

臨床試験における多重性の調整法(Gatekeeping procedure)

近年の臨床試験では複数の仮説検定を実施するデザインの臨床試験も多い。その場としては単一のエンドポイントでは評価が困難なもの、用量反応関係等 3 群以上の検定を行うもの、増大する臨床試験のコストを抑え込み、効率化を図るため同時に複数のエンドポイントを設定するもの等多岐にわたる。

しかし、複数の検定を実施した場合、誤って真の帰無仮説を棄却する第一種の過誤確率が名義水準を超えることがある。ICH E9(International Conference on Harmonization-Efficacy 9)やEMA(European Medicines Agency)、FDA(Food and Drug Administration)の各ガイドラインでも、多重性を考慮することの重要性は指摘されている。また、複数のエンドポイントを組み合わせた複合エンドポイントを用いる場合も、治療効果を過大評価する可能性等が危惧されている。

そのためこの多重性が問題となる際に、第一種の過誤確率を名義水準以下に抑える手法が検討された。古典的な多重性調整法としては Bonferroni 法、Tukey 法、Dunnnett 法、Scheffe 法等がある。例えば Bonferroni 法は、複数ある検定全体の有意水準を α に保つため、検定数を k とすれば、1 つ 1 つの検定の有意水準を α/k とするものである。この手法では確かに真の帰無仮説を 1 つでも棄却する確率(FWER: Family-Wise Error Rate)を名義水準 α 以下に抑えることが可能であるが、極めて保守的であり、検出力が大きく低下することが問題となっていた。多重性の調整を行わなければ第一種の過誤確率が高まり科学的根拠が欠ける一方で、多重性を調整を行うと有意になりにくく、逆に誤った帰無仮説を棄却できない第二種の過誤確率が高まる。ゆえに第一種の過誤確率を一定の値以下に抑えつつ、可能な限り検出力を向上させる手法の検討・考案が近年に至るまで成されてきた。

Dmitrienko によって考案された Gatekeeping procedure は現在多重性の調整で主流の手法とされる。本抄読会ではこの Gatekeeping procedure を理解するため Dmitrienko らの論文を中心に、検定における多重性の調整方法に関して整理する。信頼区間についても余力があれば纏めたいが、今回は検定の話を中心とする。

また、今後の展望に関しては以下の通りである。Gatekeeping procedure からの発展は困難であると考えた。多重性に関して一定の段階まで理解を進めた後、更に様々な分野を研究していくことで生物統計の一般的な知識を深めるとともに、課題研究に繋がるテーマの発見、検討を進める予定である。

主要文献・参考資料

Dmitrienko A, Ralph B. D'Agostino, Sr, Mohammad F Huque. Key multiplicity issues in clinical drug development. *Statistics in Medicine* 2013;32 1079-1111

2014 年度計量生物セミナー資料

森川敏彦 「閉じた検定手順の臨床応用：多重比較・多重エンドポイントからアダプティブデザインまで」第 6 回久留米大学バイオ統計学フォーラム資料 9/29/2007

Available at <http://www.med.kurume-u.ac.jp/med/biostat/forum/morikawa.pdf> (Accessed 04/17/2015)

永田 靖,吉田道弘. 統計的多重比較法の基礎. サイエンティスト社 1997