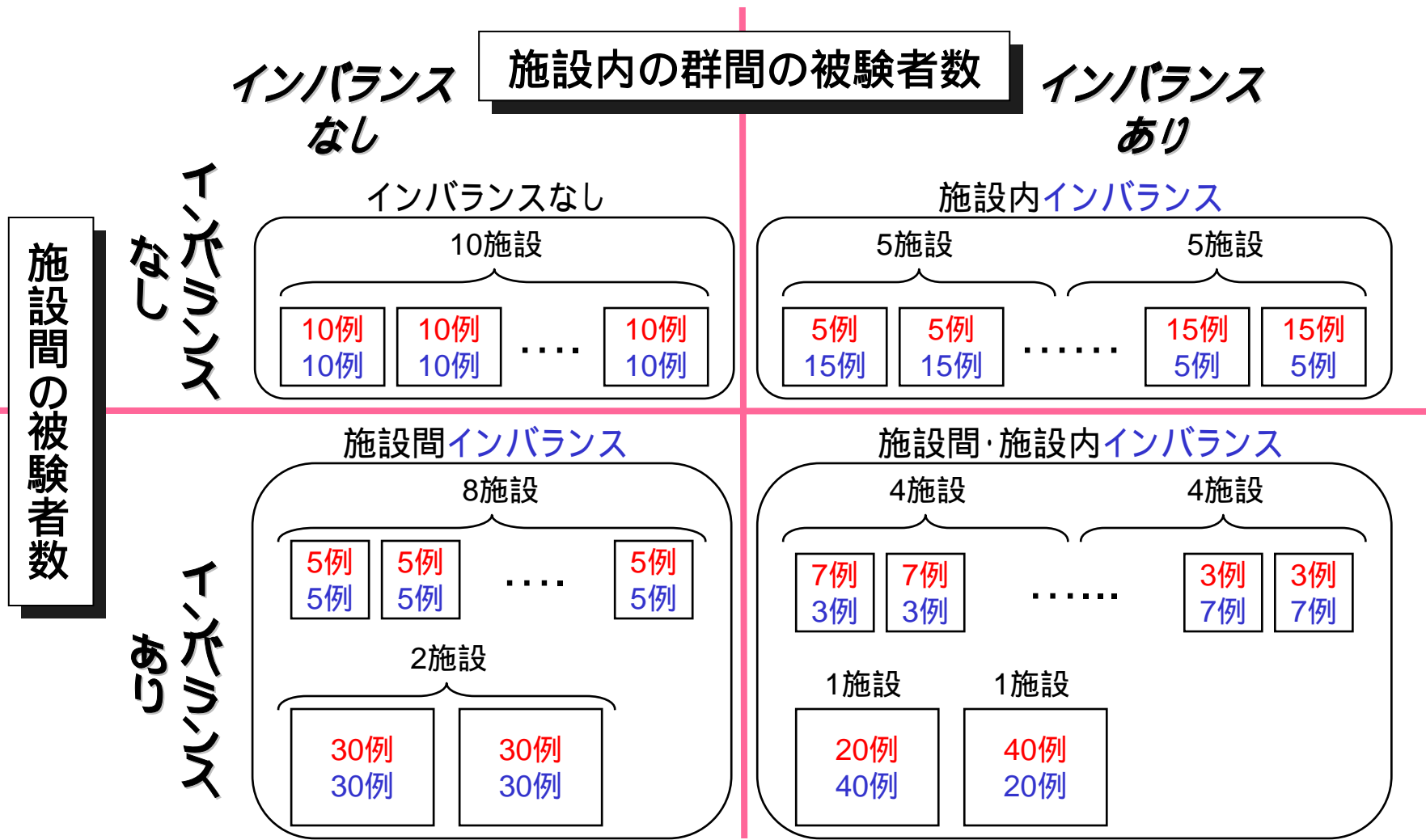


施設あたりの被験者数がインバランスである場合とインバランスがない場合とで試験の成功確率に大きな違いはなかった(Ruvuna(2004)の研究と結果が異なる)。

ただしここでは、全ての施設において被験者数が群間で釣り合っていると仮定した。

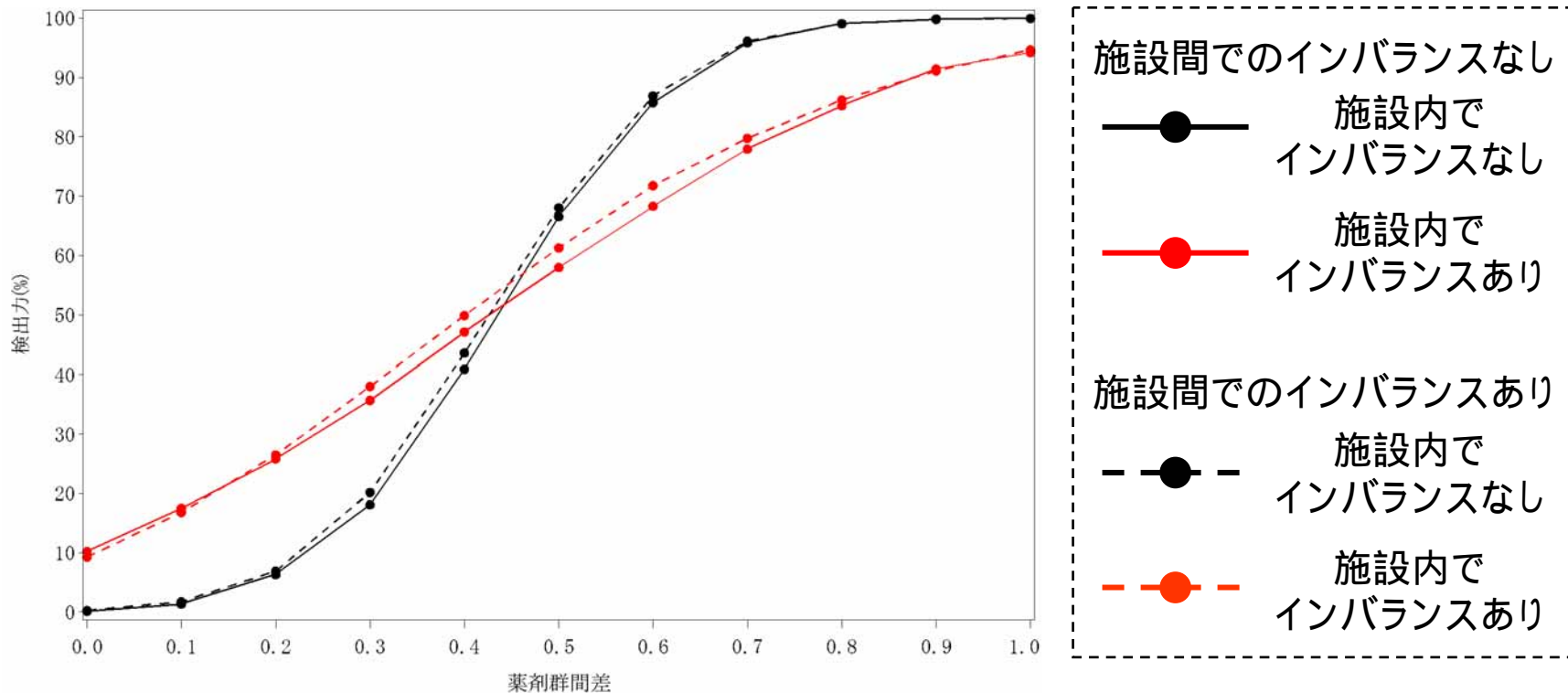
# Ruvuna(2004)の結果と本シミュレーション結果の相違

## 施設間・施設内で被験者数がインバランスであることの影響



# Ruvuna(2004)の結果と本シミュレーション結果の相違

## 結果: 施設間・施設内で被験者数がインバランスであることの影響



薬剤群間差が小さいとき, 施設内の被験者数がインバランスであることで, 施設間差を薬剤効果の差として検出した可能性が考えられる.

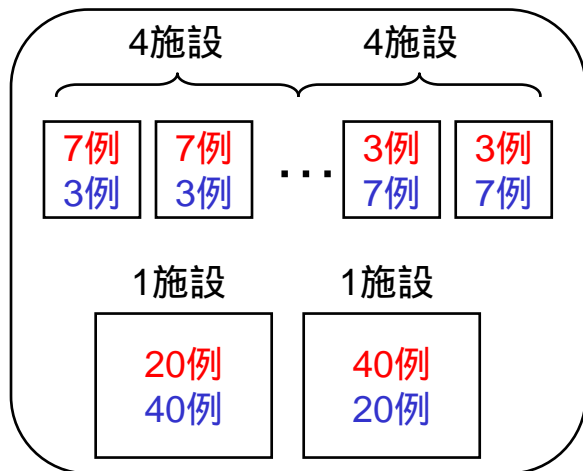
# 施設間および施設内の群間で被験者数がインバランス

## 施設内の被験者数が一方の群に偏ることの影響

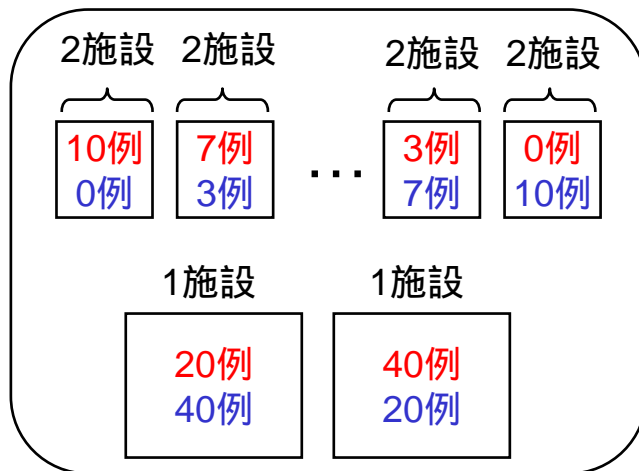
より現実的なケース

少数の大規模施設  
多数の小規模施設 → 一方の群に被験者が偏る可能性

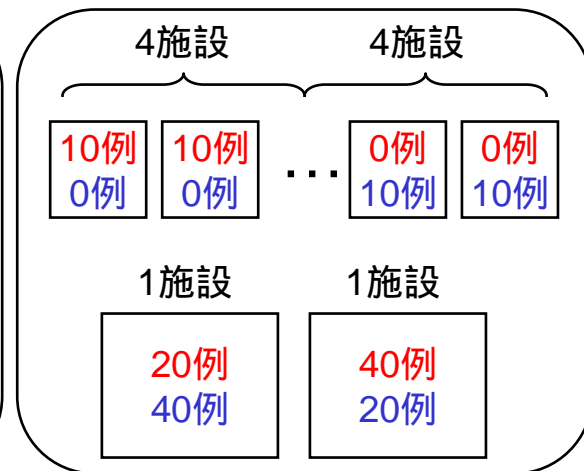
偏りはなし



偏りは小さい



偏りは大きい

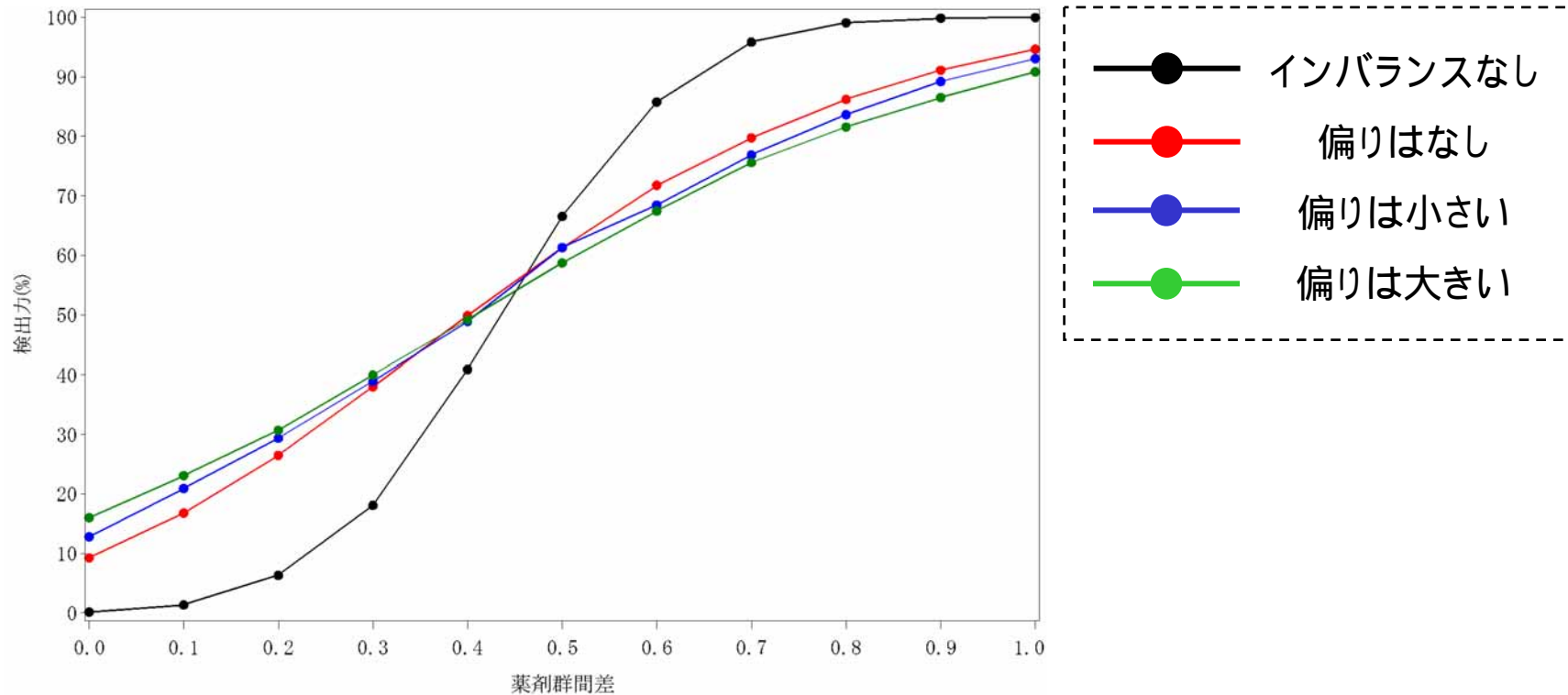


半数の小規模施設  
で偏りあり

全ての小規模施設  
で偏りあり

# 施設間および施設内の群間で被験者数がインバ ス

## 結果：施設内の被験者数が一方の群に偏ることの影響



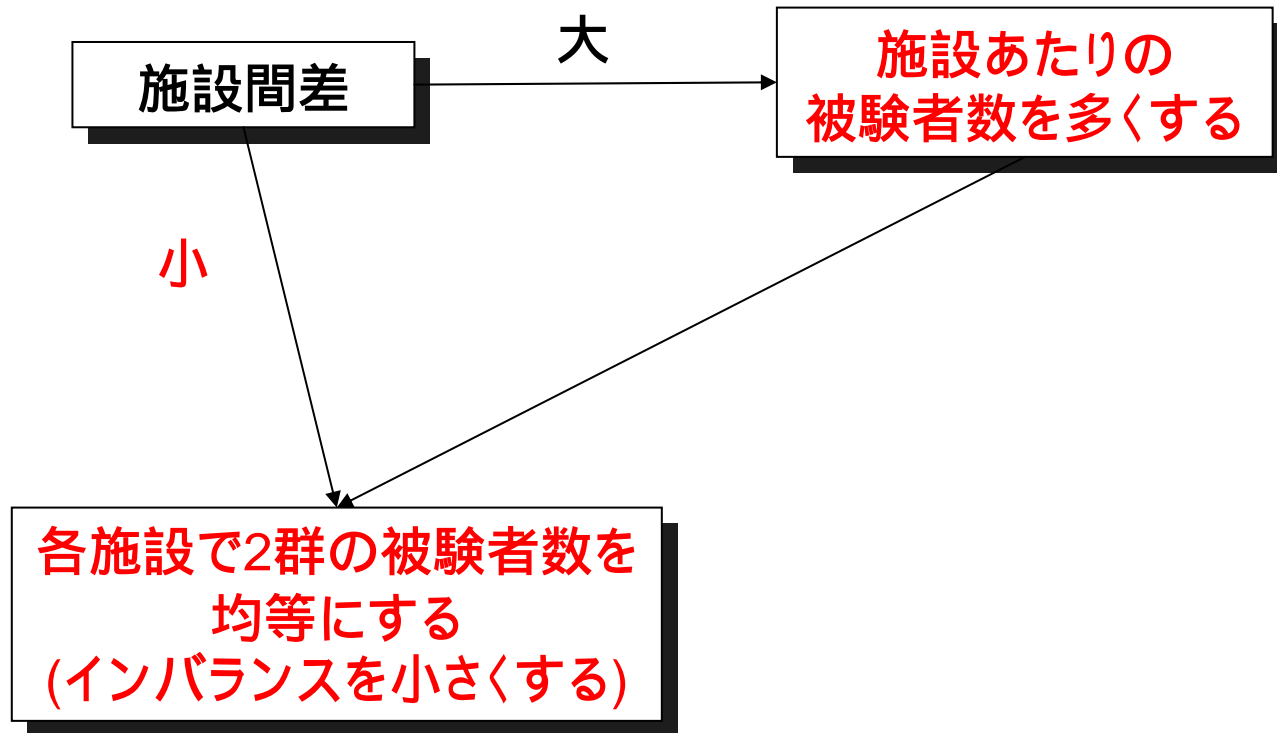
薬剤群間差が小さいとき、施設内の被験者数に偏りが生じることで、施設間差を薬剤効果の差として検出した可能性が考えられる。

偏りが大きくなるほど、その現象は顕著であった。

# 施設あたりの被験者数が試験の成功確率に及ぼす影響

## まとめ

試験の成功確率を高めるために



# 施設効果を考慮したモデルによる解析

ICH E9ガイドライン: 施設効果を**固定**効果として統計モデルに含めることを前提

Hirotsu and Hothorn(2003): 日本で実施される超多施設臨床試験では施設効果を**変量**効果として含めることが望ましい

解析モデル(混合効果モデル)

$$Y_{ijk} = \text{全平均} + \text{薬剤 } i \text{ の効果} + \text{施設 } j \text{ の効果(変量効果)} + \text{誤差}$$

シミュレーションデータはこれまでと同様のモデルにより生成

**施設あたりの被験者数の成功確率に対する影響**

施設あたりの被験者数および施設間差

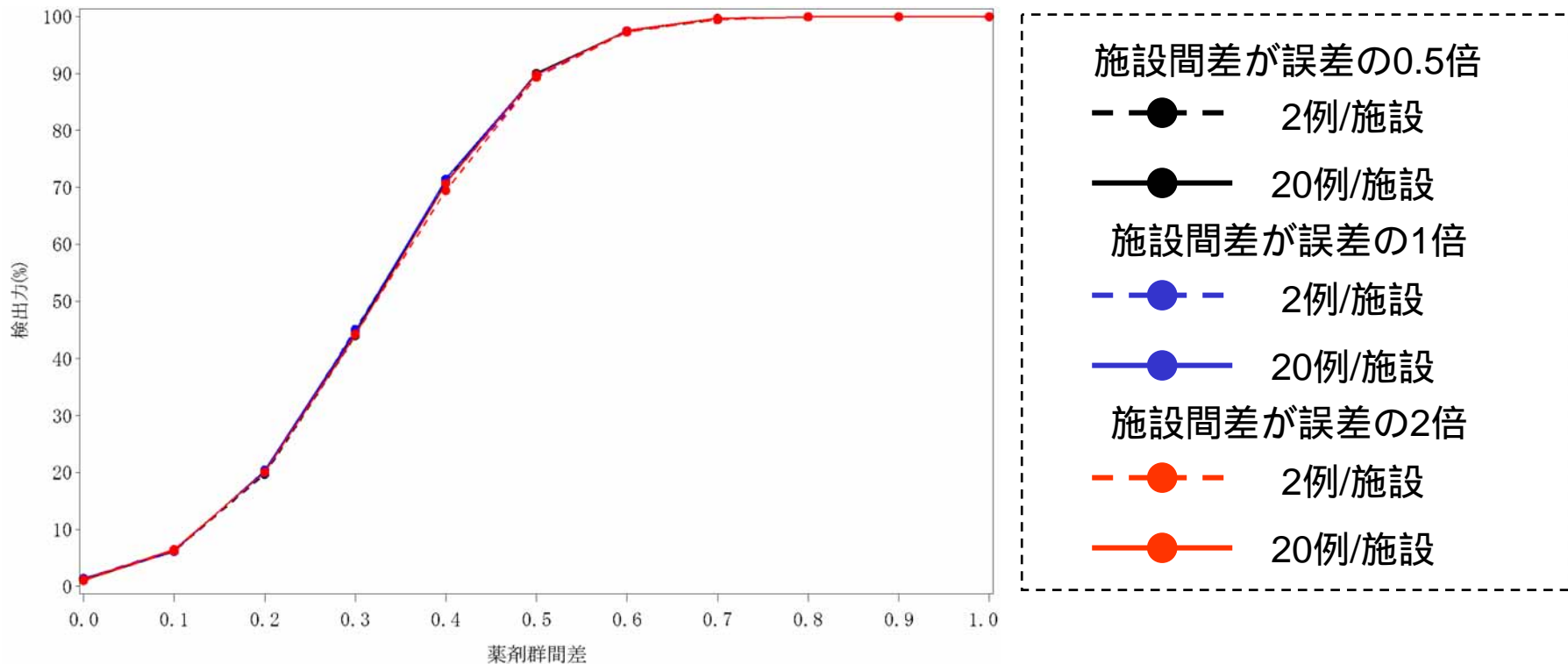
施設間・施設内で被験者数がインバランス

施設内の被験者が一方の群に偏る

各条件で薬剤群間差を0.0~1.0と変化させ、それぞれ10000回の反復を行なった

# 施設効果を考慮したモデルによる解析

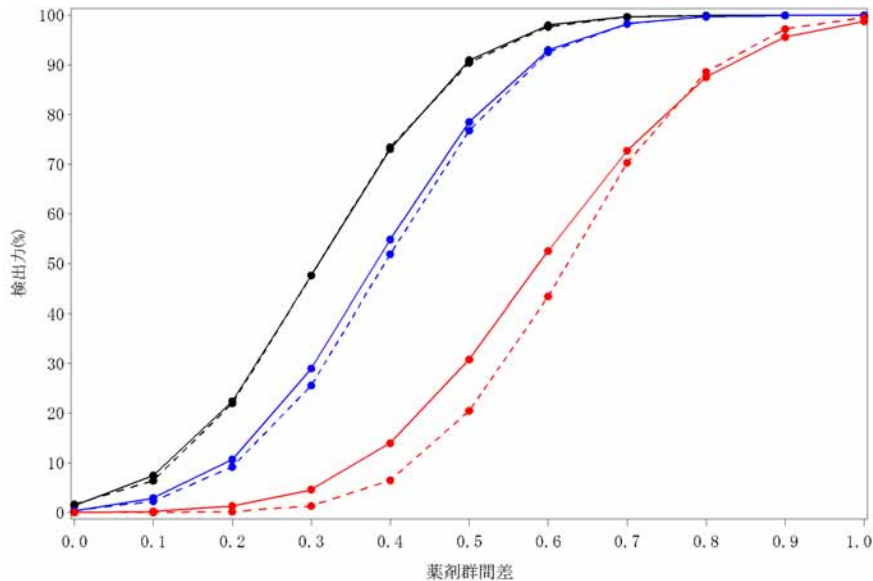
結果： 施設あたりの被験者数および施設間差の影響



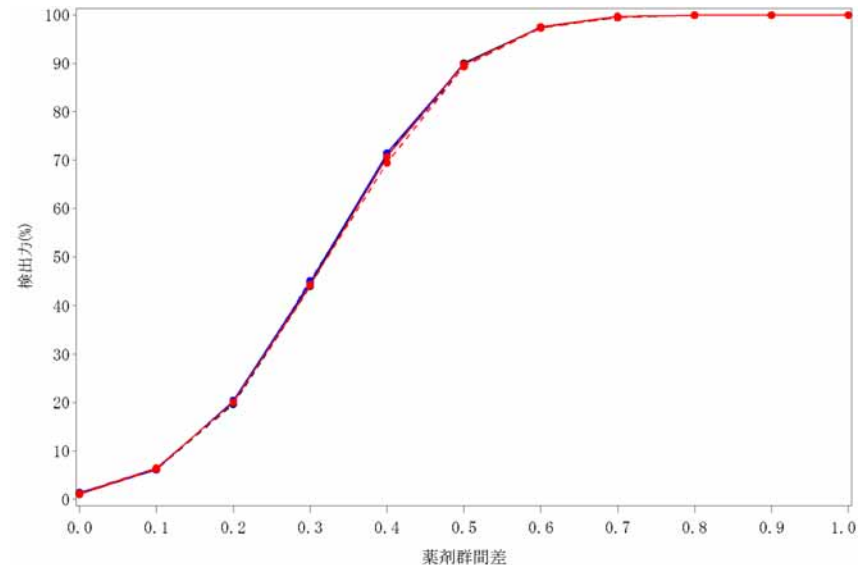
施設効果を考慮した解析では、施設間差および施設あたりの被験者数による試験の成功確率に違いがなかった。

# 施設効果を考慮したモデルによる解析

結果： 施設あたりの被験者数および施設間差の影響



施設を考慮しない解析



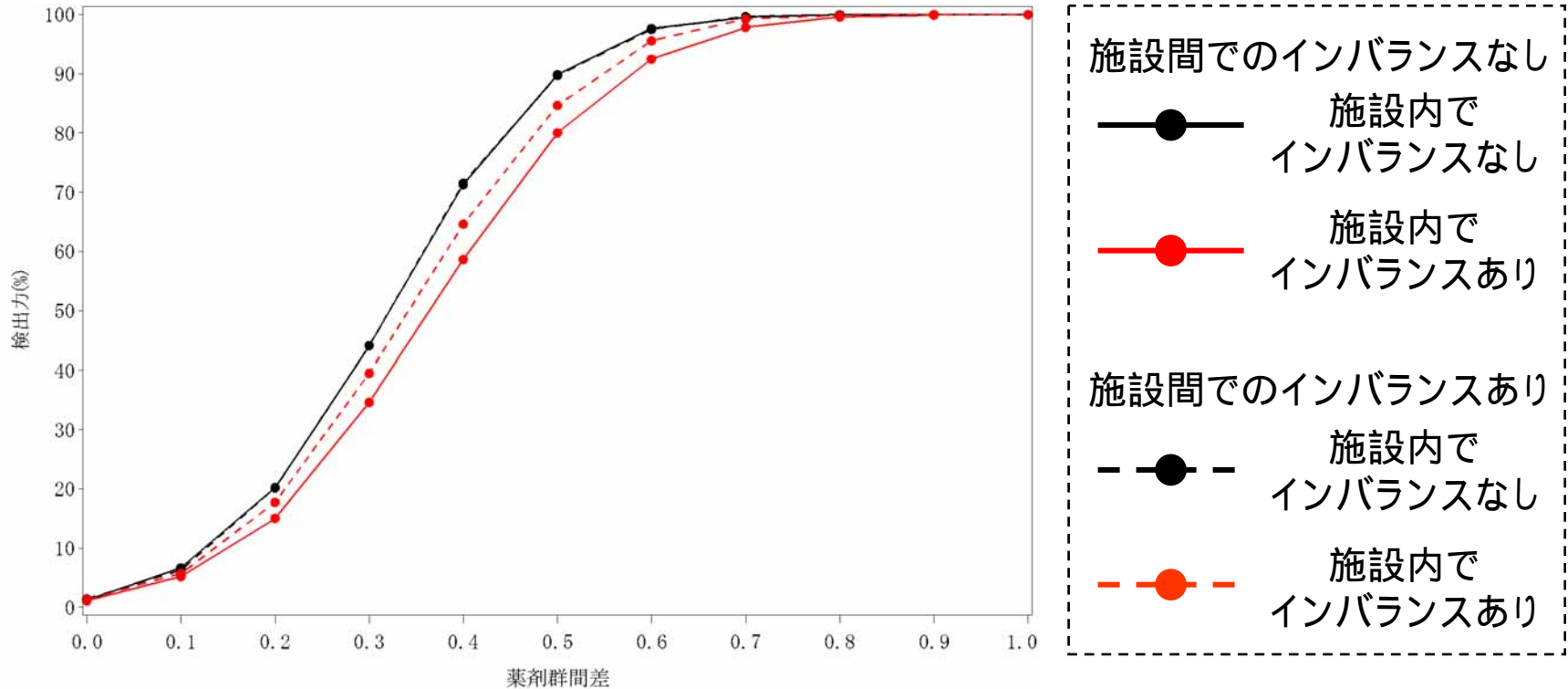
施設を考慮した解析

施設の効果を変量効果としてモデルに含めることで、施設間差の大きさおよび施設あたりの被験者数の影響を調整して評価することが可能。

➡ ただし、施設がランダムに選定されていることが前提となる。

# 施設効果を考慮したモデルによる解析

結果： 施設間・施設内で被験者数がインバランスであることの影響

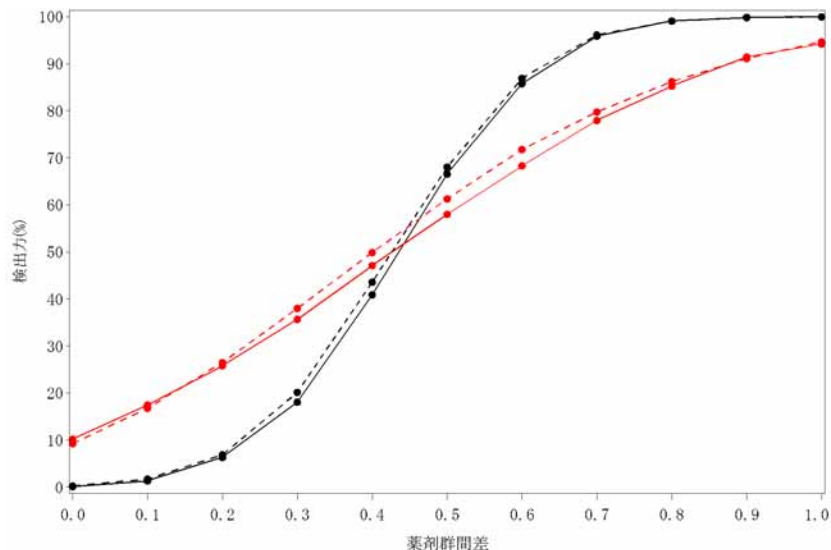


施設内で被験者数がインバランスでなければ、施設間での被験者数のインバランスは試験の成功確率に影響を及ぼさなかった。

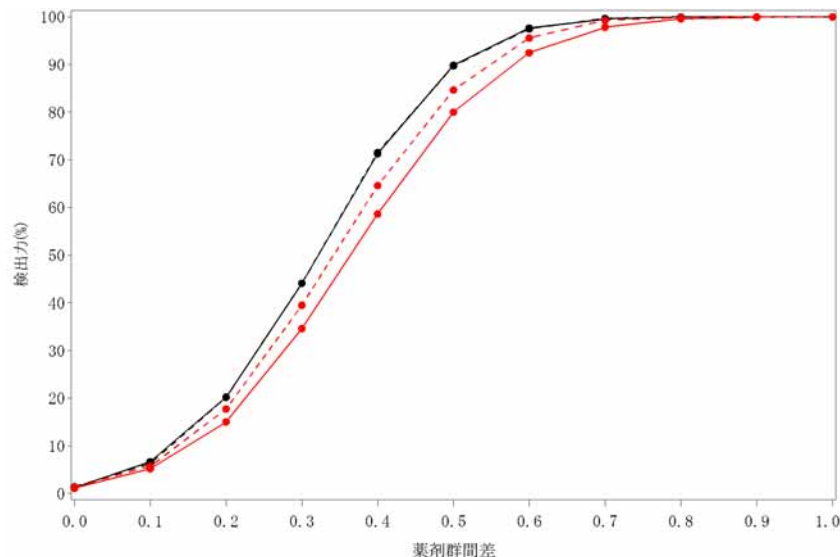
施設内で被験者数がインバランスであると試験の成功確率は低下した。

# 施設効果を考慮したモデルによる解析

結果： 施設間・施設内で被験者数がインバランスであることの影響



施設を考慮しない解析



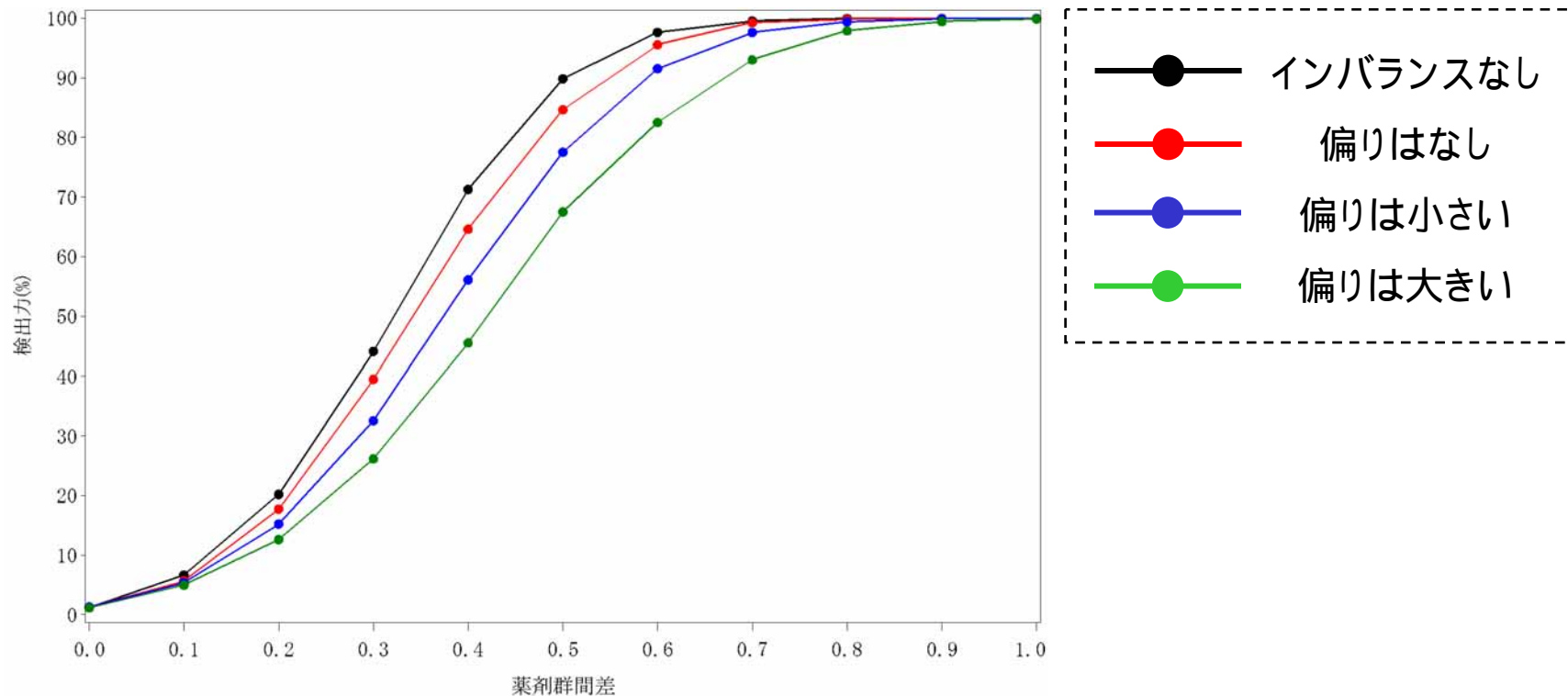
施設を考慮した解析

施設を考慮した解析により

施設内で被験者数がインバランスである場合でも、薬剤の効果をより正確に評価できた。  
施設効果を考慮する場合でも、施設内で被験者を均等に割付けることが重要。

# 施設効果を考慮したモデルによる解析

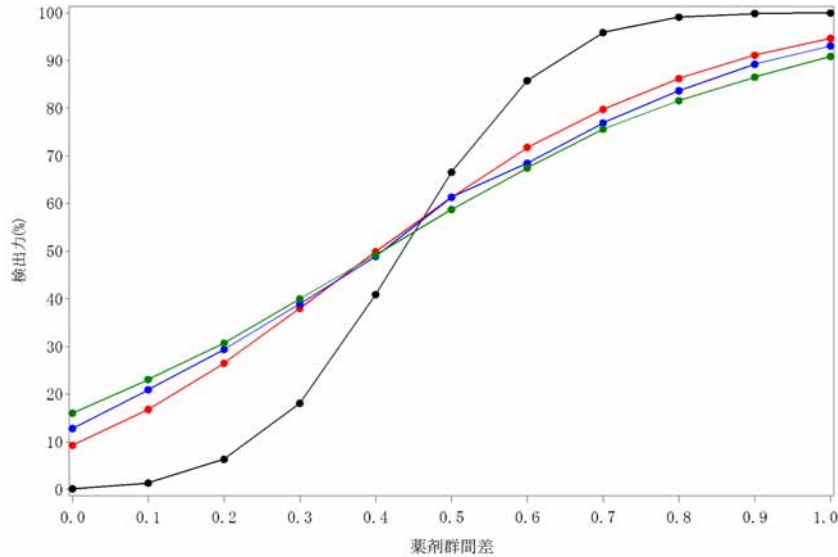
結果： 施設内の被験者数が一方の群に偏ることの影響



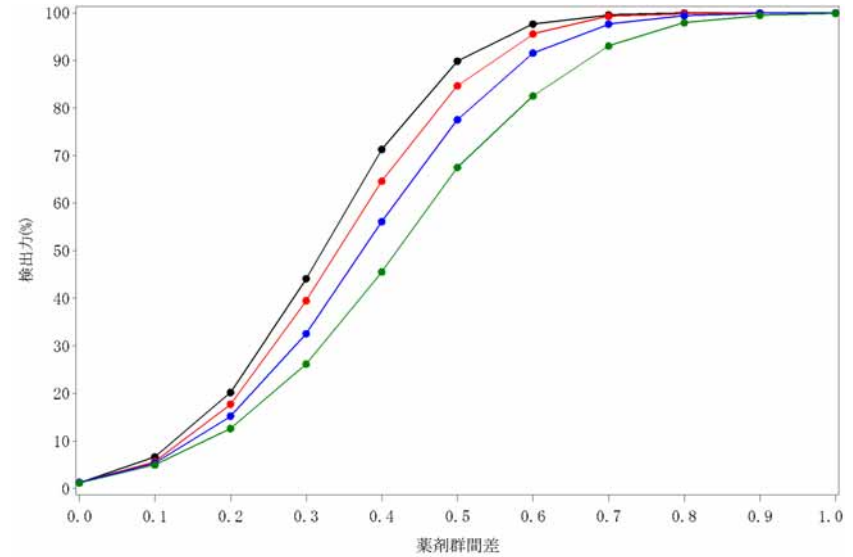
施設内の被験者数が一方の群に偏ることは試験の成功確率を低下させた。そのインバランスが大きいほどその影響は大きかった。

# 施設効果を考慮したモデルによる解析

結果： 施設内の被験者数が一方の群に偏ることの影響



施設を考慮しない解析



施設を考慮した解析

施設を考慮した解析により

薬剤の効果をより正確に評価できた。

施設内の被験者数は群間でインバランスを小さくするように割り付けることが重要。

## まとめ

施設を考慮したモデルによる解析を行う利点

試験の成功確率は,

- 施設あたりの被験者数
- 施設間差の大きさ

の影響を調整して評価される

試験に組み入れる施設がランダムに選択されていることが条件

各施設の各群で被験者を等しく割り付けることが重要  
(インバランスを小さく)

# 同一デザインの試験を2つ同時に実施したときの試験全体の成功確率に及ぼす評価方法および試験間変動の影響

うつ病のような疾患領域  成功確率が低い

海外で見られるケース

同一デザインの試験を複数実施する  
いずれかの試験で成功することを試験全体の成功と見なす

Shun et al.(2005): 1試験アプローチと2試験アプローチ

2試験のいずれも成功することを想定

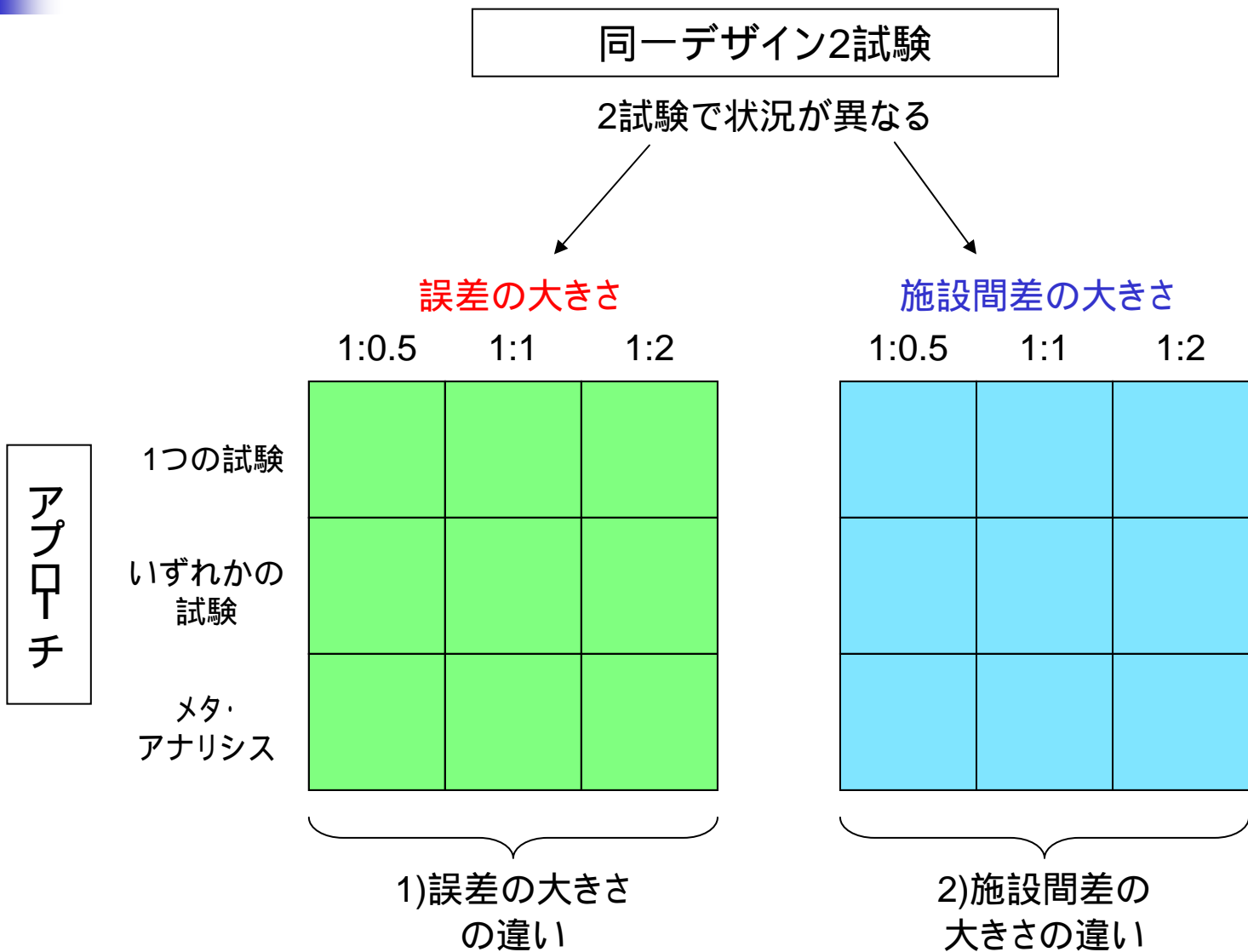
別のアプローチとして

メタ・アナリシス: 複数の試験成績を併合する

成功確率および試験成績の信頼性が向上する可能性

同一デザインの2試験で状況が異なるとき, アプローチの違いが試験全体の成功確率にどのような影響を及ぼすのか

# 同一デザインの試験を2つ同時に実施したときの試験全体の成功確率に及ぼす評価方法および試験間変動の影響



# 同一デザインの試験を2つ同時に実施したときの試験全体の成功確率に及ぼす評価方法および試験間変動の影響

被験薬(50例)と対照薬(50例)を比較する並行群間比較試験

同一デザインによる2試験を実施する

シミュレーションデータ

施設  $j$  で薬剤  $i$  による治療を受けた  
 $k$  番目の被験者の応答

$$Y_{ijk} = \text{全平均} + \text{薬剤 } i \text{ の効果} + \text{施設 } j \text{ の効果} + \text{誤差}$$

薬剤群間差  
= 被験薬の効果 - 対照薬の効果  
0.0~1.0

施設  $j$  の効果  $\sim N(0, 0.5)$   
 $N(0, 1)$   
 $N(0, 2)$

誤差  $\sim N(0, 0.5)$   
 $N(0, 1)$   
 $N(0, 2)$

各条件で薬剤群間差を0.0~1.0と変化させ、それぞれ10000回の反復を行なった

**成功確率**: t 検定によって薬剤群間差を片側有意水準0.025で検出した割合(検出力)  
(施設効果を考慮しない解析)

# 同一デザインの試験を2つ同時に実施したときの試験全体の成功確率に及ぼす評価方法および試験間変動の影響

## シミュレーション

### 同一デザインの2試験の状況

#### 1) 2試験で**誤差の大きさ**が異なるとき

1つ目の試験の誤差の大きさ:2つ目の試験の誤差の大きさ

1:0.5

1:1

1:2

#### 2) 2試験で**施設間差の大きさ**が異なるとき

1つ目の試験の施設間差の大きさ:2つ目の試験の施設間差の大きさ

1:0.5

1:1

1:2



# 同一デザインの試験を2つ同時に実施したときの試験全体の成功確率に及ぼす評価方法および試験間変動の影響

## シミュレーション

### 試験成績を評価するための3つのアプローチ

- **1つの試験**  
2試験をプールし, あたかも1つの試験のように見なす
- **いずれかの試験**  
2試験のいずれかで成功すれば試験全体の成功と見なす
- **メタ・アナリシス**  
2試験の成績をメタ・アナリシスを用いて併合する

# 同一デザインの試験を2つ同時に実施したときの試験全体の成功確率に及ぼす評価方法および試験間変動の影響

## シミュレーション

### メタ・アナリシスの方法

2試験の併合成績をADとすると

$$AD = \frac{\sum_{l=1}^2 (\bar{y}_{1l} - \bar{y}_{2l}) \cdot w_l}{\sum_{l=1}^2 w_l}$$

$l$  : 試験

$w_l$  : 試験  $l$  の重み

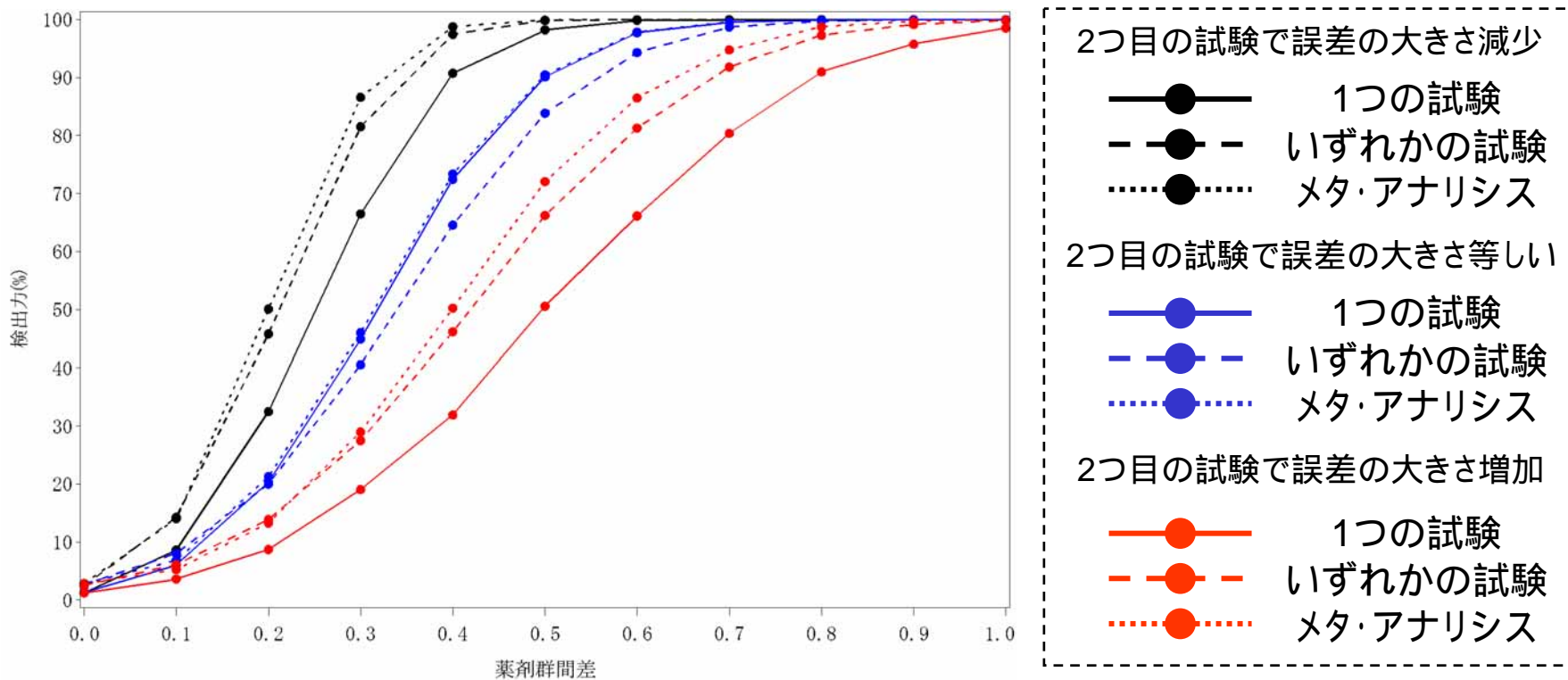
$$w_l = \left( \frac{1}{SE(\bar{y}_{1l} - \bar{y}_{2l})} \right)^2$$

### 併合成績の有意性検定

$$Q = AD^2 \cdot \sum_{l=1}^2 w_l \sim \chi_1^2(\alpha)$$

# 同一デザインの試験を2つ同時に実施したときの試験全体の成功確率に及ぼす評価方法および試験間変動の影響

結果: 1) 2試験で誤差の大きさが異なるとき

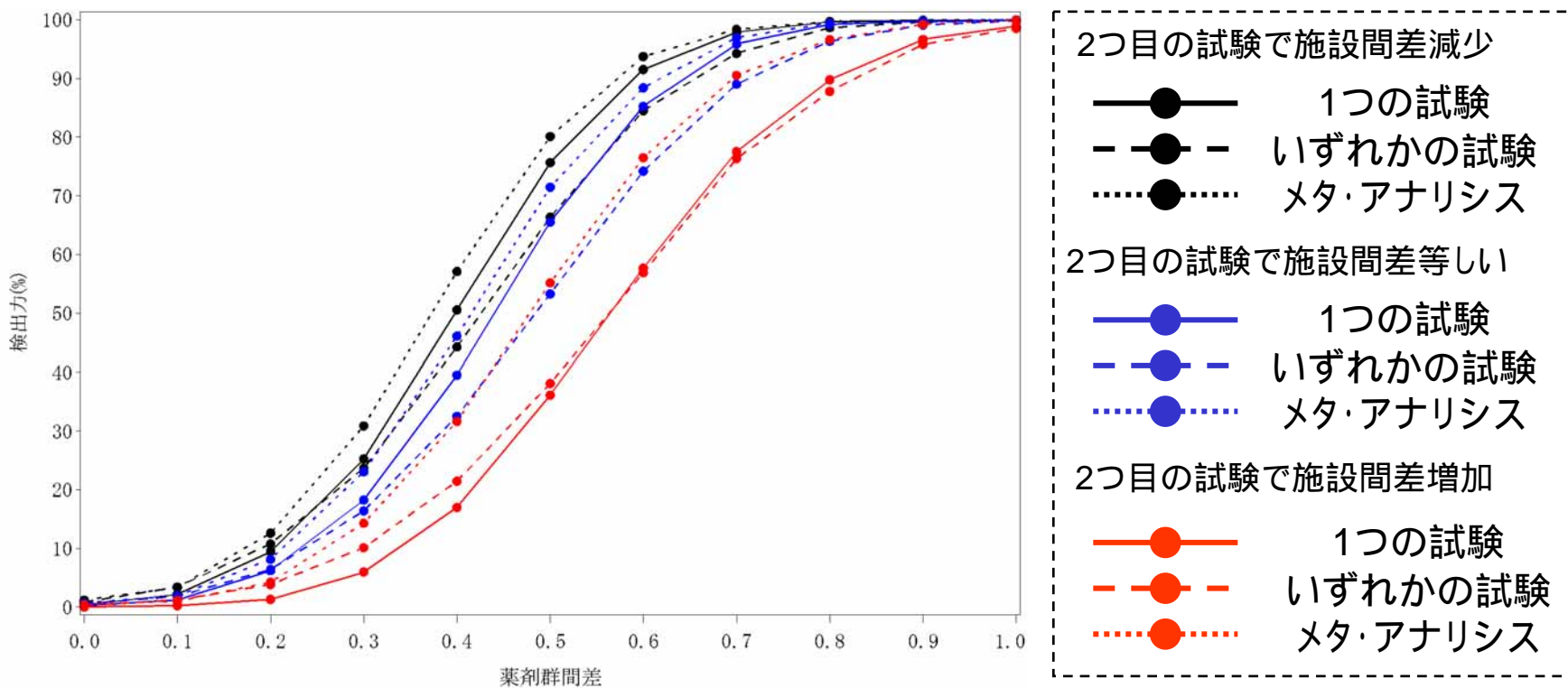


2試験で誤差の大きさが異なると、「1つの試験」アプローチによる成功確率は低下した。  
Shun et al.(2005)で検討された「いずれの試験でも成功」とする状況が自然かつ妥当。

「いずれかの試験」 選択バイアス

# 同一デザインの試験を2つ同時に実施したときの試験全体の成功確率に及ぼす評価方法および試験間変動の影響

## 結果:2) 2試験で施設間差の大きさが異なるとき



どの場合にも「メタ・アナリシス」アプローチによる成功確率が最も高かった。

「1つの試験」と「いずれかの試験」アプローチで比較すると、施設間差が減少したときに「1つの試験」、増加したときに「いずれかの試験」アプローチで高い傾向にあった。

# 同一デザインの試験を2つ同時に実施したときの試験全体の成功確率に及ぼす評価方法および試験間変動の影響

## まとめ

### 試験の成功確率を高めるために

**メタ・アナリシス**により2試験の成績を併合して評価すること  
成功確率

誤差の大きさが異なる

「メタ」>「いずれか」>「1つ」

施設間差の大きさが異なる

「メタ」>「いずれか」>「1つ」(施設間差が大のとき)

「メタ」>「1つ」>「いずれか」(施設間差が小のとき)

**誤差が小さいこと・施設間差を小さくすること**

「いずれかの試験」よりも「いずれの試験でも成功」とする状況が妥当

- 選択報告バイアス



- [1] Francis Ruvuna(2004). Unequal Center Sizes, Sample Size and Power in Multicenter Clinical Trials. *Drug Information Journal*, 57, 387-394
- [2] C.Hirotsu and L. A. Hothorn(2003). Impact of the ICH E9 Guideline *Statistical Principles for Clinical Trials* on the Conduct of Clinical Trials in Japan. *Drug Information Journal*, 37, 381-395
- [3] Z. Shun, E. Chi, S. Durrleman and L. Fisher(2005). Statistical Consideration of the Strategy for Demonstrating Clinical Evidence of Effectiveness - One Larger vs Two Smaller Pivotal Studies. *Statistics in Medicine*, 24, 1619-1637