

国際共同試験での東アジア地域の症例数調査 (民族差と統計学的手法)

研究分担者 竹内 正弘 北里大学薬学部 教授
研究協力者 吉田 顕臣 北里大学薬学部 研究員
" 内田 めぐみ 北里大学薬学部 助教

研究要旨：

治療効果の民族差の存在が示唆される国際共同試験において、地域間の治療効果の一貫性を評価することは重要である。地域間で一貫した治療効果を有する場合には、他の地域の情報を利用して推定を行う経験ベイズ法が有用である。我々は経験ベイズ法の利用妥当性を評価することによって一貫性を評価する方法を提案した。モンテカルロシミュレーションによって提案法の性能評価を行った結果、従来法よりも有用な場合があることが示唆された。

A. 研究目的：

近年医薬品開発の国際化に伴い、多国籍、多地域で臨床試験を実施する機会が増加している。こういった国際共同試験において、地域間あるいは国間で一貫した治療効果が得られているか評価し、参加した地域あるいは国で被験薬を使用できるか判断することは重要である。

一貫した治療効果が得られない原因としては、遺伝子多型や人種といった内因性要因および医療設備や社会環境などの外因性要因がある。これらの要因によって生じる地域間の治療効果の違いを可能な限り小さくするため、国際共同試験は同一のプロトコールで実施されている。このように同一プロトコールのもとで既知の要因を管理したとしても、未知の要因によって生じる地域間の治療効果の違いまでは管理することができない。

未知の要因が存在するかどうかを評価するために統計学的手法が提案された。2007年に厚生労働省は「国際共同試験に関する基本的考え方について」のQ&Aにおいて、興味のある地域の治療効果とそれ以外の地域の治療効果の一貫性を評価する2つの方法を例示した。1つ目の方法(厚労省法1)は興味のある地域の治療効果の点推定値を D_1 、全地域における治療効果の点推定値を D_{all} とすると、 $D_1/D_{all} > \pi$ (π は0.5以上の定数)となる場合に治療効果の一貫性があると判断する。2つ目の方法(厚労省法2)は、地域($k=1, \dots, K$)の治療効果の点推定値を D_1, D_2, \dots, D_K とすると、全ての k において $D_k > 0$ または $D_k < 0$ となった場合に治療効果の一貫性があると判断する。これらの手法で一貫性評価を行う試験がこれまでに数多く実施されてきた。しかしながら、これらの方法では治療効果の一貫性をその点推定値のみ

で評価しており、点推定値の精度は考慮されていない。

2010年 Hung ら (Pharm.stat.2010; 9: 173-178) は治療効果の点推定値の精度が一貫性評価に与える影響を調べるために、全地域において対照群と被験薬群で有意な差が検出される状況で、治療効果が反転する確率を理論的に計算した。これによると厚労省法2で4地域の一貫性評価を行った際に、判断を誤る確率が約50%となることが指摘された。点推定値の精度は各地域の症例数にも依存しており、症例数の少ない地域が増加することによって、一貫性評価を誤る確率がさらに高くなると考えられる。

そこで我々は、情報量の少ないつまり症例数の少ない地域の推定精度を補う手法として経験ベイズ法に注目した。経験ベイズ法を用いることで一貫性があると判断しやすくなるが、これはあくまで経験ベイズ法の仮定が妥当である場合に限られる。経験ベイズ法の仮定が妥当でない場合には、一貫性を過大評価してしまう。経験ベイズ法の仮定が妥当であるか評価する指標が Efron (JASA 1996; 91: 538-550) によって提案された。本研究では、Efronの方法を発展させ、治療効果の一貫性を評価する新たな手法を提案する。提案法の性能を評価するために、1. 提案法, 2. 厚労省法1, 3. 厚労省法2, 4. 経験ベイズ法に基づく他の手法をモンテカルロシミュレーションを通して比較することを目的とした。

B. 研究方法：

提案法の概要を説明する。地域 i の真の治療効果を θ_i とする($i=1, \dots, K$)。各地域の治療効果の点推定値 $\hat{\theta}_i$ はそれぞれ正規分布 $\hat{\theta}_i \sim N(\theta_i, \sigma_i^2)$ に従うと仮定する($\sigma_i^2 = \sigma^2 * (1/N_{ic} + 1/N_{it})$)。こ

ここで N_{iT} と N_{iC} はそれぞれ地域 i の被験薬群，対照群の症例数である。ここで未観測母数 θ_i は独立に同一の分布 $N(M, A)$ に従うと仮定する。この仮定が経験ベイズ法を用いる際に必要である。この仮定が妥当であるかどうか評価する指標が Efron (JASA 1996; 91: 538-550) によって提案されている。この指標は 2 つの事前分布と興味のある地域 ($i=1$) の尤度を用いて計算される。1 つ目の事前分布は仮定が妥当である下での事前分布 g_A つまり $N(M, A)$ であり，もう 1 つは仮定が妥当でないときの事前分布 g_B 例えば $N(M, \infty)$ である (本研究では ∞ は 10^4 としている)。経験ベイズ法では興味のある地域の真の治療効果 θ_1 に対する事前分布の未観測母数 M および A を推定する際には，興味のある地域以外の観測値 $\hat{\theta}_i (i=2, \dots, K)$ を用いる。 $\hat{\theta}_i (i=2, \dots, K)$ と $\theta_i (i=2, \dots, K)$ の同時分布を $\theta_i (i=2, \dots, K)$ で積分消去し，周辺密度

$$\prod_{i=2}^K \frac{1}{\sqrt{2\pi(A+\sigma_i^2)}} \exp\left(-\frac{(\hat{\theta}_i - M)^2}{2(A+\sigma_i^2)}\right)$$

を得る。制限付き最尤法 (REML 法) によりこの式を制限付きに最大化する推定値 $\hat{M}, \hat{A}, \hat{\sigma}_i^2$ を得る。この推定値を興味のある地域の真の治療効果 θ_1 の事前分布 $N(M, A)$ および $N(M, \infty)$ に代入し，それぞれ事後分布 $f_A(\theta_1 | data), f_B(\theta_1 | data)$ を計算する。 θ_1 の事前分布は g_A または g_B のいずれかであるため，それぞれを選択する事前確率を $h_A, h_B (=1-h_A)$ とする。これらの事前確率は解析者が指定する。それぞれの事前分布を選択する事後確率は，

$$h_A(\hat{\theta}_1) = h_A \int f_A(\theta_1 | data) d\theta_1 / d(\hat{\theta}_1)$$

$$h_B(\hat{\theta}_1) = h_B \int f_B(\theta_1 | data) d\theta_1 / d(\hat{\theta}_1)$$

となる。ただし，

$$d(\hat{\theta}_1) = h_A \int f_A(\theta_1 | data) d\theta_1 + h_B \int f_B(\theta_1 | data) d\theta_1$$

である。ベイズファクターとも呼ばれる事後確率の比 $BF_{AB}(\hat{\theta}_1) = h_A(\hat{\theta}_1) / h_B(\hat{\theta}_1)$ が経験ベイズ法の仮定の妥当性を評価する指標である。本研究ではより精密に経験ベイズ法の仮定の妥当性を評価するため，

$$BF_{AB} = \min(BF_{AB}(\hat{\theta}_i), i=1, \dots, K)$$

を用いる。

ベイズファクターの解釈は Jeffreys (Theory of Probability. Oxford, 1961) および Kass & Raftery

(JASA, 1995; 90: 773-795) で与えられており，それぞれベイズファクターが 10 以上，3 以上となったときに，分子の状況すなわち経験ベイズ法の仮定が妥当であることに強い根拠があるとしている。本研究では BF_{AB} の基準として 3 および 10 を採用する。

提案法の性能を評価するためにモンテカルロシミュレーションを行い，他の 3 つの方法と比較する。

1) 厚生労働省が例示する 1 つ目の方法の変法

$$\{\hat{\theta}_i / \hat{\theta}_{all} > 1/K, \text{ for all } i\}$$

以下，厚労省法 1 とよぶ。

2) 厚生労働省が例示する 2 つ目の方法

$$\{\hat{\theta}_i > 0, \text{ for all } i\}$$

以下，厚労省法 2 と呼ぶ。

3) Huang ら (Pharm.Stat, 2013; 12: 59-64) が提案している経験ベイズ法に基づいた方法

本研究の提案法と同様に異なる 2 つの事前分布によって計算される 2 つの事後分布 $f_A(\theta_1 | data), f_B(\theta_1 | data)$ と h_A, h_B を用いて混合分布を構成する。

$$f(\theta_1 | data) =$$

$$\frac{h_A f_A(\theta_1 | data) + h_B f_B(\theta_1 | data)}{\int \{h_A f_A(\theta_1 | data) + h_B f_B(\theta_1 | data)\} d\theta_1}$$

M および A は REML 法ではなく，最尤法によって推定される。また，事後分布 $f_B(\theta_1 | data)$ の計算に用いる事前分布 g_B は $N(0, \hat{A})$ とする。この混合分布が以下の式を満たす場合に θ_1 は他の地域と一貫した治療効果を有すると判断する。

$$\int_0^\infty f(\theta_1 | data) d\theta_1 > 0.8$$

以下，Huang1 と呼ぶ。

また，提案法との比較のために事前分布の設定 g_B を提案法と同じ $N(M, \infty)$ として混合分布を構成する Huang2, Huang2 の M と A の推定を REML 法で行う Huang3 も比較対照とした。

地域数，治療効果，症例数分配比を様々に変化させてモンテカルロシミュレーションを実施した。

評価指標には一貫性証明同時確率を用いた。これは全地域を併合した際の Welch 検定で有意な治療効果が認められ，かつ一貫性あり，経験ベイズ法の仮定が妥当となった割合と定義する

C. 研究結果：

3 地域で症例数の少ない地域がない場合には提

案法，厚労省が例示する 2 つ目の方法（厚労省法 2），Huang らの提案する方法（Huang2, 3）がいずれも同程度の性能であり，一貫性評価に用いることが推奨される。しかしながら，症例数割合が不均一になったり，地域数が増加すると厚労省法 2 と Huang2 は性能が低下するため，提案法または Huang3 によって一貫性評価を行うことが推奨される。また，提案法および Huang らの提案する方法を利用する際には，経験ベイズ法の仮定が妥当であるとする事前確率 h_A を解析者が決定する必要がある。Huang2, 3 および地域数が少ない場合の提案法（基準値 3）を利用する際には， h_A による変動はほとんどないため，頑健性が保たれる。それ以外の場合には h_A によってある程度変動するため，慎重に選択を行う必要がある。多くの場合には無情報を表す 0.5 を用いることが適切であると考えられる。

D. 今後の展望

トルテロジンの日本人及び韓国人を対象とした第 III 相比較試験において，12 週間投与における尿失禁回数/週における変化率 (%) の平均を主要評価項目として有効性を評価したところ，日本人と韓国人とで平均値の変化の傾向は同様であったが，その変化量は日本人の方が大きい傾向であった。日本人及び韓国人のデータを統合して解析した結果では，有効性が示されず，本国での承認申請データとしては日本人のデータのみが利用されている。申請者の見解ではプラセボ効果やカプサイシンの摂取量の違いを可能性として挙げているが，その要因については明らかにされていない。

トルテロジンの非結合型活性体の薬物動態パラメータ (C_{max} 及び AUC_{0-24}) の値は「日本人，欧米人及び韓国人において類似していた」との評価であった。前述のように，日本人と韓国人とでの有効性評価が異なった点について，その違いの原因を薬物動態的観点から統計的に検討していきたいと考える。

E. 健康危険情報：

該当項目なし

F. 研究発表：

1. M. Takeuchi, H. Uno, A. Yoshida, M. Takeuchi, F. Takahashi, Statistical Analysis for Multi-National Clinical Trials in Inter-Variations among Regions, FDA/Industry Workshop, 2012 年 9 月, USA
2. 吉田顕臣, 高橋史朗, 竹内正弘, 地域一貫性指標に関する考察, RIMS 共同研究 Asymptotic Expansions for Various Models and Their Related

Topics 研究会, 2013 年 3 月 4 日, 京都

3. 竹内正弘, 統計学的手法と民族差, 第 34 回日本臨床薬理学会学術総会 シンポジウム, 2013 年 12 月 5 日, 東京

G. 知的財産権の出願・登録状況：

該当項目なし

H. 添付資料

該当項目なし