

中間解析結果に基づく被験者数再設定方法の特性の検討

○高沢 翔¹⁾, 堂園 剛司¹⁾, 富金原 悟¹⁾, 森川 敏彦²⁾

1) 小野薬品工業株式会社 データサイエンス室

2) 元久留米大学

Consideration of the characteristics of sample size re-estimation based on the interim results

Akira Takazawa¹⁾, Koji Dozono¹⁾, Satoru Fukinbara¹⁾, Toshihiko Morikawa²⁾

1) Data Science, ONO Pharmaceutical Co., Ltd.

2) Former Kurume University

要旨

中間解析結果に基づいて被験者数再設定 (sample size re-estimation : SSR) を行う場合, 第一種の過誤確率が增大することが問題となる. 近年, この問題を解決する方法として, 最終解析時の検定統計量を調整する方法 (Cui, Hung and Wang, 1999), 中間解析時の条件付検出力 (conditional power : CP) が 50%以上であれば SSR を行う 50%-CP 法 (Chen, Demets and Lan, 2004) 及び 50%-CP 法の基準となる条件付検出力を 50%より低い値に変更する方法 (Uemura, Matsuyama and Ohashi, 2008) が提案されている. 一方, 被験者数再算出方法としては, 試験開始前に仮定した薬剤効果を中間解析時の観測値に置き換え試験計画時の被験者数算出方法に従って被験者数再算出を行う方法 (事前検出力基準) 及び中間解析時に観測した薬剤効果が真の薬剤効果であると仮定した下で中間解析時の情報を固定条件として, 最終解析を行った場合の検出力 (条件付検出力) に基づいて被験者数再算出を行う方法 (条件付検出力基準) が提案されている. 第一種の過誤確率を制御する方法において, Uemura らにより事前検出力基準を用いて被験者数再算出を行う場合の性能比較が行われた. 本論文では, 中間解析結果に基づいて SSR を行う場合の第一種の過誤確率の制御方法, 中間解析結果に基づく被験者数再算出方法を解説し, 実際の臨床試験を想定したシミュレーションによって, 特に, 事前検出力基準を用いた場合と条件付検出力基準を用いた場合の評価指標の挙動に着目し, 各方法の特性を検討する.

キーワード : 臨床試験, Adaptive Design, 被験者数再設定, 事前検出力, 条件付検出力

1 緒言

真に有用性をもつ新薬を一日でも早く上市するためには, 臨床試験の成功確率を向上することが重要な課題となるが, 近年, 臨床試験の成功確率は著しく低下している (政策研ニュース, 2010). そのような中, 試験の成功確率を向上するための方法として, 中間解析結果に基づき必要に応じて試験デザインを変更する Adaptive Design が注目されており, 研究が活発に行われている.

臨床試験において、被験者数を何例に設定するかは試験の成功確率を左右する重要な問題である。被験者数を誤って少なく設定すると検出力が不足し、臨床的に意味のある薬剤効果を検出できない可能性がある。一方、被験者数を多く設定するとオーバーパワーにより不要な被験者をリスクに曝すことになり倫理的に好ましくないという問題がある。治験薬の有効性の検証を目的とする臨床試験では、過去の臨床試験の結果や他剤の情報などを参考に、検出すべき対照薬との差やその分散を仮定し、被験者数を設定する。しかしながら、これらの情報は不確実な場合が多く、試験計画立案時に適切な被験者数を設定することが困難である場合が多い。中間解析により試験開始後に観測した対照薬との差やその分散が仮定した値と乖離していることが判明した場合、観測した値に基づいて被験者数を再設定することが成功確率の向上に繋がるのが期待される。そこで、試験の成功確率を向上させる方法の一つとして、試験開始後に得られた薬剤効果に基づいて被験者数を再設定する Adaptive Design が提案されている。このような Design を用いることにより、真の薬剤効果が仮定した薬剤効果より小さい場合に、被験者数不足による検出力不足を防止することが可能となる。しかしながら、単純にこのような方法を用いると、第一種の過誤確率が增大することが問題視されている。近年、この問題を解決する方法として、最終解析時に用いる検定統計量を調整する方法 (Cui, Hung and Wang, 1999)、検定統計量の調整を行わずに中間解析時の条件付検出力が 50%以上であれば第一種の過誤確率を名義水準以下に保ちながら SSR を行う Chen らの方法 (Chen, Demets and Lan, 2004) 及び Chen らの方法における条件付検出力の基準を 50%より低い値に変更する方法 (Uemura, Matsuyama and Ohashi, 2008) が提案されている。一方、被験者数の算出方法としては、試験開始前に仮定した対照薬との差及びその分散を中間解析時の観測値に置き換え試験計画時の被験者数算出方法に従って被験者数再算出を行う方法 (Cui, Hung and Wang, 1999) 及び中間解析時に観測した値が真の薬剤効果であると仮定した下で中間解析までの情報を固定条件として、最終解析を行った場合の検出力 (条件付検出力) に基づいて被験者数再算出を行う方法 (Proschan and Hunsberger, 1995 ; Wassmer, Eisebitt and Coburger 2001) が提案されている。このように、SSR における被験者数再算出方法及び第一種の過誤確率を名義水準以下に制御する方法は数多く提案されている。実際に SSR を臨床試験に適応する際には、可能な限り高性能な方法を採用したい。そのため、臨床試験に適応する前に、各手法の性能を評価し、特徴を把握することが重要である。これまでに第一種の過誤確率を名義水準以下に制御する各方法において、Uemura らにより、Cui らが提案した方法を用いて被験者数再算出を行う場合の性能比較が行われた (Uemura, Matsuyama and Ohashi, 2008)。そこで本稿では、第一種の過誤確率を名義水準以下に制御する 3 つの方法について、Cui らと Proschan らによる 2 つの被験者数再算出方法の特性を検討した。

本論文では、2 章で第一種の過誤確率を名義水準以下に制御する方法、3 章で中間解析結果に基づく被験者数再算出方法を紹介し、4 章で性能比較を行うための方法、5 章で結果を紹介する。最後に、6 章で本論文の要約を行う。

2 第一種の過誤確率を制御する方法

本論文では、治験薬の対照薬に対する有効性を検証することを目的とした、Key Open を伴う中間解析を 1 回行う二重盲検無作為化並行群間比較試験を想定し、帰無仮説 $H_0: \delta = 0$, 対立仮説 $H_1: \delta > 0$ について第一種の過誤確率を片側 $\alpha = 2.5\%$ と設定した下で片側検定を行うことを考える。ここで本論文中で使用する変数を以下のように定義する。

- i : 薬剤群 (治験薬群 : $i=1$, 対照薬群 : $i=2$).
- j : 被験者番号 ($j=1, \dots, N^*$).
- μ_i : 薬剤群 i の反応の平均.
- σ^2 : 各群の反応の群内分散 (共通で既知と仮定). (一般性を失くことなく, $\sigma^2 = 1$ と仮定)
- X_{ij} : 薬剤群 i に割り付けられた被験者 j の反応変数. $X_{ij} \sim N(\mu_i, 1)$.
- δ : 治験薬の対照薬に対する効果.

$$\delta = \frac{\mu_1 - \mu_2}{\sigma} = \mu_1 - \mu_2 (= \delta_0) \quad (\delta_0 : \text{試験計画時に仮定する薬剤効果}).$$

- $\hat{\delta}$: 中間解析時の薬剤効果の推定値.
- z_u : 標準正規分布の上側 u % 点.
- $1 - \beta$: 試験計画時の検出力.
- N_0 : 試験計画時の各群の被験者数.

$$N_0 = 2 \left(\frac{z_\alpha + z_\beta}{\delta_0 / \sigma} \right)^2 = 2 \left(\frac{z_\alpha + z_\beta}{\delta_0} \right)^2 \quad (\because \sigma^2 = 1).$$

- M : 中間解析結果に基づき算出した被験者数再算出後の各群の被験者数.
- N^* : SSR 後の最終的な各群の被験者数.
- n_1 : 中間解析前に集積される各群の被験者数.
- n_2 : SSR を行わなかった場合に, 中間解析後に集積する各群の被験者数 ($n_2 = N_0 - n_1$).
- n_2^* : SSR を行った場合に, 中間解析後に集積する各群の被験者数 ($n_2^* = N^* - n_1$).
- t : 中間解析を行う時点. 各群において n_1 名の被験者データが集積された時点で中間解析を行う ($t = n_1 / N_0$).

2.1 Cui らの重み付き検定統計量を用いる方法

中間解析結果に基づき SSR を行う前及び行った後に集積された被験者データから得られる検定統計量をそれぞれ Z_{n_1} 及び $Z_{N^*-n_1}$, SSR 後の各群の最終的な被験者数 N^* に基づく最終解析時の検定統計量 (中間解析前後の各被験者データに対する重みは等しい) を Z_{fix}^* とすると, それぞれ以下の式で与えられる.

$$\begin{aligned} Z_{n_1} &= \frac{1}{\sqrt{2n_1}} \sum_{j=1}^{n_1} (X_{1j} - X_{2j}), \quad Z_{N^*-n_1} = \frac{1}{\sqrt{2(N^* - n_1)}} \sum_{j=n_1+1}^{N^*} (X_{1j} - X_{2j}) \\ Z_{fix}^* &= \frac{1}{\sqrt{2N^*}} \sum_{j=1}^{N^*} (X_{1j} - X_{2j}) \\ &= \frac{1}{\sqrt{2N^*}} \sum_{j=1}^{n_1} (X_{1j} - X_{2j}) + \frac{1}{\sqrt{2N^*}} \sum_{j=n_1+1}^{N^*} (X_{1j} - X_{2j}) \\ &= \sqrt{\frac{n_1}{N^*}} Z_{n_1} + \sqrt{\frac{N^* - n_1}{N^*}} Z_{N^*-n_1} \end{aligned}$$

従って, 検定統計量 Z_{fix}^* は Z_{n_1} と $Z_{N^*-n_1}$ の独立した検定統計量の重み付き和で表わされる. しかし, 中間解析前後の各被験者データに対する重み (検定統計量の重み) が中間解析結果に依存して決まる SSR 後の被験者数 N^* に依存しているため, Z_{fix}^* は標準正規分布に従わず, 単純に Z_{fix}^* を用いて最終解析を行うと第一種

の過誤確率が増大する．そこで，中間解析前後の各被験者データに対する重み（検定統計量の重み）を中間解析結果に依存しない重みに変更することで，第一種の過誤確率を制御する方法が提案された（Cui, Hung and Wang, 1999）．Cui らが提案した最終解析時に用いる検定統計量を Z_{CHW}^* とすると， Z_{CHW}^* は以下の式で表わされる．

$$\begin{aligned} Z_{CHW}^* &= \sqrt{\frac{n_1}{N_0}} Z_{n_1} + \sqrt{\frac{N_0 - n_1}{N_0}} Z_{N^* - n_1} \\ &= \sqrt{t} Z_{n_1} + \sqrt{1-t} Z_{N^* - n_1} \end{aligned}$$

中間解析前後の各被験者データに対する重み（検定統計量 Z_{n_1} と $Z_{N^* - n_1}$ に対する重み）はそれぞれ \sqrt{t} 及び $\sqrt{1-t}$ であり，共に中間解析結果に依存しない重みであるため， Z_{CHW}^* を用いて最終解析を行った場合，第一種の過誤確率は名義水準以下に制御される．Cui らの提案した方法は，中間解析後に被験者数を増やした場合，中間解析後の各被験者データに対する重みが SSR 後の被験者数ではなく，試験計画時の被験者数で重み付けられるため，中間解析前の各被験者データに対して，中間解析後の各被験者データに対する重みを少なくした方法である．検定統計量 Z_{CHW}^* は被験者数を変更するか否かに関わらず，帰無仮説の下で標準正規分布に従うので対応する棄却限界値は，中間解析時に有効・無効中止を考えず SSR のみを行う場合，固定デザインにおける棄却限界値 z_α となる．

なお，Cui らの方法を用いて第一種の過誤確率の制御を行った場合，SSR 後の被験者数 N^* を以下のように決定する．ここで， M は中間解析結果に基づいて推定した被験者数， N_{\max} は想定する臨床試験における SSR 後の被験者数の最大値であり，臨床的に意味のある最小の対照薬との差を試験計画時の検出力で検出する際に必要となる被験者数である．

$$N^* = \begin{cases} N_0 & \text{if } \hat{\delta} \geq \delta_0 \\ M & \text{if } \hat{\delta} < \delta_0 \text{ and } N_0 < M < N_{\max} \\ N_{\max} & \text{if } \hat{\delta} < \delta_0 \text{ and } N_{\max} < M \end{cases} \quad (1)$$

2.2 重み付けない検定統計量を用いる Chen らの方法（50%-CP 法）

中間解析結果に基づく条件付検出力 $CP(t, z, \delta = \hat{\delta})$ を以下のように定義する．ただし， $t = n_1/N_0$ である．

$$CP(t, z, \delta = \hat{\delta}) = \Pr(Z_{N_0} > z_\alpha \mid Z_{n_1} = z, t; \delta = \hat{\delta}) = \Phi\left(\frac{z/\sqrt{t} - z_\alpha}{\sqrt{1-t}}\right)$$

（ Φ は標準正規分布の累積分布関数を， Z_{N_0} は試験計画時の被験者数で最終解析を行った場合の検定統計量を示し， $Z_{N_0} = (2N_0)^{-1/2} \sum_{j=1}^{N_0} (X_{1j} - X_{2j})$ で与えられる）

上記の条件付検出力は，中間解析時に得られた薬剤効果が真の薬剤効果であると仮定した下で中間解析時を固定条件として，最終解析を試験計画時の被験者数で行った場合の検出力であり，この条件付検出力が 50% 以上であれば第一種の過誤確率を制御しながら SSR を行うことが可能であることが示されている（Chen, Demets and Lan, 2004）．本論文では，Chen らが提案した方法を 50%-CP 法と呼ぶことにする．50%-CP 法を用いて SSR を行った場合に最終解析に用いる検定統計量は，固定デザインで用いる検定統計量と同様である．

なお、50%-CP法を用いた場合は、SSR後の最終的な被験者数 N^* を以下のように決定する。

$$N^* = \begin{cases} N_0 & \text{if } CP(t, z, \delta = \hat{\delta}) < 0.5 \\ N_0 & \text{if } CP(t, z, \delta = \hat{\delta}) \geq 0.5 \text{ and } M \leq N_0 \\ M & \text{if } CP(t, z, \delta = \hat{\delta}) \geq 0.5 \text{ and } N_0 < M < N_{\max} \\ N_{\max} & \text{if } CP(t, z, \delta = \hat{\delta}) \geq 0.5 \text{ and } N_{\max} < M \end{cases} \quad (2)$$

2.3 重み付けない検定統計量を用いる Uemura らの方法 (20%-CP法)

Chen らが提案した 50%-CP法は、第一種の過誤確率を名義水準以下に制御し過ぎており、保守的な手法であることが Chen らのシミュレーションにより示されている。そこで、制御し過ぎた第一種の過誤確率の残りを利用し、50%-CP法の条件付検出力の閾値を 20%まで低下させる方法が提案された (Uemura, Matsuyama and Ohashi, 2008)。本論文では、Uemura らの方法を 20%-CP法と呼ぶことにする。Uemura らは、被験者数再算出を行った結果、被験者数を試験計画時の被験者数より増やす場合には、各群の被験者数が試験全体で $r_{\min} \times N_0$ 以上になるように設定するという条件を設けることで、50%-CP法の条件付検出力の閾値を 20%まで低下することを可能とした。従って、20%-CP法を用いた場合は、SSR後の最終的な被験者数 N^* は以下のように決定する。

$$N^* = \begin{cases} N_0 & \text{if } CP(t, z, \delta = \hat{\delta}) < 0.2 \\ N_0 & \text{if } CP(t, z, \delta = \hat{\delta}) \geq 0.2 \text{ and } M \leq N_0 \\ r_{\min} N_0 & \text{if } CP(t, z, \delta = \hat{\delta}) \geq 0.2 \text{ and } N_0 < M \leq r_{\min} N_0 \\ M & \text{if } CP(t, z, \delta = \hat{\delta}) \geq 0.2 \text{ and } r_{\min} N_0 < M < N_{\max} \\ N_{\max} & \text{if } CP(t, z, \delta = \hat{\delta}) \geq 0.2 \text{ and } M \geq N_{\max} \end{cases} \quad (3)$$

ただし、 r_{\min} は第一種の過誤確率を制御するために設定する条件であり (Uemura, Matsuyama and Ohashi, 2008)、 r_{\min} は被験者数再算出後の被験者数 M が中間解析結果により一意に推定される場合であれば、中間解析後の被験者数の再算出方法に依らず決まる値である。なお、20%-CP法を用いて第一種の過誤確率を制御しながら SSRを行った場合、最終解析に用いる検定統計量は、50%-CP法同様、固定デザインで用いる検定統計量となる。

3 被験者数再算出方法

3.1 事前検出力基準

中間解析結果に基づき被験者数再算出を行う方法として、試験開始前に仮定した薬剤効果 δ_0 を中間解析時に観測した薬剤効果の推定値 $\hat{\delta}$ に置き換え試験計画時の被験者数再算出方法に従って被験者数再算出を行う

方法が提案されている (Cui, Hung and Wang, 1999). 本論文では, 本方法を事前検出力基準と呼ぶことにする. 本基準を用いて推定される各群の再算出後の被験者数 M は, 第一種の過誤確率の制御方法に依らず, 以下の式で与えられる.

$$M = 2 \left(\frac{z_\alpha + z_\beta}{\hat{\delta}} \right)^2 = \left(\frac{\delta_0}{\hat{\delta}} \right)^2 N_0$$

3.2 条件付検出力基準

中間解析結果に基づき被験者数再算出を行う別の方法として, 中間解析時に観測した薬剤効果が真の薬剤効果であると仮定した下で, 中間解析までの情報を固定条件として, 最終解析を行った場合の検出力 (条件付検出力) に基づいて被験者数再算出を行う方法が提案されている (Proschan and Hunsberger, 1995; Wassmer, Eisebit and Coburger, 2001). 本論文では, 本方法を条件付検出力基準と呼ぶことにする. 本基準を用いて推定される再算出後の被験者数 M は, 第一種の過誤確率の制御方法により異なる.

3.2.1 重み付き検定統計量を用いる場合

Cui らの重み付き検定統計量を用いて第一種の過誤確率の制御を行う場合, 中間解析時に観測した薬剤効果が真の薬剤効果であると仮定した下で中間解析時の情報を固定条件として, SSR 後の被験者数で最終解析を行った場合の検出力 (条件付検出力) $CP_{CHW}(t, z, n_2^*, \delta = \hat{\delta})$ は以下の式で与えられる.

$$CP_{CHW}(t, z, n_2^*, \delta = \hat{\delta}) = \Pr(Z_{CHW}^* > z_\alpha \mid Z_{n_1} = z, t, n_2^*; \delta = \hat{\delta}) = \Phi \left(\hat{\delta} \sqrt{\frac{n_2^*}{2}} - \frac{z_\alpha - z\sqrt{t}}{\sqrt{1-t}} \right)$$

この条件付検出力が試験計画時に設定した検出力 $1 - \beta$ に一致するような中間解析後に集積する被験者数 n_2^* は,

$$CP_{CHW}(t, z, n_2^*, \delta = \hat{\delta}) = 1 - \beta$$

を n_2^* について解き,

$$n_2^* = \frac{2}{\hat{\delta}^2} \left(\frac{z_\alpha - z\sqrt{t}}{\sqrt{1-t}} + z_\beta \right)^2$$

で与えられる. 従って, Cui らの重み付き検定統計量による方法を用いて第一種の過誤確率の制御を行う場合の被験者数再算出後の各群の被験者数 M は, $M = n_1 + n_2^*$ で推定される.

3.2.2 50%-CP 法又は 20%-CP 法を用いる場合

Chen らの 50%-CP 法を用いて第一種の過誤確率の制御を行う場合, 中間解析時に観測した薬剤効果が真の薬剤効果であると仮定した下で, SSR 後の被験者数で最終解析を行った場合に得られる条件付検出力 $CP(t, z, M, \delta = \hat{\delta})$ は以下の式で与えられる.

$$CP(t, z, M, \delta = \hat{\delta}) = \Pr(Z_{fix}^* > z_\alpha | Z_{n_1} = z, t, M, \delta = \hat{\delta}) = \Phi\left(\frac{z/\sqrt{n_1/M} - z_\alpha}{\sqrt{1 - n_1/M}}\right)$$

この条件付検出力が試験計画時に設定した検出力 $1 - \beta$ に一致するような被験者数再算出後の被験者数 M は、

$$CP(t, z, M, \delta = \hat{\delta}) = 1 - \beta$$

を M について解くことで推定することができる。上式は M の 4 次方程式であるため、本論文では IML プロシジャの polyroot 関数を用いて、 M を推定した。

なお、20%-CP 法を用いて SSR を行った場合の再算出後の被験者数 M は、50%-CP 法の場合と同様になる。

4 シミュレーション方法

前述の第一種の過誤確率を制御する方法（Cui らの方法、50%-CP 法、20%-CP 法）と被験者数再算出方法（事前検出力基準、条件付検出力基準）の全ての組み合わせ（全 6 組）について、シミュレーションを実施して、これらの手法の挙動及び特性について検討した。なお、シミュレーションには SAS 9.2 を使用した。

想定する臨床試験は、被験薬の対照薬に対する優越性を検討することを目的とした、二重盲検無作為化並行群間比較試験とした。ここで、反応変数は正規分布に従っているとし、SAS 9.2 の RANNOR 関数を用いて乱数を発生させた。シミュレーションの設定条件を表 1 に示す。また、手法の挙動及び性能について評価するための評価指標を表 2 に示す。

表 1 シミュレーションの設定条件 ($\sigma^2 = 1$)

片側有意水準： α	$\alpha = 2.5\%$
検出力： $1 - \beta$	$1 - \beta = 80\%$
被験薬の対照薬に対する真の薬剤効果： δ	0.15, 0.19, 0.23, 0.27, 0.31, 0.35
試験計画時に仮定する薬剤効果： δ_0	$\delta_0 = 0.25$
1 群あたりの初期被験者数： N_0	$N_0 = 252$
被験者数の再算出を行う時点： $t = n_t/N_0$	0.25, 0.5, 0.75
20%-CP 法で被験者数を増やす場合に最低でも増やす被験者数の下限の $r_{min}N_0$ の係数 r_{min}	$t = 0.25$ のとき：1.2 $t = 0.50, 0.75$ のとき：1.1 (Uemura, Matsuyama and Ohashi, 2008 より引用)
SSR 後の被験者数の最大値： N_{max}	698 ($\delta = 0.15$ を臨床的に意味のある差の下限とし、設定した)

表 2 評価指標

評価指標	定義
第一種の過誤確率 (<i>Alpha Error</i>)	$\text{Alpha Error} = \frac{\sum_{s=1}^S A_s}{S}$ <p>S はシミュレーション回数, A_s は s 回目のシミュレーションで帰無仮説: $\delta = 0$ の下で統計的に有意差が観察された場合は 1 を, 観察されなかった場合は 0 をとる指示変数とする.</p>
SSR を行った場合の検出力 (<i>Power (N*)</i>)	$\text{Power}(N^*) = \frac{\sum_{s=1}^S R_s}{S}$ <p>S はシミュレーション回数, R_s は s 回目のシミュレーションで対立仮説の下で統計的に有意差が観察された場合は 1 を, 観察されなかった場合は 0 をとる指示変数とする.</p>
被験者数を増やした割合	$\text{Pr}(N^* > N_0) = \frac{\sum_{s=1}^S U_s}{S}$ <p>S はシミュレーション回数, U_s は s 回目のシミュレーションで SSR 後の被験者数 N^* が初期被験者数 N_0 を超えた場合は 1 を, 達しなかった場合は 0 をとる指示変数とする.</p>
期待被験者数 (<i>ASN</i>)	$\text{ASN} = \frac{\sum_{s=1}^S N_s^*}{S}$ <p>S はシミュレーション回数, N_s^* は s 回目のシミュレーションにおける最終被験者数とする.</p>
最大被験者数 (N_{\max}) に達した割合	$\text{Pr}(N^* = N_{\max}) = \frac{\sum_{s=1}^S Q_s}{S}$ <p>S はシミュレーション回数, Q_s は s 回目のシミュレーションで最大被験者数 N_{\max} に達した場合は 1 を, 達しなかった場合は 0 をとる指示変数とする.</p>
期待被験者数 100 例あたりの検出力 (<i>Power (ASN₁₀₀)</i>)	$\text{Power}(\text{ASN}_{100}) = \frac{\text{Power}(N^*)}{\text{ASN}} \times 100$
固定デザインの下で SSR を行った場合の検出力を得るのに必要な被験者数 (N_0^*)	$N_0^* = \frac{2(z_{\alpha} + z_{\beta'})^2}{\delta^2}$ <p>$1 - \beta'$ は SSR を行った場合の検出力を示し, δ は真の薬剤効果を示す.</p>
被験者数再算出方法の効率 (<i>Efficiency</i>)	$\text{Efficiency} = \frac{\text{ASN}}{N_0^*}$

第一種の過誤確率を制御する方法（Cui らの方法，50%-CP 法，20%-CP 法）と被験者数再算出方法（事前検出力基準，条件付検出力基準）の各組み合わせに対して，以下の手順でシミュレーションを実施した。

- 1) 試験計画時に仮定する薬剤効果 δ_0 ，及び 1 群あたりの初期被験者数 N_0 に対して，真の薬剤効果 δ ，及び被験者数の再算出を行う時点 t を与え，被験薬群 $n_1 = N_0 \times t$ 名のデータを薬剤効果 δ ，分散 $\sigma^2 = 1$ に従う正規乱数より発生させる。
- 2) 対照薬群 $n_1 = N_0 \times t$ 名のデータを平均 0，分散 $\sigma^2 = 1$ に従う正規乱数より発生させる。
- 3) 上記の 1) 及び 2) で発生させた計 $2n_1$ 名のデータを用いて，被験者数の再算出時の薬剤効果の推定値 $\hat{\delta}$ を算出し，被験者数再算出方法に基づき 1 群あたりの被験者数 M を算出する。
- 4) 第一種の過誤確率を制御する方法に基づき，SSR 後の最終的な 1 群あたりの被験者数 N^* を「2 第一種の過誤確率を制御する方法の (1) ~ (3) 式」に示すルールに従い算出する。
- 5) 被験薬 $n_2^*(= N^* - n_1)$ 名のデータを薬剤効果 δ ，分散 $\sigma^2 = 1$ に従う正規乱数より発生させる。
- 6) 対照薬 $n_2^*(= N^* - n_1)$ 名のデータを平均 0，分散 $\sigma^2 = 1$ に従う正規乱数より発生させる。
- 7) 1 群あたり N^* 名のデータに対して検定統計量^{a)} を算出し，有意水準片側 2.5% で両側検定を行い，有意判定を行う。
 - a) Cui らの方法を用いて第一種の過誤確率の制御を行う場合，中間解析前及び後の被験者データから得られた検定統計量を用いて，重み付き検定統計量を算出する。
- 8) 上記の 1) から 7) を 10,000 回繰り返す。
- 9) シミュレーション結果に対して，それぞれの評価指標を算出する。

5 シミュレーション結果と考察

表 3 に Cui らの方法，20%-CP 法及び 50%-CP 法を用いた第一種の過誤確率の制御について，事前検出力基準及び条件付検出力基準を用いて被験者数再算出を行った場合の第一種の過誤確率をシミュレーションにより評価した結果を示した。

表 3 第一種の過誤確率

第一種の過誤確率の制御方法	被験者数再算出方法	第一種の過誤確率
Cui らの方法	事前検出力基準	2.18
	条件付検出力基準	2.14
20%-CP法	事前検出力基準	2.36
	条件付検出力基準	2.36
50%-CP法	事前検出力基準	2.27
	条件付検出力基準	2.29

表 3 より，いずれの場合においても，第一種の過誤確率は事前に設定した有意水準片側 2.5% 未満であり，第一種の過誤確率を制御していることが確認された。

表 4 に，「4 シミュレーション方法」で示したシミュレーション方法に従い，「表 2 評価指標」で示した評価指標を算出した結果を示す。表中の総合評価は被験者数を増やさなかった場合と増やした場合の両方を含めた場合の結果であり，条件付評価は被験者数を増やした場合のみの結果である。なお，中間解析時点に依らず同様の結果が得られたため，ここでは $t = 0.50$ の場合の結果のみを示した。また，Cui らの方法，20%-CP

法及び 50%-CP 法を用いて第一種の過誤確率の制御を行い、SSR を行った場合の真の薬剤効果 δ に対する検出力の推移をそれぞれ図 1-1～図 1-3 に、期待被験者数の推移を図 2-1～図 2-3 に、期待被験者数 100 例あたりの検出力を図 3-1～図 3-3 に示した。

最初に、被験者数再算出方法である事前検出力基準と条件付検出力基準の性能を比較する。臨床試験においては、真の薬剤効果が仮定した薬剤効果よりも小さい場合に SSR を行い、検出力の向上を図りたいと考えるため、ここでは真の薬剤効果が仮定した薬剤効果 ($\delta_0 = 0.25$) より小さい場合 ($\delta = 0.15, 0.19, 0.23$) の結果に注目する。この場合、第一種の過誤確率の制御方法に依らず、総合評価において検出力 $Power(N^*)$ は事前検出力基準が条件付検出力基準と比較して若干高かった (図 1-1～1-3)。これは、事前検出力基準の方が条件付検出力基準と比較して被験者数を増やした割合 $Pr(N^* > N_0)$ 及び最大被験者数に達した割合 $Pr(N^* = N_{max})$ が高く、期待被験者数 ASN が多かったためであると考えられる (図 2-1～2-3)。なお、期待被験者数 100 例あたりの検出力 $Power(ASN_{100})$ 及び SSR による固定デザインに対する効率 $Efficiency$ は、両再算出方法においてほぼ同等であった (図 3-1～図 3-3)。同様に、条件付評価においても、事前検出力基準が条件付検出力基準と比較して期待被験者数 ASN が多く検出力 $Power(N^*)$ が高かった (図 1-1～1-3, 2-1～2-3)。なお、期待被験者数 100 例あたりの検出力 $Power(ASN_{100})$ 及び SSR による固定デザインに対する効率 $Efficiency$ は両再算出方法においてほぼ同等であった (図 3-1～図 3-3)。以上の結果より、第一種の過誤確率の制御方法に依らず、事前検出力基準は条件付検出力基準と比較して期待被験者数 100 例あたりの検出力及び固定デザインに対する効率はほぼ同程度であるが、期待被験者数が多いため検出力が高い方法であると考えられる。

次に、真の薬剤効果が仮定した薬剤効果よりも小さい場合における第一種の過誤確率の制御方法である Cui らの方法、20%-CP 法及び 50%-CP 法の性能を比較する。この場合、被験者数再算出方法に依らず、総合評価における検出力 $Power(N^*)$ は、高い順に Cui らの方法、20%-CP 法、50%-CP 法であった (図 1-1～1-3)。これは、Cui らの方法、20%-CP 法、50%-CP 法の順で被験者数を増やした割合 $Pr(N^* > N_0)$ 及び最大被験者数に達した割合 $Pr(N^* = N_{max})$ が高く、期待被験者数 ASN が多かったためであると考えられる (図 2-1～2-3)。期待被験者数 100 例あたりの検出力 $Power(ASN_{100})$ は、高い順に 50%-CP 法、20%-CP 法、Cui らの方法であった (図 3-1～3-3)。なお、SSR による固定デザインに対する効率 $Efficiency$ は、いずれの方法においても 1 に近い値であり、被験者効率は固定デザインとほぼ同等であった。一方、条件付評価においては、期待被験者数 ASN は総合評価と同様、多い順に Cui らの方法、20%-CP 法、50%-CP 法であったが、検出力 $Power(N^*)$ は 50%-CP 法が最も低く、場合により Cui らの方法と 20%-CP 法の検出力の大小関係は異なっていた (図 1-1～1-3, 2-1～2-3)。期待被験者数 100 例あたりの検出力 $Power(ASN_{100})$ は総合評価と同様、高い順に 50%-CP 法、20%-CP 法、Cui らの方法であった (図 3-1～3-3)。SSR による固定デザインに対する効率 $Efficiency$ について、Cui らの方法は、真の薬剤効果に依らず 1 より高い値であり、被験者効率は固定デザインと比較して良くなかった。20%-CP 法及び 50%-CP 法については、真の薬剤効果が仮定した薬剤効果よりも小さい場合では、1 より小さく、被験者効率は固定デザインと比較して良かった。

本シミュレーションでは、Cui らの方法について、中間解析時の治験薬群の薬剤効果が対照薬群と比較して少しでも大きければ SSR を行うルール (「2 第一種の過誤確率を制御する方法」の (1) 式) を設けたことから、20%-CP 法及び 50%-CP 法と比較して、特に真の薬剤効果が想定よりも著しく低かった場合に SSR を行う割合が高いため期待被験者数が多く、それに付随して検出力も高くなっていると考えられる。実際の臨床試験においては、被験者数に比例して臨床試験に要する時間及び費用が上昇するため、本シミュレーションにおける SSR を行うルールに従い第一種の過誤確率の制御方法として Cui らの方法を用いる場合は、費用対

効果をより慎重に見極める必要があると考えられる。一方、20%-CP法は、Cuiらの方法と比較して検出力は低いものの、50%-CP法と比較して検出力が向上している。また、被験者数はCuiらの方法の被験者数と試験計画時の初期被験者数の中間程度であり、検出力の向上と費用のバランスの取れた方法であると考えられる。なお、50%-CP法は、検出力が固定デザインの検出力と大きく変わらず、被験者数も試験計画時の初期被験者数と大きく変わらないことから、SSRによる効果が得られにくい方法であると考えられる。なお、これらの結論はUemuraらが行ったシミュレーション結果に対する考察と同様であった。

6 まとめ

本稿では、中間解析結果に基づきSSRを行う方法の性能を比較するため、SSRを行う場合に問題となる第一種の過誤確率を制御する方法（Cuiらの方法、50%-CP法、20%-CP法）と被験者数再算出方法（事前検出力基準、条件付検出力基準）の各組み合わせに対して、シミュレーションを行い、特に、事前検出力基準を用いた場合と条件付検出力基準を用いた場合の挙動に着目し、特性の検討を行った。その結果、状況に依らず事前検出力基準は条件付検出力基準と比較して検出力が高い方法であると考えられた。第一種の過誤確率の制御方法については、費用対効果のバランスを考慮して使い分けることが望ましく、費用よりも検出力の向上に重点を置く場合はCuiらの方法を用い、検出力の向上と費用の両方に重点を置く場合は20%-CP法を用いることが望ましいと考える。

謝辞

本論文の作成にあたっては、独立行政法人医薬品医療機器総合機構の上村鋼平氏から貴重な示唆をいただいた。ご協力に深く感謝申し上げます。

参考文献

- 1) 八木 崇, 大久保 昌美, 小野 俊介 (2010) . 医薬品開発の期間と費用—アンケートによる実態調査. 政策研ニュース **29**,1-9.
- 2) Chen, Y. H. J., Demets, D. L., and Lan, K. K. G. (2004) . Increasing sample size when the unblinded interim result is promising. *Statistics in Medicine* **23**, 1023–1038.
- 3) Chow, S.-C., and Chang, M. (2007) . *Adaptive Design Methods in Clinical Trials*. Chapman & Hall/CRC.
- 4) Cui, L., Hung, H. M. J., and Wang, S. J. (1999) . Modification of sample size in group sequential clinical trials. *Biometrics* **55**, 853–857.
- 5) K. Uemura, Y. Matsuyama and Y. Ohashi (2008) . A Modification of the 50%-Conditional Power Approach for Increasing the Sample Size Based on an Interim Estimate of Treatment Difference. *Japanese Journal of Biometrics* **29**, 1, 19-34.
- 6) Chang, M. (2008) . *Adaptive Design Theory and Implementation Using SAS and R*. Chapman & Hall/CRC.
- 7) Proschan, M., and Hunsberger, S. (1995) . Designed extension of studies based on conditional power. *Biometrics* **51**, 1315-1324.
- 8) Wassmer, G., Eisebitt, R., and Coburger, S. (2001) . Flexible interim analyses in clinical trials using multistage adaptive test designs. *Drug Information Journal* **35**, 1131-1146.

表4 シミュレーション結果 ($\delta_0 = 0.25, N_0 = 252, t = 0.50, r_{\min} = 1.1, N_{\max} = 698$)

δ	第一種の過誤確率の制御方法	被験者数再算方法	総合評価								条件付評価					
			Power (N_0)	Power (N^*)	Pr ($N^* > N_0$)	Pr ($N^* = N_{\max}$)	ASN	Power (ASN ₁₀₀)	N_0^*	Efficiency	Power (N_0)	Power (N^*)	ASN	Power (ASN ₁₀₀)	N_0^*	Efficiency
0.15	Cuiらの方法	事前検出力基準	39.0	63.5	67.6	39.3	477.7	13.3	474	1.01	30.9	68.4	585.9	11.7	530	1.11
		条件付検出力基準	39.0	60.3	67.6	26.8	443.5	13.6	440	1.01	30.9	63.7	535.4	11.9	476	1.12
	20%-CP	事前検出力基準	39.0	48.9	37.1	8.8	341.8	14.3	333	1.03	43.4	72.0	494.0	14.6	576	0.86
		条件付検出力基準	39.0	47.4	31.8	7.6	330.3	14.4	321	1.03	41.1	69.8	498.3	14.0	548	0.91
	50%-CP	事前検出力基準	39.0	41.6	20.7	—	276.1	15.1	273	1.01	50.7	66.4	368.5	18.0	506	0.73
		条件付検出力基準	39.0	40.4	15.4	—	267.8	15.1	264	1.01	48.4	61.7	354.6	17.4	455	0.78
0.19	Cuiらの方法	事前検出力基準	56.7	82.3	61.6	30.8	443.7	18.6	463	0.96	46.2	85.9	563.2	15.2	512	1.10
		条件付検出力基準	56.7	80.3	61.6	19.5	409.4	19.6	440	0.93	46.2	82.6	507.5	16.3	467	1.09
	20%-CP	事前検出力基準	56.7	68.6	39.1	8.3	343.5	20.0	333	1.03	58.4	85.8	485.8	17.7	511	0.95
		条件付検出力基準	56.7	67.6	33.2	6.9	330.9	20.4	325	1.02	56.3	85.5	489.3	17.5	506	0.97
	50%-CP	事前検出力基準	56.7	61.3	22.5	—	277.3	22.1	281	0.99	65.1	80.2	364.4	22.0	439	0.83
		条件付検出力基準	56.7	60.4	16.6	—	268.3	22.5	275	0.98	63.2	78.0	349.9	22.3	415	0.84
0.23	Cuiらの方法	事前検出力基準	73.1	91.3	53.4	22.8	403.5	22.6	418	0.97	59.9	93.7	536.0	17.5	462	1.16
		条件付検出力基準	73.1	90.1	53.4	13.8	374.9	24.0	400	0.94	59.9	91.3	482.3	18.9	418	1.15
	20%-CP	事前検出力基準	73.1	81.9	37.4	6.8	332.3	24.7	313	1.06	69.8	92.6	466.8	19.8	440	1.06
		条件付検出力基準	73.1	80.9	30.9	5.6	320.0	25.3	305	1.05	67.1	91.6	472.1	19.4	423	1.12
	50%-CP	事前検出力基準	73.1	76.4	23.4	—	277.2	27.6	273	1.02	76.4	89.1	359.7	24.8	387	0.93
		条件付検出力基準	73.1	75.5	16.9	—	267.8	28.2	267	1.00	73.9	86.5	345.9	25.0	356	0.97
0.27	Cuiらの方法	事前検出力基準	85.7	96.1	42.8	15.6	362.9	26.5	382	0.95	73.0	97.9	511.3	19.1	439	1.16
		条件付検出力基準	85.7	95.0	36.1	14.6	347.4	27.3	358	0.97	70.0	97.0	520.1	18.7	406	1.28
	20%-CP	事前検出力基準	85.7	90.9	32.4	5.2	316.8	28.7	299	1.06	80.3	97.0	451.9	21.5	406	1.11
		条件付検出力基準	85.7	90.3	25.8	4.3	306.0	29.5	293	1.04	78.1	96.7	461.3	21.0	397	1.16
	50%-CP	事前検出力基準	85.7	87.8	21.6	—	274.2	32.0	269	1.02	84.9	95.8	354.7	27.0	375	0.95
		条件付検出力基準	85.7	87.2	15.0	—	265.7	32.8	264	1.01	83.0	94.4	343.3	27.5	347	0.99
0.31	Cuiらの方法	事前検出力基準	93.5	98.4	30.3	8.9	321.9	30.6	352	0.91	83.7	99.2	482.9	20.5	399	1.21
		条件付検出力基準	93.5	97.8	24.1	8.4	309.9	31.6	330	0.94	80.6	98.7	497.8	19.8	366	1.36
	20%-CP	事前検出力基準	93.5	96.3	24.7	3.4	297.2	32.4	294	1.01	88.7	98.9	435.0	22.7	377	1.15
		条件付検出力基準	93.5	96.0	18.6	2.9	289.0	33.2	288	1.00	86.3	98.6	450.5	21.9	361	1.25
	50%-CP	事前検出力基準	93.5	94.9	17.6	—	269.2	35.3	270	1.00	91.9	98.6	349.6	28.2	361	0.97
		条件付検出力基準	93.5	94.6	11.5	—	262.4	36.1	266	0.99	89.9	97.4	342.1	28.5	319	1.07
0.35	Cuiらの方法	事前検出力基準	97.5	99.4	20.8	5.2	296.7	33.5	328	0.90	89.8	99.4	466.9	21.3	328	1.42
		条件付検出力基準	97.5	99.2	16.2	4.7	288.0	34.4	313	0.92	88.1	99.5	480.1	20.7	337	1.42
	20%-CP	事前検出力基準	97.5	98.5	17.8	2.1	283.3	34.8	280	1.01	93.0	99.3	428.0	23.2	320	1.34
		条件付検出力基準	97.5	98.4	13.2	1.7	277.2	35.5	277	1.00	92.1	99.6	442.9	22.5	349	1.27
	50%-CP	事前検出力基準	97.5	98.0	13.1	—	264.8	37.0	265	1.00	94.6	99.0	349.8	28.3	301	1.16
		条件付検出力基準	97.5	97.8	8.5	—	259.8	37.7	259	1.00	93.9	99.2	343.4	28.9	313	1.10

$\delta_0 = 0.25, N_0 = 252, t = 0.50, r_{\min} = 1.1, N_{\max} = 698$

図1-1 Cuiらの方法を用いた場合の検出力

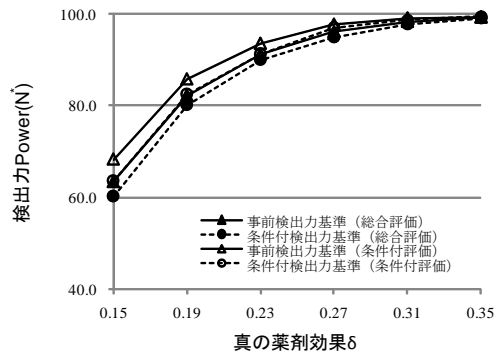


図1-2 20%-CP法を用いた場合の検出力

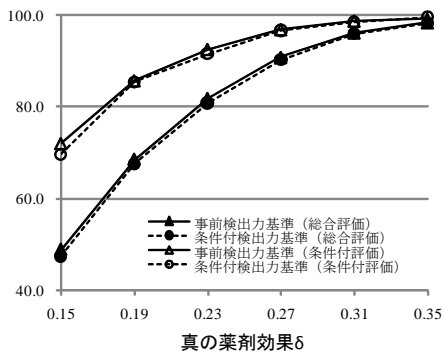


図1-3 50%-CP法を用いた場合の検出力

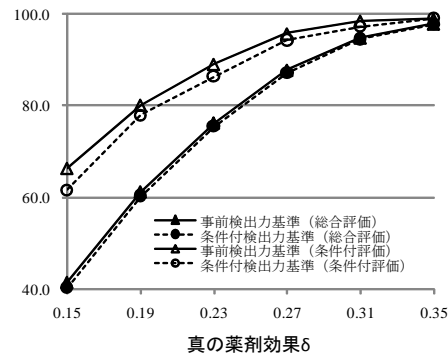


図2-1 Cuiらの方法を用いた場合の期待被験者数

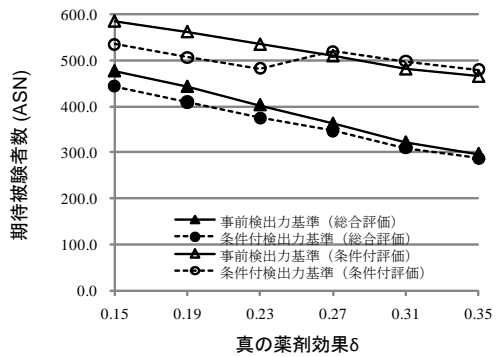


図2-2 20%-CP法を用いた場合の期待被験者数

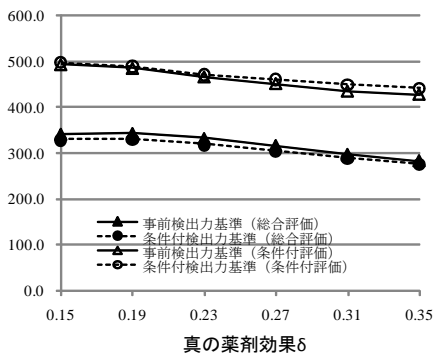


図2-3 50%-CP法を用いた場合の期待被験者数

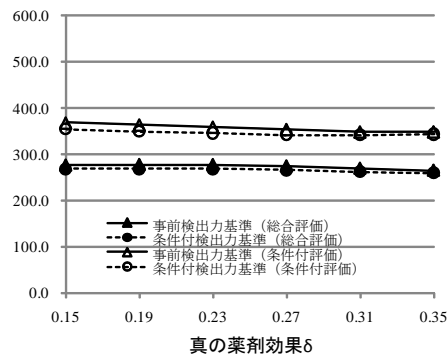


図3-1 Cuiらの方法を用いた場合の期待被験者数100例あたりの検出力

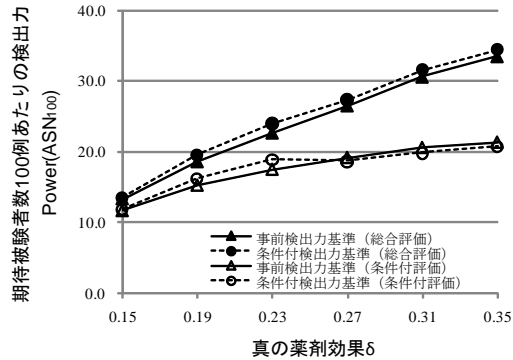


図3-2 20%-CP法を用いた場合の期待被験者数100例あたりの検出力

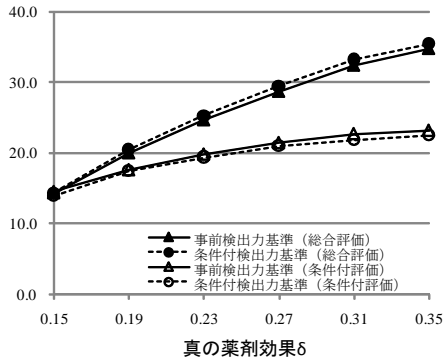


図3-3 50%-CP法を用いた場合の期待被験者数100例あたりの検出力

