

## 第Ⅱ相試験と第Ⅲ相試験を結合した試験デザインの提案 ～エンドポイントが二項分布に従う場合への拡張～

北里大学大学院 薬学研究科 臨床統計部門 中條 愛 (Ai Nakajo)  
Division of Biostatistics, Graduate School of Pharmaceutical Sciences,  
Kitasato University

北里大学 薬学部 臨床統計部門 高橋 史朗 (Fumiaki Takahashi)

北里大学 薬学部 臨床統計部門 竹内 正弘 (Masahiro Takeuchi)

Division of Biostatistics, School of Pharmaceutical Sciences, Kitasato University

### 1. はじめに

2004年, 米国食品医薬局<sup>1)</sup>は, ここ10年間(1993~2003年)に医薬品開発費用が高騰していることを示した(図1-1, 1-2). この原因としては, ヒトゲノムの解読や再生医療等の生命科学技術の急速な発展に伴い, ゲノム創薬やテーラーメイド医療などの新たな医薬品開発に多額の費用を投じる必要性が生じたことが挙げられる. また, このようにして開発された医薬品は有効性や安全性などに十分な科学的根拠がないため, 臨床試験における患者一人当たりが高額な費用がかかることもこの原因として挙げられる. これらの結果として, 新薬承認申請の件数は減少している(図1-3). また, 新たな医薬品開発の時代に突入したにもかかわらず, 新薬を評価するための最適な統計手法がないことも問題とされている. 日本においては, 開発費用の高騰をはじめ, 試験効率の悪さなどが海外の状況以上に深刻となっている<sup>2)</sup>(図1-4, 1-5). そこで, 現在必要とされているのが, 柔軟かつ効率的な試験デザインと最適な統計手法である.

これらの要求を満たす試験デザインとして提案されているのがAdaptive Designである. Adaptive Designとは, 開発中の薬が実際に効果のなかった場合に早期に試験を中止したり, 症例数の再計算を行うことにより試験に必要な検出力を十分に担保するなど, 試験が行われている最中に試験デザインを変更することが可能な柔軟な試験デザインのことである. 多くのAdaptive Designが提案されているが, Liu & Pledger<sup>3)</sup>は, 第Ⅱ相と第Ⅲ相を結合することによって症例数を削減し, さらに試験期間を短縮する方法を提案した. この方法は, いずれの評価項目とも正規分布に従う連続データの場合に対してのみ提案されている. 骨粗鬆症を対象とした試験では, 第Ⅱ相で骨密度といった連続データを, 第Ⅲ相では骨折の有無といった二値データを評価項目として用いることが多く, この方法を直接適用することができない.

そこで本研究では, より柔軟で効率的な試験デザインおよび最適な統計手法を提案するため, 第Ⅲ相の評価項目が二値データの場合へ拡張することとした. また, 提案するデザ

インの有用性を示すため、症例数削減と時間短縮の観点から提案する方法と従来の方法を比較する。

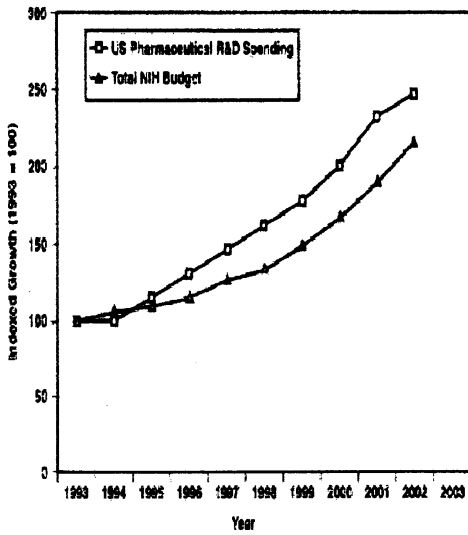
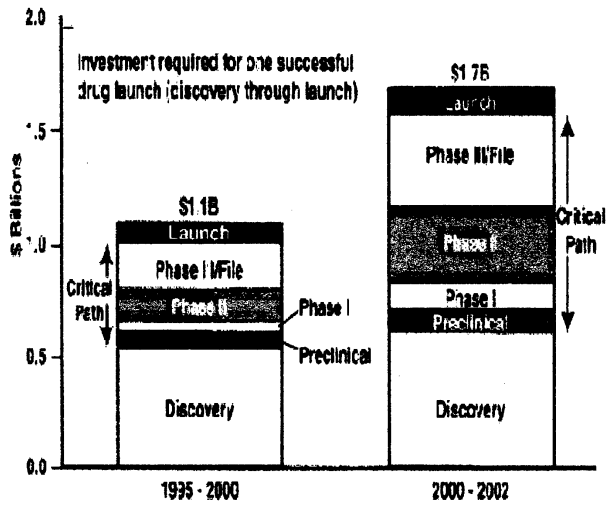


図 1-1 10年間の医薬品開発費用



SOURCE: Windhamer's in Vivo: The Business & Medicine Report, Bain drug economics model, 2003

図 1-2 開発の各段階における費用の拡大

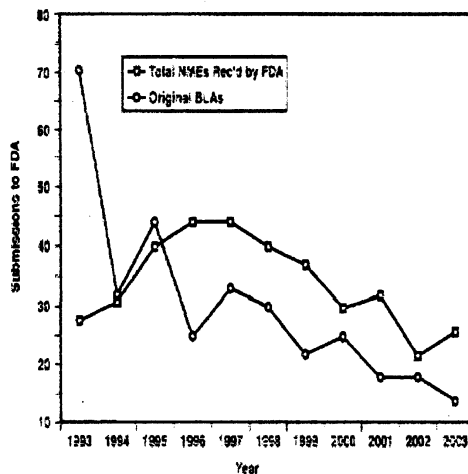


図 1-3 10年間のFDAへの申請件数

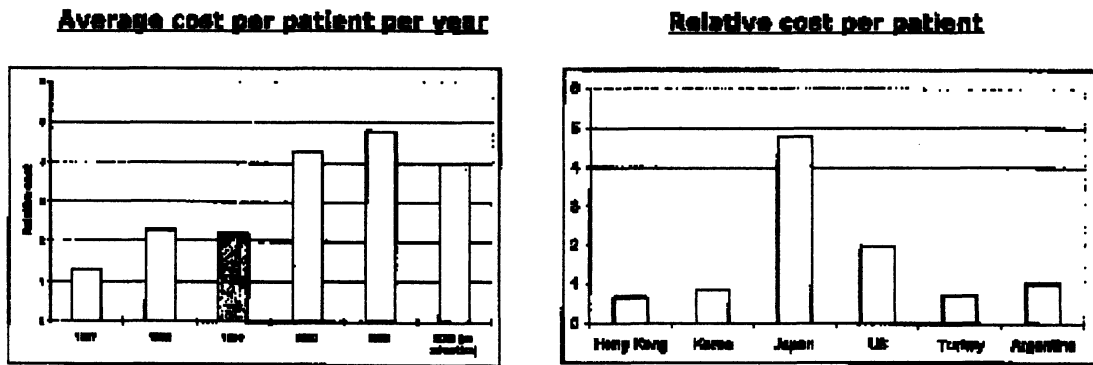


図 1-4 日本における開発費用の推移（左）および他国との比較（右）

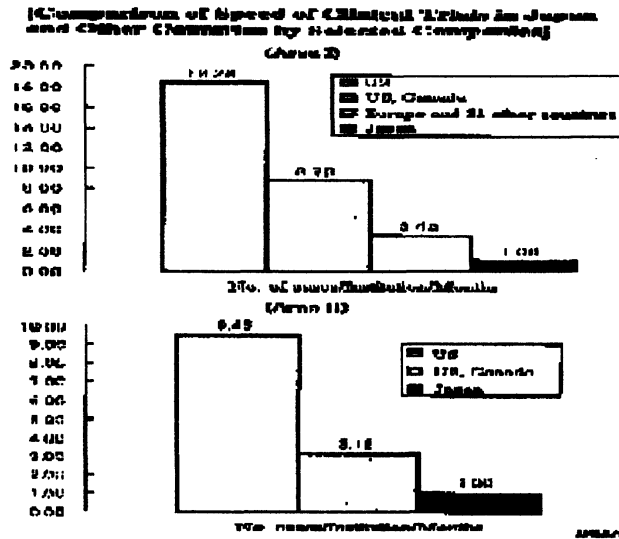


図 1-5 日本と他国との試験期間の比較

## 2. 第Ⅱ相/第Ⅲ相結合デザイン

Liu & Pledger<sup>3)</sup> によって提案された『第Ⅱ相/第Ⅲ相結合デザイン』の概要は以下のとおりである。

- ① 第Ⅲ相比較試験を2段階とし、第1段階に第Ⅱ相用量反応試験を組み込んだ2段階試験デザイン。
- ② 第1段階終了時に、早期エンドポイント（代替エンドポイント）を用いて早期無効中止の決定を行う。
- ③ 最終検定の検定統計量は第1段階と第2段階の両臨床エンドポイント（臨床的に興味のあるエンドポイント）のデータを用いて算出する。
- ④ 第1段階に登録される症例においては、早期エンドポイントと臨床エンドポイントの両方のデータを収集する。第2段階に登録される症例においては、臨床エンドポイントのみを収集する。
- ⑤ 早期エンドポイントに関する検定は行わないため、最終検定に対する有意水準の調整は必要としない。

また、このデザインの流れは図 2-1 のとおりである。

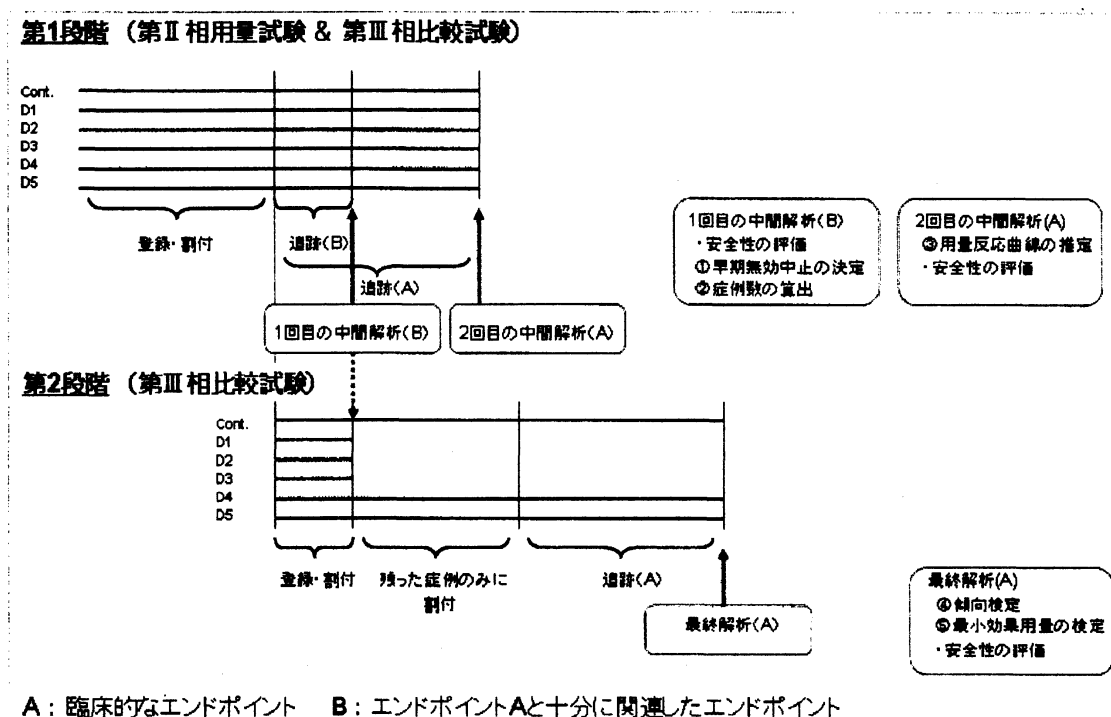


図 2-1 第Ⅱ相/第Ⅲ相結合デザインの流れ

Liu & Pledger は、多群比較デザインを提案する前段階として 2 群比較のデザインを提案した。本研究では 2 群比較での拡張方法を提案することを目的としているため、以下に 2 群比較のデザインの流れを示す (図 2-2)。

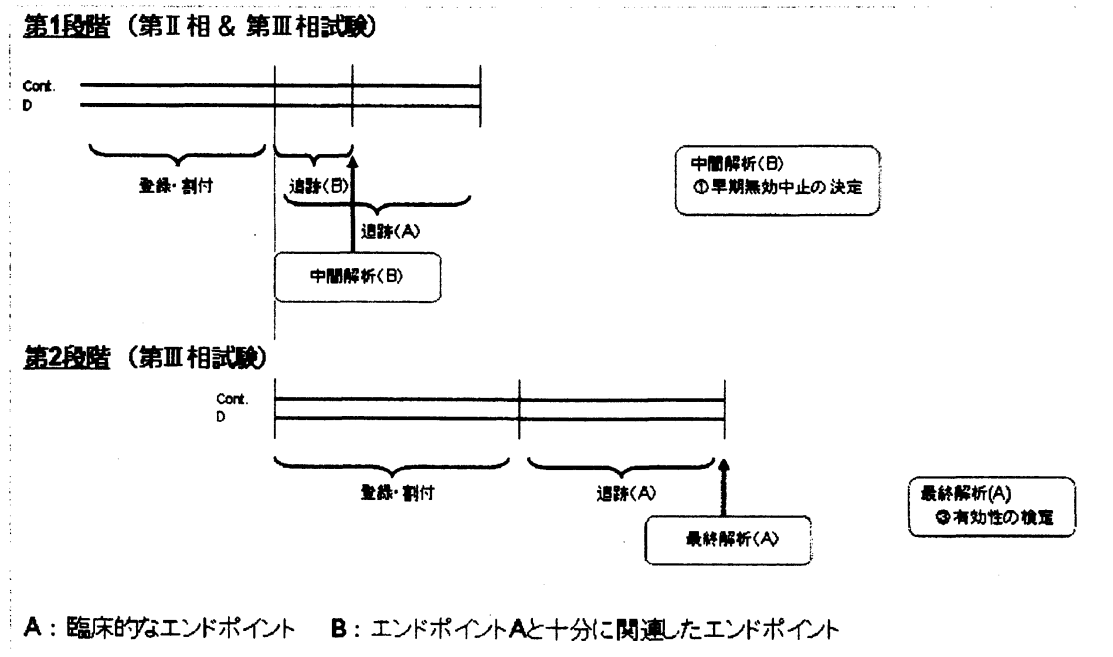


図 2-2 2 群比較の場合のデザインの流れ

試験デザインは、早期無効中止を許す 2 段階試験とする。第 2 段階のエンドポイントは臨床エンドポイントとし、第 1 段階のエンドポイントは、臨床的なエンドポイントと十分に関連があり、それよりも早期にデータを収集することが可能な早期エンドポイント (代替エンドポイント) とする。

第 1 段階では、各群に  $n_1$  人の患者をランダムに割り付ける。この  $n_1$  人においては、早期・臨床の両エンドポイントを収集する。第 1 段階の登録が終了したのち、試験を中断することなく第 2 段階の登録を開始する。従って、中間解析が終了するまでは、第 1 段階と第 2 段階は時間的に重なっていることになる。

早期無効中止の決定は、第 1 段階終了時の中間解析により行う。このとき用いるデータは早期エンドポイントのデータとする。早期エンドポイントのデータから算出された検定統計量の p 値が、あらかじめ定めた早期無効中止とする確率より小さい場合、被験薬は無効中止と判断し、その時点で試験を早期中止する。それ以外の場合には第 2 段階へと進む。

第 2 段階に進んだ場合には、引き続き症例登録を行い、各群計  $n_2$  人の患者をランダムに割り付ける。最終検定は、第 1 段階と第 2 段階の臨床エンドポイントのデータから算出した検定統計量を用い、試験全体の帰無仮説に対して検定を行う。

このデザインの欠点としては、被験薬が無効の場合には、中間解析が終了するまでに集積した症例が無駄になってしまう、という点が挙げられる。一方、利点としては、被験薬が有効の場合には、第Ⅱ相と第Ⅲ相の間の時間、および従来の第Ⅲ相試験を行う場合よりも第Ⅱ相の症例数が削減できるという点が挙げられる。

### 3. 第Ⅲ相のエンドポイントのみが二項分布に従う場合

#### ～ 潜在変数を用いた拡張法の提案 ～

本研究では、骨粗鬆症の臨床試験を想定し、第1段階の早期エンドポイントとして骨密度（連続データ）、第2段階の臨床エンドポイントとして骨折の発現の有無（二値データ）を考える。骨密度が上昇するにつれて骨折のリスクは低下すると考えられるため、両エンドポイントの関連を考える必要がある。この場合重要となるのは、早期エンドポイントと臨床エンドポイントの関連の考え方であるが、正規分布に従うデータと二項分布に従うデータの相関を直接考えることは複雑である。そこで、臨床エンドポイントの二値データが、潜在的に正規分布に従うデータであると仮定し、潜在分布に従うデータと早期エンドポイントの相関を考えることとした。

この拡張方法においても、試験デザインは Liu & Pledger と同様、2段階試験とし、第1段階の  $n_1$  人においては早期・臨床の両エンドポイントを、第2段階の  $n_2$  人においては臨床エンドポイントのみを収集する。

#### 3. 1 潜在変数を用いた拡張方法

まず、被験薬群とコントロール群の第  $i$  症例 ( $i=1, \dots, n$ ) の早期エンドポイント  $X_i$  および  $Y_i$  は、それぞれ以下の正規分布に独立に従う確率変数であると仮定する。

$$X_i \sim N(\mu_1, \sigma_1^2) \quad , \quad Y_i \sim N(\mu_2, \sigma_2^2)$$

ただし、実際に観測されるのは  $n_1$  人のデータであるが、潜在的には  $n$  人のデータが存在すると仮定する。

次に、被験薬群とコントロール群の第  $s$  段階 ( $s=1, 2$ ) の第  $i$  症例 ( $i=1, \dots, n_s$ ) の臨床エンドポイント  $A_i$  および  $B_i$  は、それぞれ以下の成功確率を持つ独立なベルヌーイ確率変数であると仮定する。

$$A_i \sim \text{Bernoulli}(p_A) \quad , \quad B_i \sim \text{Bernoulli}(p_B)$$

ここで、 $X_i$  と  $A_i$  および  $Y_i$  と  $B_i$  の関連性を考慮するために、 $A_i$  および  $B_i$  に対する潜在変数  $\tilde{A}_i$ 、 $\tilde{B}_i$  を考える。それぞれの潜在変数  $\tilde{A}_i$ 、 $\tilde{B}_i$  は以下の正規分布に従う確率変数であると

仮定する.

$$\tilde{A}_i \sim N(\nu_1, \tau_1^2) \quad , \quad \tilde{B}_i \sim N(\nu_2, \tau_2^2)$$

このとき, ベルヌーイ確率変数  $A$  は,  $\tilde{A} \geq 0$  のときに 1,  $\tilde{A} < 0$  のときに 0 をとる確率変数と考えられ, その成功確率  $p_A$  は,

$$p_A = P(A=1) = \int_0^{\infty} f(\tilde{a}) d\tilde{a}$$

から求まる. 同様に, ベルヌーイ確率変数  $B$  は,  $\tilde{B} \geq 0$  のときに 1,  $\tilde{B} < 0$  のときに 0 をとる確率変数と考えられ, その成功確率  $p_B$  は,

$$p_B = P(B=1) = \int_0^{\infty} f(\tilde{b}) d\tilde{b}$$

から求まる.

$X_i$  および  $\tilde{A}_i$  の相関係数を  $\rho_1$  とすると, 同時分布は以下のとおりである.

$$\begin{pmatrix} X_i \\ \tilde{A}_i \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_1 \\ \nu_1 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_1 \tau_1 \rho_1 \\ & \tau_1^2 \end{pmatrix} \right)$$

ここで, 計算を簡単に行うために,  $\tau_1^2 = 1/(1-\rho_1^2)$  とおくと,

$$\begin{pmatrix} X_i \\ \tilde{A}_i \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_1 \\ \nu_1 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_1 \rho_1 / \sqrt{1-\rho_1^2} \\ & 1/(1-\rho_1^2) \end{pmatrix} \right)$$

となる. このとき,  $X_i$  で条件付けた  $\tilde{A}_i$  の条件付分布は,

$$\tilde{A}_i | X_i \sim N \left( \nu_1 + \frac{\rho_1 (X_i - \mu_1)}{\sigma_1 \sqrt{1-\rho_1^2}}, 1 \right)$$

となる. 標準正規分布の分布関数を  $\Phi(\cdot)$  で表わすと,

$$P(A_i = 1 | X_i) = \int_0^{\infty} f(\tilde{a}_i | x_i) d\tilde{a}_i = \Phi \left( \nu_1 + \frac{\rho_1 (X_i - \mu_1)}{\sigma_1 \sqrt{1-\rho_1^2}} \right)$$

となり, ベルヌーイ確率変数の成功確率  $p_A$  は以下のとおり算出される.

$$p_A = \int_{-\infty}^{\infty} \Phi \left( \nu_1 + \frac{\rho_1 (X_i - \mu_1)}{\sigma_1 \sqrt{1-\rho_1^2}} \right) f(x_i) dx_i$$

となる. 同様に,  $Y_i$  および  $\tilde{B}_i$  の相関係数を  $\rho_2$  とし,  $\tau_2^2 = 1/(1-\rho_2^2)$  とすると, 同時分布は,

$$\begin{pmatrix} Y_i \\ \tilde{B}_i \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_2 \\ \nu_2 \end{pmatrix}, \begin{pmatrix} \sigma_2^2 & \sigma_2 \rho_2 / \sqrt{1 - \rho_2^2} \\ \sigma_2 \rho_2 / \sqrt{1 - \rho_2^2} & 1 / (1 - \rho_2^2) \end{pmatrix} \right)$$

となる。\$Y\_i\$ で条件付けた \$\tilde{B}\_i\$ の条件付分布は、

$$\tilde{B}_i | Y_i \sim N \left( \nu_2 + \frac{\rho_2 (Y_i - \mu_2)}{\sigma_2 \sqrt{1 - \rho_2^2}}, 1 \right)$$

となる。標準正規分布の分布関数を \$\Phi(\cdot)\$ で表わすと、

$$P(B_i = 1 | Y_i) = \int_0^\infty f(\tilde{b}_i | y_i) d\tilde{b}_i = \Phi \left( \nu_2 + \frac{\rho_2 (Y_i - \mu_2)}{\sigma_2 \sqrt{1 - \rho_2^2}} \right)$$

となり、ベルヌーイ確率変数の成功確率 \$p\_B\$ は以下のとおり算出される。

$$p_B = \int_{-\infty}^\infty \Phi \left( \nu_2 + \frac{\rho_2 (Y_i - \mu_2)}{\sigma_2 \sqrt{1 - \rho_2^2}} \right) f(y_i) dy_i$$

### 3. 2 検定統計量の算出

この試験において最も興味のあることは、臨床エンドポイントに関して被験薬群とコントロール群で差があるか、ということである。そこで、試験全体の帰無仮説は、

$$H_0 : p_A = p_B = p$$

となる。これに対し、対立仮説は、

$$H_1 : p_A < p_B$$

とし、有意水準 \$\alpha\$、検出力 \$1 - \beta\$ の片側検定を行う。

臨床エンドポイントに関する検定統計量 \$Z\$ は、第 1 段階および第 2 段階の両方の臨床エンドポイントを用いて以下のとおり算出する。

$$Z = \frac{\sum_{i=1}^n B_i / n - \sum_{i=1}^n A_i / n}{\sqrt{\frac{2}{n} \tilde{p}(1 - \tilde{p})}} = \frac{\sqrt{n} (\sum_{i=1}^n B_i / n - \sum_{i=1}^n A_i / n)}{\sqrt{2 \tilde{p}(1 - \tilde{p})}}$$

ここで、\$A\_i\$、\$B\_i\$ はそれぞれ被験薬群とコントロール群の臨床エンドポイントに関するデータ、\$p\_A\$、\$p\_B\$ はそれぞれ被験薬群とコントロール群の成功確率、\$n\_1\$、\$n\_2\$ はそれぞれ第 1 段階および第 2 段階の各群の症例数とし、\$n = n\_1 + n\_2\$、また、\$\tilde{p}\$ は両群の併合した分散とし、

\$\tilde{p} = (\sum\_{i=1}^n A\_i + \sum\_{i=1}^n B\_i) / 2n\$ である。\$n\$ が十分大きく、かつ帰無仮説のもとでは、\$Z\$ は漸近



的に標準正規分布に従う。最終的な判断は、 $Z \geq z_\alpha$  となった場合には帰無仮説  $H_0$  を棄却して両群に差があるとし、 $Z < z_\alpha$  となった場合には、帰無仮説  $H_0$  を受容して両群に差は見られないとする。

次に、早期エンドポイントに関する検定統計量  $\tilde{Z}$  の算出方法を示す。早期エンドポイントの検定統計量  $\tilde{Z}$  は、第1段階の  $n_1$  人から集積されたデータのみを用いて以下のように算出する。

$$\tilde{Z} = \frac{\sum_{i=1}^{n_1} X_i/n_1 - \sum_{i=1}^{n_1} Y_i/n_1}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n_1}} = \frac{\sqrt{n_1}(\sum_{i=1}^{n_1} X_i/n_1 - \sum_{i=1}^{n_1} Y_i/n_1)}{\sqrt{(\sigma_1^2 + \sigma_2^2)}}$$

ただし、 $X_i$  および  $Y_i$  はそれぞれ被験薬群およびコントロール群の早期エンドポイントに関するデータ、 $\sigma_1^2$  および  $\sigma_2^2$  はそれぞれ被験薬群およびコントロール群の分散、 $n_1$  は第1段階の各群の症例数とする。 $n_1$  が十分大きいとき、 $\tilde{Z}$  は漸近的に、

$$\tilde{Z} \sim N\left(\frac{\sqrt{n_1}(\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}}, 1\right)$$

の正規分布に従う。

### 3.3 第1段階の症例数 $n_1$ および早期無効中止の判断基準の選択

第1段階の症例数  $n_1$  および早期無効中止の判断基準の選択については、Liu & Pledger と同様の方法により決定することとした。つまり、第1段階の症例数  $n_1$  は早期無効中止の確率を選択することによって決定する。

帰無仮説  $H_0$  が真の時に試験を早期に中止する確率を  $(1 - \alpha_1^*)$  とし、 $\tilde{Z} < z_{\alpha_1^*}$  となった場合に試験を早期に中止することとする。また、対立仮説  $H_1$  が真の時に試験を早期に中止する確率を  $\beta_1^*$  とすると、第1段階の症例数  $n_1$  は、

$$P_{H_1} \left\{ \tilde{Z} \geq z_{\alpha_1^*} \right\} = 1 - \beta_1^*$$

を満たすよう算出することができる。つまり、

$$\begin{aligned} & P_{H_1} \left\{ \tilde{Z} \geq z_{\alpha_1^*} \right\} \\ &= P_{H_1} \left\{ \frac{\sqrt{n_1}(\sum_{i=1}^{n_1} X_i/n_1 - \sum_{i=1}^{n_1} Y_i/n_1)}{\sqrt{\sigma_1^2 + \sigma_2^2}} - \frac{\sqrt{n_1}(\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} \geq z_{\alpha_1^*} - \frac{\sqrt{n_1}(\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right\} = 1 - \beta_1^* \end{aligned}$$

より,

$$z_{\alpha_1^*} - \frac{\sqrt{n_1}(\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} = z_{1-\beta_1^*}$$

となり,

$$n_1 = \{z_{\alpha_1^*} + z_{\beta_1^*}\}^2 \cdot (\sigma_1^2 + \sigma_2^2) / (\mu_1 - \mu_2)^2$$

となる. この計算に必要な $\alpha_1^*$ および $\beta_1^*$ の選択方法については, 4. 5で示す.

### 3. 4 第2段階の症例数 $n_2$ の算出

Liu & Pledgerの方法では, 第2段階の症例数 $n_2$ は, 試験全体の検出力 $P_{r_{\min}}\{Z \geq z_\alpha, \tilde{Z}_1 \geq z_{\alpha_1^*}\}$ が少なくとも $1-\beta$ となるよう, 近似的な方法を用いて算出していた.

しかし, ここでは, 試験全体の検出力が $1-\beta$ , つまり,  $P_{H_1}\{Z \geq z_\alpha, \tilde{Z} \geq z_{\alpha_1^*}\} = 1-\beta$ となるよう直接計算を行うこととした.

まず,  $\tilde{Z}$ と $Z$ の同時分布を考える.  $\tilde{Z}$ と $Z$ の同時分布は,

$$\begin{pmatrix} \tilde{Z} \\ Z \end{pmatrix} \sim N \left( \begin{pmatrix} \sqrt{n_1}(\mu_1 - \mu_2) / \sqrt{\sigma_1^2 + \sigma_2^2} \\ \sqrt{n}(p_B - p_A) / \sqrt{D} \end{pmatrix}, \begin{pmatrix} 1 & \text{cov}[\tilde{Z}, Z] \\ & 1 \end{pmatrix} \right) \quad (3. 1)$$

となる. ここで,  $D = p_A(1-p_A) + p_B(1-p_B)$ ,  $\text{cov}[\tilde{Z}, Z]$ は $\tilde{Z}$ と $Z$ の共分散である.

$\tilde{Z}$ と $Z$ の共分散 $\text{cov}[\tilde{Z}, Z]$ は, 共分散の定義より以下のように算出する.

$$\begin{aligned} & \text{cov}[\tilde{Z}, Z] \\ &= E\left[\left[\tilde{Z} - E(\tilde{Z})\right] \cdot [Z - E(Z)]\right] \\ &= E\left[\left[\left\{\frac{\sqrt{n_1}\left(\sum_{i=1}^{n_1} X_i / n_1 - \sum_{i=1}^{n_1} Y_i / n_1\right)}{\sqrt{\sigma_1^2 + \sigma_2^2}} - \frac{\sqrt{n_1}(\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right\} \right. \right. \\ & \quad \left. \left. \times \left[\left\{\frac{\sqrt{n}\left(\sum_{i=1}^n B_i / n - \sum_{i=1}^n A_i / n\right)}{\sqrt{D}} - \frac{\sqrt{n}(p_B - p_A)}{\sqrt{D}}\right\}\right]\right]\right] \end{aligned}$$

$$\begin{aligned}
&= \frac{\sqrt{n_1 n}}{\sqrt{(\sigma_1^2 + \sigma_2^2)D}} E \left\{ \left[ \left( \sum_{i=1}^{n_1} X_i / n_1 - \mu_1 \right) - \left( \sum_{i=1}^{n_1} Y_i / n_1 - \mu_2 \right) \right] \right. \\
&\quad \left. \times \left[ \left( \sum_{i=1}^n B_i / n - p_B \right) - \left( \sum_{i=1}^n A_i / n - p_A \right) \right] \right\} \\
&= - \frac{\sqrt{n_1 n}}{\sqrt{(\sigma_1^2 + \sigma_2^2)D}} \\
&\quad \times E \left\{ \left( \sum_{i=1}^{n_1} X_i / n_1 - \mu_1 \right) \cdot \left( \sum_{i=1}^n A_i / n - p_A \right) + \left( \sum_{i=1}^{n_1} Y_i / n_1 - \mu_2 \right) \cdot \left( \sum_{i=1}^n B_i / n - p_B \right) \right\} \\
&= - \frac{\sqrt{n_1 n}}{\sqrt{(\sigma_1^2 + \sigma_2^2)D}} \left\{ \frac{1}{n_1 n} \sum_{i=1}^{n_1} E(X_i - \mu_1) \cdot (A_i - p_A) + \frac{1}{n_1 n} \sum_{i=1}^{n_1} E(Y_i - \mu_2) \cdot (B_i - p_B) \right\} \\
&= - \sqrt{\frac{n_1}{n}} \frac{1}{\sqrt{(\sigma_1^2 + \sigma_2^2)D}} \{ E(X - \mu_1) \cdot (A - p_A) + E(Y - \mu_2) \cdot (B - p_B) \}
\end{aligned}$$

ここで,  $E(X - \mu_1) \cdot (A - p_A)$  および  $E(Y - \mu_2) \cdot (B - p_B)$  について考える. まず,

$E(X - \mu_1) \cdot (A - p_A)$  は共分散の定義より,

$$\begin{aligned}
&E(X - \mu_1) \cdot (A - p_A) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_1) \cdot (a - p_A) dF(x, a) \\
&= \int_{-\infty}^{\infty} \sum_{j=0}^1 (x - \mu_1) \cdot (a - p_A) P(a = j | x) f(x) dx \\
&= \int_{-\infty}^{\infty} \{ (x - \mu_1) \cdot (0 - p_A) P(a = 0 | x) + (x - \mu_1) \cdot (1 - p_A) P(a = 1 | x) \} f(x) dx
\end{aligned}$$

となる. ここで, 4. 1の潜在変数の考え方をい用いると,  $P(a_i = 0 | x_i)$  および  $P(a_i = 1 | x_i)$  は,

$$P(a_i = 0 | x_i) = 1 - \Phi \left( v_1 + \frac{\rho_1(x_i - \mu_1)}{\sigma_1 \sqrt{1 - \rho_1^2}} \right) \quad (3.2)$$

および,

$$P(a_i = 1 | x_i) = \Phi \left( v_1 + \frac{\rho_1(x_i - \mu_1)}{\sigma_1 \sqrt{1 - \rho_1^2}} \right) \quad (3.3)$$

となる. これを代入すると,

$$E(X - \mu_1) \cdot (A - p_A)$$

$$= \int_{-\infty}^{\infty} (x - \mu_1) \left\{ \Phi \left( v_1 + \frac{\rho_1(x - \mu_1)}{\sigma_1 \sqrt{1 - \rho_1^2}} \right) - p_A \right\} f(x) dx \quad (3.4)$$

となる。同様に  $E(Y - \mu_2) \cdot (B - p_B)$  は、

$$E(Y - \mu_2) \cdot (B - p_B) = \int_{-\infty}^{\infty} (y - \mu_2) \left\{ \Phi \left( v_2 + \frac{\rho_2(y - \mu_2)}{\sigma_2 \sqrt{1 - \rho_2^2}} \right) - p_B \right\} f(y) dy \quad (3.5)$$

となる。

以上より、 $\tilde{Z}$  と  $Z$  の同時分布が求まる。この同時分布を用いて、

$$P_{H_1} \{ Z \geq z_\alpha, \tilde{Z} \geq z_{\alpha_1} \} = \int_{z_\alpha}^{\infty} \int_{z_{\alpha_1}}^{\infty} f(\tilde{z}, z) d\tilde{z} dz = 1 - \beta \quad (3.6)$$

を満たす  $n$  を数値計算により求める。また、第2段階の症例数  $n_2$  は、

$$n_2 = n - n_1 \quad (3.7)$$

により算出する。

### 3.5 最適な $\alpha_1^*$ と $\beta_1^*$ の選択

臨床試験をポートフォリオの観点からみると、症例数を最小限に抑えることは臨床試験の費用を抑えることにつながり、大変有意義なことである。そこで、症例数を算出する際に必要となる  $\alpha_1^*$  と  $\beta_1^*$  を、試験をとおしての期待総症例数が最小となるように選択する。

仮に、第2段階の症例数を  $\tilde{n}_2$  とする。第1段階の登録が終了してから第1段階の中間解析の結果が出るまでの間に第2段階に割り付けられる症例数を  $n_{20}$  とすると、第2段階の症例数は以下の式で与えられる。

$$\begin{cases} \tilde{Z}_1 < z_{\alpha_1} \text{ なら, } \tilde{n}_2 = n_{20} \\ \tilde{Z}_1 \geq z_{\alpha_1} \text{ なら, } \tilde{n}_2 = n_2 \end{cases}$$

$\delta$  の事前確率を、 $P\{\delta = \delta_{\min}\} = \pi$  および  $P\{\delta = 0\} = 1 - \pi$  とする。各群の総症例数は  $n = n_1 + \tilde{n}_2$  であるので、期待総症例数  $E_{(\pi)}(N)$  は以下のように算出することができる。

$$E_{(\pi)}(N) = n_1 + \pi E_{\tau_{\min}}(\tilde{N}_2) + (1 - \pi) E_0(\tilde{N}_2)$$

ここで、

$$E_{\tau_{\min}}(\tilde{N}_2) = n_{20} \times \beta_1^* + n_2 \times (1 - \beta_1^*)$$

$$E_0(\tilde{N}_2) = n_{20} \times (1 - \alpha_1^*) + n_2 \times \alpha_1^*$$

であることに注意すると、期待総症例数は、

$$E_{(\pi)}(N) = (n_1 + n_{20}) + (n_2 - n_{20}) \{ \pi(1 - \beta_1^*) + (1 - \pi)\alpha_1^* \} \quad (3.8)$$

この式から、期待総症例数  $E_{(\pi)}(N)$  を最小とする  $\alpha_1^*$  および  $\beta_1^*$  を選択する。ただし、 $\alpha \leq \alpha_1^* < 1$  および  $0 < \beta_1^* < \beta$  である。

### 3.6 シミュレーション結果および考察

#### 3.6.1 症例数の算出

式(3.1)の同時分布を用い、式(3.6)を満たす  $n$  を算出するため、SASを用いた数値計算を行った。数値計算を行うにあたり、各エンドポイントの effect size を

$$\text{早期エンドポイント; } \varepsilon_1 = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

$$\text{臨床エンドポイント; } \varepsilon_2 = (p_B - p_A) / \sqrt{p_A(1 - p_A) + p_B(1 - p_B)}$$

とし、これらの値を変えることによってそのときの症例数を算出した。

- ①  $\alpha$ ,  $\beta$ ,  $\varepsilon_2$  ( $p_A$ ,  $p_B$ ),  $\varepsilon_1$ ,  $\rho_1$ ,  $\rho_2$  を与える。
- ② 式(3.2) および (3.3) を満たす  $v_1$ ,  $v_2$  を算出する。
- ③ 式(3.4) および (3.5) より  $E(X - \mu_1) \cdot (A - p_A)$  および  $E(Y - \mu_2) \cdot (B - p_B)$  を算出する。
- ④  $\alpha \leq \alpha_1^* < 1$  および  $0 < \beta_1^* < \beta$  の範囲内で  $\alpha_1^*$  および  $\beta_1^*$  のさまざまな組み合わせを与え、第1段階の症例数  $n_1$  を算出する。
- ⑤ SAS の probnorm 関数を用いて  $n$  を  $n \geq n_1 + 1$  の範囲で動かしながら式(3.6)を満たす  $n$  を求める。ただし、probnorm は以下の同時分布の確率を算出する SAS 関数である。

$$\text{probnorm}(a, b, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^a \int_{-\infty}^b \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) dx dy$$

- ⑥ 式(3.7)より  $n_2$  を求め、式(3.8)より期待総症例数  $E_{(\pi)}(N)$  を算出する。

$E_{(\pi)}(N)$ が最小となるときの $\alpha_1^*$ および $\beta_1^*$ を最適な $\alpha_1^*$ および $\beta_1^*$ とし、そのときの $n$ を総症例数とする。ただし、このとき用いる $n_{20}$ は3. 6. 3に示した理由から、 $n_{20} = 50$ とした。また $\pi$ は、どちらの仮説も等しい確率で生じると仮定し、 $\pi = 0.5$ とした。

さまざまな $\alpha$ ,  $\beta$ ,  $\varepsilon_2$  ( $p_A$ ,  $p_B$ ),  $\varepsilon_1$ ,  $\rho_1$ ,  $\rho_2$ を与えた場合の総症例数 $n$ , およびそのときの $\alpha_1^*$ ,  $\beta_1^*$ の値を【表 3-1】に示す。

### 3. 6. 2 検出力曲線

算出された症例数で試験を実施した場合に検出力が担保されているかを確認するため、SAS を用いてモンテカルロ・シミュレーションを行った。算出された症例数、およびこのとき与えた値を用い、第2段階の effect size を変更しながら検出力曲線を作成した。結果を【図 3-1】に示す。

### 3. 6. 3 潜在変数を用いたデザインと従来のデザインとの比較

提案した潜在変数を用いたデザインと従来のデザインとの比較を行うために、期待総症例数および期待総期間を算出した。期待総症例数および期待総期間を算出するために必要な設定は以下のとおりとした。

#### 【設定】

- 1) 早期エンドポイントの追跡期間は2ヶ月とし、データ固定および解析には1ヶ月を要するものとする。
- 2) 臨床エンドポイントの追跡期間は1年(12ヶ月)とする。
- 3) 従来の試験においては、第II相と第III相の間の期間は、データ固定、解析、および第III相の計画を含め、6ヶ月を要するものとする。
- 4) 月々各群25人の患者が登録されるものとする。ただし、提案するデザインにおける追跡・データ固定・解析に要する3ヶ月間は登録が消極的になると仮定し、 $n_{20} = 50$ 人とする。その後第2段階および従来の第III相においては、再び月々各群25人の患者が登録されるものとする。

以上の設定より、総症例数(人)および期待総期間(ヶ月)を以下のように算出する。

#### ① 期待総症例数

提案する方法 ; 式(3. 8)

従来の方法 ;  $E_{c(\pi)}(N) = n_{c1} + n_{c2} \{ \pi(1 - \beta) + (1 - \pi)\alpha \}$

② 対立仮説のもとでの期待総症例数

$$\text{提案する方法 ; } E_{H1}(N) = n_1 + \beta_1^* \cdot n_{20} + (1 - \beta_1^*) \cdot n_2$$

$$\text{従来の方法 ; } E_{cH1}(N) = n_{c1} + (1 - \beta) \cdot n_{c2}$$

③ 対立仮説のもとでの期待総期間

$$\text{提案する方法 ; } E_{H1}(T) = \left( \frac{n_1}{25} + 3 \right) + (1 - \beta_1^*) \cdot \left\{ \frac{(n_2 - n_{20})}{25} + 12 \right\}$$

$$\text{従来の方法 ; } E_{cH1}(T) = \left( \frac{n_{c1}}{25} + 6 \right) + (1 - \beta) \cdot \left\{ \frac{n_{c2}}{25} + 12 \right\}$$

④ 帰無仮説のもとでの期待総症例数

$$\text{提案する方法 ; } E_{H0}(N) = n_1 + (1 - \alpha_1^*) \cdot n_{20} + \alpha_1^* \cdot n_2$$

$$\text{従来の方法 ; } E_{cH0}(N) = n_{c1} + \alpha \cdot n_{c2}$$

⑤ 帰無仮説のもとでの期待総期間

$$\text{提案する方法 ; } E_{H0}(T) = \left( \frac{n_1}{25} + 3 \right) + \alpha_1^* \cdot \left\{ \frac{(n_2 - n_{20})}{25} + 12 \right\}$$

$$\text{従来の方法 ; } E_{cH0}(T) = \left( \frac{n_{c1}}{25} + 6 \right) + \alpha \cdot \left\{ \frac{n_{c2}}{25} + 12 \right\}$$

また、このときの  $n_{c1}$  および  $n_{c2}$  は以下のとおり算出する。

$$n_{c1} = \frac{(z_\alpha + z_\beta)^2}{(\mu_1 - \mu_2)^2} \cdot (\sigma_1^2 + \sigma_2^2)$$

$$n_{c2} = \frac{(z_\alpha + z_\beta)^2}{(p_B - p_A)^2} \cdot \{p_A(1 - p_A) + p_B(1 - p_B)\}$$

これらの手順に従って算出した期待総症例数および期待総期間の結果を【表 3-2】および【表 3-3】に示す。

### 3. 6. 4 結果

$\alpha = 0.025$ ,  $\beta = 0.2$ ,  $\rho_1 = \rho_2 = -0.9$  とした場合の結果を以下に示す。

表 3-1 症例数および $\alpha_1^*$ ,  $\beta_1^*$ 

$\varepsilon_1$	$\varepsilon_2$	$E_r(N)$	$\alpha_1^*$	$\beta_1^*$	$n_1$	$n_2$	$n$
0.2	0.14	321	0.251	0.055	127	290	417
0.3	0.14	287	0.116	0.048	90	324	414
0.3	0.22	150	0.33	0.044	51	126	177
0.3	0.29	88	0.956	0.197	8	83	91
0.4	0.14	267	0.067	0.037	67	340	407
0.4	0.22	141	0.192	0.043	42	135	177
0.4	0.29	87	0.954	0.199	5	86	91
0.4	0.38	55	0.994	0.034	3	52	55

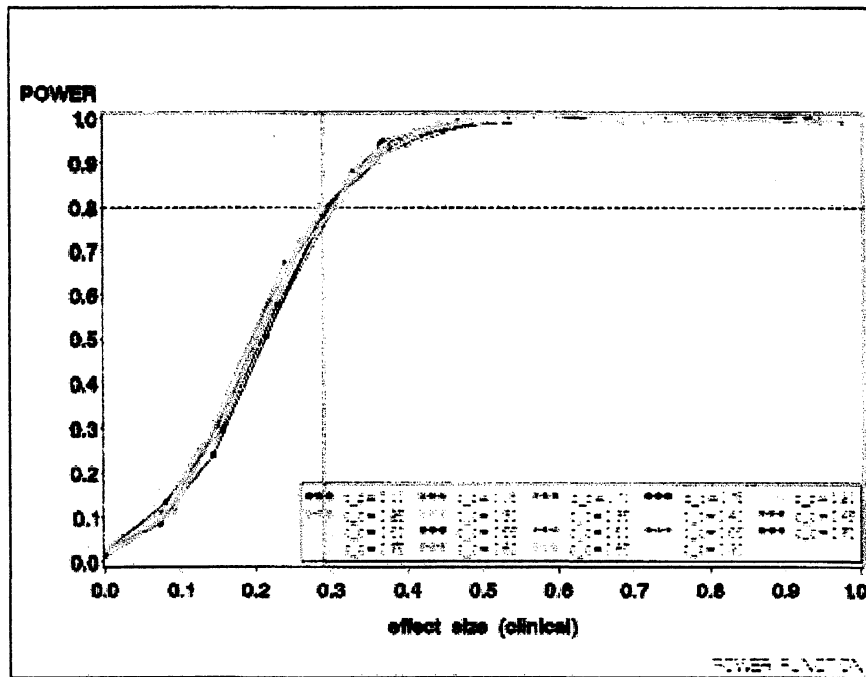
図 3-1 検出力曲線 ( $\varepsilon_1 = 0.4$ ,  $\varepsilon_2 = 0.29$ , シミュレーション回数: 1000 回)



表 3-2 帰無仮説のもとでの期待総症例数および期待総期間

$\varepsilon_1$	$\varepsilon_2$	$\alpha_1^*$	$\beta_1^*$	期待総症例数		期待総期間	
				提案	従来	提案	従来
0.2	0.14	0.251	0.055	238	203	13.5	14.41
0.3	0.14	0.116	0.048	172	96	9.26	10.13
0.3	0.22	0.33	0.044	127	91	10	9.91
0.3	0.29	0.956	0.197	90	89	16.05	9.83
0.4	0.14	0.067	0.037	137	59	7.26	8.65
0.4	0.22	0.192	0.043	109	54	7.64	8.43
0.4	0.29	0.954	0.199	90	52	16.02	8.35
0.4	0.38	0.994	0.034	55	51	15.13	8.32

表 3-3 対立仮説のもとでの期待総症例数および期待総期間

$\varepsilon_1$	$\varepsilon_2$	$\alpha_1^*$	$\beta_1^*$	期待総症例数		期待総期間	
				提案	従来	提案	従来
0.2	0.14	0.251	0.055	404	501	28.49	35.64
0.3	0.14	0.116	0.048	401	394	28.46	31.36
0.3	0.22	0.33	0.044	174	220	19.42	24.38
0.3	0.29	0.956	0.197	85	159	14.02	21.95
0.4	0.14	0.067	0.037	397	357	28.41	29.88
0.4	0.22	0.192	0.043	174	183	19.42	22.90
0.4	0.29	0.954	0.199	84	122	13.97	20.47
0.4	0.38	0.994	0.034	55	93	14.79	19.32

### 3. 6. 5 考察

【表 3-1】より、早期エンドポイントの effect size が等しい場合には、臨床エンドポイントの effect size が大きくなるにつれ、ほとんどの場合において無効中止の確率  $(1-\alpha_1^*)$  が小さくなることがわかった。しかし、 $\beta_1^*$  には傾向が見られなかった。臨床エンドポイントの effect size が大きくなるにつれて、無効中止の確率は小さくなり、第 2 種の過誤は小さくなると予想したが、一部の  $(1-\alpha_1^*)$  と  $\beta_1^*$  に関しては予想に反した結果となった。この原因については、今後、他の設定においてシミュレーションを行い、十分に検討する必要がある。

【図 3-1】より、算出された症例数およびこのときの effect size により、検出力が担保で

きていることがわかった。また、effect size が過大評価されていた場合には十分な検出力を得ることができなかった。どの試験にも共通していえることだが、effect size の選択には十分な考慮が必要であることがこの結果からも示唆された。

【表 3-2】より、帰無仮説のもとでの期待総症例数は、すべての effect size の組み合わせにおいて提案するデザインのほうが多かった。これは、第 1 段階が終了してから中間解析の結果が出るまでに集積される症例数が原因であると考えられる。

【表 3-3】より、対立仮説のもとでの期待総症例数は、effect size の組み合わせにより、提案するデザインのほうが大きい場合と小さい場合の結果が得られた。対立仮説のもとでの期待総症例数が、提案するデザインのほうが少ないものに注目すると、早期・臨床両エンドポイントの effect size の間にある程度の関連がある場合には、期待総症例数および期待総期間ともに削減できることがわかった。また、期待総症例数を削減できなかった effect size の組み合わせにおいても、期待総症例数を削減できていることがわかった。したがって、症例数の集積が多く、時間の短縮のみに焦点をおきたい試験においては、どの effect size の組み合わせでもこのデザインが有用であるといえる。しかし、症例数と試験期間の両方に焦点をおく場合には、両エンドポイントの関連が重要となってくるため、これらについて十分検討する必要があるといえる。

また、帰無仮説および対立仮説のもとでの期待総症例数および期待総期間の両方を併せて考えると、提案するデザインにはトレードオフの関係が存在しており、したがって、このデザインを用いる場合には、適用する試験の目的に応じて、この点を十分に検討する必要があるといえる。

以上のシミュレーション結果より、今回提案したデザインは、早期・臨床両エンドポイントに十分な関連のある場合には、対立仮説のもとでの期待総症例数および期待総期間を削減し、十分な関連がない場合でも、試験期間を短縮できるということが示された。したがって、被験薬の有効性を評価する試験においては大変有用なデザインであるといえる。また、今回は、骨密度と骨折の有無のような負の相関を持つエンドポイントに焦点を当て、症例数の算出を行った。しかし、血圧や高脂血症などにおける血圧降下率またはコレステロール値降下率と心血管障害の関係のような正の相関のあるエンドポイントについても、正の相関係数を用いることによって適用可能なデザインとなっている。したがって、このデザインを適用する際には、エンドポイントの特性と、早期・臨床の両エンドポイントの関連を十分に検討した上で適用することが望まれる。

#### 4. おわりに

本研究では、症例数の削減と時間の短縮に焦点をあて、さまざまな臨床試験に適用可能となるよう、エンドポイントが二値データである場合に拡張したデザインを提案した。そ

の方法として、潜在変数を用いた方法を提案したが、早期・臨床の両エンドポイントに十分な関連がある場合には、症例数および時間を削減することが可能であることがわかった。限られた設定でのシミュレーションであるが、いずれの方法においても、対立仮説のもとで症例数と時間を削減するという観点から、大変有用なデザインであるといえる。今後は、実際に行われた臨床試験のデータに、提案する方法を事後的に適用し、その性能を検討してこのデザインの有用性を示したい。また、その際には、症例数と時間だけでなく、臨床試験に関する多くの側面、例えば、試験が中止に終わった場合のリスクや、患者だけでなく、試験を実施する側の費用についても十分に検討する必要がある。それらすべての総合的な判断を加え、新たな医薬品開発のアプローチに適する柔軟かつ効率的なデザインを検討していくことが必要であると考えられる。

また、今回提案したデザインは、どちらも 2 群比較の場合での拡張方法であった。実際の第 II 相試験は、主に用量反応関係を検証する試験であり、多群比較を行う場合が多い。したがって、今後はこれらの方法を 3 群以上の比較の場合に拡張した方法についても検討したい。また、二値データだけでなく生存時間などの他のエンドポイントを用いる場合への拡張も可能であると考えられる。

## 5. 参考文献

1. 米国食品医薬品局 Whitepaper (2004), "Innovation Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products,"  
<http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html#execsummary>
2. 竹内正弘 (2005), "The Impact of Critical Path Initiative on R&D in Japan,"  
[http://www.pharm.kitasato-u.ac.jp/biostatist/presentation2005/25/Session2/01Prof\\_Ta\\_keuchi\\_final.pdf](http://www.pharm.kitasato-u.ac.jp/biostatist/presentation2005/25/Session2/01Prof_Ta_keuchi_final.pdf)
3. Liu, Q., and Pledger, G. W. (2005), "Phase 2 and 3 Combination Designs to Accelerate Drug Development," *Journal of the American Statistical Association*, 100, 493-502.
4. Thall, P. F., Simon, R., and Ellenberg, S. S. (1988), "Two-stage selection and testing designs for comparative clinical trials," *Biometrika*, 75, 303-310.
5. Thall, P. F., Simon, R., and Ellenberg, S. S. (1989), "A Two-stage Design for Choosing Among Several Experimental Treatments and a Control in Clinical Trials," *Biometrics*, 45, 537-547.
6. Lehmann, E. L., "Elements of Large-Sample Theory," Springer
7. Molenberghs, G., Geys, H., and Buyse, M. (2001), "Evaluation of surrogate endpoints in randomized experiments with mixed discrete and continuous outcomes," *Statistics in Medicine*, 20, 3023-3038.

8. Armitage, P., McPherson, C. K., and Rowe, B. C. (1969), "Repeated Significance Tests on Accumulating Data," *Journal of the Royal Statistical Society. Series A*, 132, 235-244
9. Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., "NUMERICAL RECIPES in Fortran 77," CAMBRIDGE, Second Edition.
10. 日本公定書協会 編 (2004), "新薬評価ガイドライン," 薬事日報社.
11. Jennison, C., and Turnbull, B. W., "GROUP SEQUENTIAL METHOD with APPLICATIONS to CLINICAL TRIALS," CHAPMAN & HALL/CRC.
12. 竹村彰通 (1991), "現代数理統計学," 創文社.