

国際共同治験における外れ値となる地域の検出と影響力診断の方法

青木 誠

総合研究大学院大学 統計科学専攻 博士課程(3年次編入学)4年

背景と目的

最近では医薬品開発のグローバル化のため、国際共同治験が増加してきている。この国際共同治験の主目的は全ての地域における新薬の有効性を検証する事であるが、地域間で様々な関連する因子の違いにより異なる地域で治療効果が一貫しない場合がある。そのため、治療効果の一貫性の評価が国際共同治験における重要な統計的課題となっている。これまで極端な性質を検出し、その影響を診断する有効な方法が調査されているが、国際共同治験での方法は議論されていない。本研究ではこの方法を応用し、治療効果が全体集団と一貫していない事が疑われる地域を検出する方法を提案する。

国際共同治験における変量効果モデルと固定効果モデル

本研究では、 i 番目の地域の部分集団解析から得られる各地域の要約統計量 y_i ($i = 1, \dots, k$)を用いる。 k 個の地域で共通の治療効果を仮定し、 y_i に対して正規分布

$$y_i \sim N(\theta, \sigma_i^2)$$

を仮定する。ここで、 θ は共通の治療効果のパラメータであり、 σ_i^2 は部分集団解析から得られる各地域の分散とする。このモデルはメタアナリシスにおける固定効果モデルとなる。もう一つの標準的な統計モデルとして地域間の異質性を考慮した変量効果モデル

$$y_i \sim N(\theta_i, \sigma_i^2), \theta_i \sim N(\mu, \tau^2) \quad (*)$$

を仮定する。ここで、 μ は k 個の地域の治療効果の全体平均であり、 τ^2 は地域間の分散とする。 $\tau^2 = 0$ の時、モデル(*)が固定効果モデルに対応するため、本研究では主に変量効果モデルに対して提案法を考える。

Studentized residualを用いた統計量

従来の回帰診断で用いられるdfbetaと同様のLeave-one-out cross validation (LOOCV)型の影響尺度について考える(Belsley et al. [1])。 $\hat{\mu}^{(-i)}$ と $\tau^{2(-i)}$ は i 番目の地域を除いた $k - 1$ 個の地域のデータに基づいた変量効果モデル(*)からのREML推定量とする。この時、LOOCV型のStudentized residualは以下で定義される。

$$t_i = \frac{y_i - \hat{\mu}^{(-i)}}{\sqrt{\text{Var}[y_i - \hat{\mu}^{(-i)}]}}$$

ここで、 $\text{Var}[y_i - \hat{\mu}^{(-i)}] = (\hat{w}_i^{(-i)})^{-1} + (\sum_{j \neq i} \hat{w}_j^{(-i)})^{-1}$ 、 $\hat{w}_j^{(-i)} = (\hat{\tau}^{2(-i)} + \hat{\sigma}_j^2)^{-1}$ となり、 $\hat{\mu}^{(-i)}$ と $\hat{\tau}^{2(-i)}$ は i 番目の地域を除いた $k - 1$ 個の地域のデータによって推定される。そのため、 t_i は $k - 1$ 個の地域によって推定された変量効果モデル(*)からの i 番目の地域における予測されたstudentized residualとして解釈される。影響の評価に対して、基準となる閾値は t_i の標本分布によって得られ、想定される変量効果モデル(*)が正しければ、 t_i は標準正規分布に従い、 t_i の絶対値は1.96と比較される。基準を満たす時、対応する地域が偶然のばらつきを超える潜在的な外れ値として考えられる。しかし、この基準は大標本近似に基づくため、ばらつきが適切に定量化されていない可能性がある。そのため、ばらつきを考慮するため、パラメトリックブートストラップ法を適用する。

尤度比検定に基づく統計量

平均シフトモデルを用いたモデルベースの尤度比検定を考える。この方法はメタアナリシスに対してNegeri and Beyene [2]により提案されている。変量効果モデル(*)に対して、 j 番目の地域の変量効果モデルが $\theta_j \sim N(\mu + \zeta, \tau^2)$ となり、他の $k - 1$ 個の地域の変量効果モデルが $\theta_i \sim N(\mu, \tau^2)$ ($i \neq j$)となる事を仮定する。この時、以下の検定問題を考える。

$$H_0: \zeta = 0 \text{ vs. } H_1: \zeta \neq 0$$

帰無仮説が棄却される時、 i 番目の治療効果は全体平均から逸脱し、潜在的な外れ値となる地域と判断する。これに対して、尤度比検定を考える。帰無仮説の下で対数尤度関数を $l_0(\mu, \tau^2)$ とし、対立仮説の下での対数尤度関数を $l_{1[j]}(\mu, \tau^2, \zeta)$ とすると、尤度比検定統計量は

$$T_{[j]} = -2\{l_0(\hat{\mu}, \hat{\tau}^2) - l_{1[j]}(\hat{\mu}_{[j]}, \hat{\tau}_{[j]}^2, \hat{\zeta}_{[j]})\}$$

で与えられる。ここで、 $\{\hat{\mu}, \hat{\tau}^2\}$ は帰無仮説の下でのモデル(*)における最尤推定量であり、 $\{\hat{\mu}_{[j]}, \hat{\tau}_{[j]}^2, \hat{\zeta}_{[j]}\}$ は j 番目の地域に対する平均シフトモデルの最尤推定量とする。尤度比検定統計量 $T_{[j]}$ は帰無仮説の下で自由度1のカイ二乗分布に従い、95%点である3.84が基準となる。しかし、大標本近似に基づいているため、パラメトリックブートストラップ法を適用する。

地域間分散と地域内分散を用いた統計量

LOOCV型スキームを用いて μ の全体の推定量の分散の相対的な変化を評価する。この方法はViechtbauer and Cheung [3]によって、メタアナリシスの変量効果モデル(*)に対して提案されており、本研究で国際共同治験に適用する。 j 番目の地域における統計量は

$$VRATIO_j = \frac{\text{Var}[\hat{\mu}^{(-j)}]}{\text{Var}[\hat{\mu}]} = \frac{\sum_{i=1}^k \hat{w}_i}{\sum_{i \neq j} \hat{w}_i}$$

で表され、この統計量によって一つの地域を除いたデータセットと全ての地域を含めたデータセット間の分散の相対的な変化を評価する。 $VRATIO_j$ が1より小さいときに j 番目の地域が全体の治療効果の分散を大きくしていると考え、潜在的な外れ値と判断する。外れ値と判断する基準を決定するため、パラメトリックブートストラップ法によって $VRATIO_j$ の分布を推定し、その分布の下側5%点を基準値とする。

地域間分散を用いた統計量

$VRATIO_j$ と同様のアプローチとして、変量効果モデル(*)の異質性に対して、一つの地域を除いたデータと全地域のデータに対する τ^2 の推定量の比を用いたViechtbauer and Cheung [3]が提案した以下の統計量を考える。

$$TRATIO_j = \frac{\hat{\tau}^{2(-j)}}{\hat{\tau}^2}$$

$TRATIO_j$ は $VRATIO_j$ と同様に1より小さい場合に潜在的な外れ値と考える。パラメトリックブートストラップ法で $TRATIO_j$ の分布を推定し、下側5%点を外れ値の判断の基準とする。

事例解析

ロサルタンの2型糖尿病患者における糖尿病性腎症に対する国際共同治験の事例[4]において4つの提案法を適用した。ロサルタンの国際共同治験では28か国が参加し、アジア地域において他地域と比べて大きな治療効果が得られていた(図1)。アジア地域、欧州地域、中南米地域、北米地域の4地域で提案法を適用した結果、全ての提案法においてアジア地域だけが有意水準5%の閾値を超えていた(表1~3)。アジア地域を除く事により、全体のハザード比とその95%信頼区間が0.85(0.73, 0.99)から0.94(0.79, 1.11)となり、有意であった結果がアジア地域を除く事によって有意差が消失していた。そのため、アジア地域が外れ値となる地域となり、全体の結果に影響していたと考えられる。

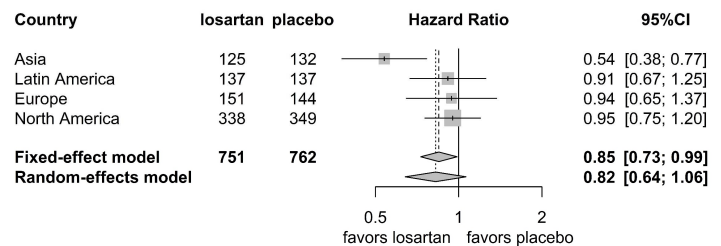


図1 地域ごとのハザード比におけるフォレストプロット

表1 LOOCV studentized residualを用いた統計量の結果

固定効果モデル			変量効果モデル				
地域	t_i	ブートストラップ		地域	t_i	ブートストラップ	
		下側2.5%点	上側2.5%点			下側2.5%点	上側2.5%点
アジア	-2.796	-2.058	1.991	アジア	-2.796	-2.058	1.991
北米	1.290	-1.924	1.912	北米	1.290	-1.924	1.912
欧州	0.599	-2.032	2.002	欧州	0.599	-2.032	2.002
中南米	0.532	-1.907	1.921	中南米	0.532	-1.907	1.921

表2 尤度比検定統計量の結果

固定効果モデル				変量効果モデル			
地域	尤度比検定	ブートストラップ		地域	尤度比検定	ブートストラップ	
		統計量	95%点			p値	統計量
アジア	7.815	3.889	0.009	アジア	6.935	5.013	0.014
北米	1.663	3.725	0.182	北米	0.899	5.381	0.476
欧州	0.358	4.090	0.544	欧州	0.339	5.035	0.624
中南米	0.283	3.676	0.608	中南米	0.257	4.836	0.651

表3 地域間分散と地域内分散に基づく統計量、及び地域間分散に基づく統計量の結果

地域間分散と地域内分散に基づく統計量			地域間分散に基づく統計量		
地域	VRATIO	ブートストラップ下側5%点	地域	TRATIO	ブートストラップ下側5%点
アジア	0.418	0.461	アジア	0.000	0.003
欧州	1.788	0.448	北米	1.529	0.001
北米	1.877	0.521	欧州	1.608	0.002
中南米	1.954	0.462	中南米	1.733	0.002

参考文献

- Belsley DA, Kuh E, Welsch RE. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: Wiley; 1980.
- Negeri ZF, Beyene J. Statistical methods for detecting outlying and influential studies in meta-analysis of diagnostic test accuracy studies. Stat Methods Med Res. 2019.
- Viechtbauer W, Cheung M W. Outlier and influence diagnostics for meta-analysis. Res Synth Methods. 2010; 1(2):112-125.
- Pharmaceuticals and Medical Devices Agency. Review report for Losartan. 2006; Available at: pmda.go.jp/drugs/2006/P200600021/63015300_21000AMZ00678_Q101_1.pdf.