

非劣性試験の症例数設計方法の紹介 -生存時間データの場合-

張 方紅・寺尾 工
グラクソ・スミスクライン(株)バイオメディカルデータサイエンス部

An introduction to sample size estimation for non-inferiority trial with time to event data

Fanghong Zhang and Takumi Terao
Biomedical Data Sciences Department, GlaxoSmithKline K.K.

背景:非劣性試験での報告が少ない

生存時間解析における症例数設計

- 優越性:
 - 浜田・藤井(2003)
 - 浜田・安藤(2005)
 - 中西・五条・菅波(2006)
 - 水澤・浜田(2008)
- 非劣性:
 - 浜田・藤井(2003), 76-78枚目

背景:なぜ少ないか?

以下の点について調べた

- 非劣性と優越性の関係
- 非劣性の特徴
- 考察
- 課題

非劣性と優越性 (1/2)

- 両群で必要なイベント数

優越性(Schoenfeld)

$$D = \frac{4(Z_{\alpha} + Z_{\beta})^2}{(\log(\theta))^2}$$

Collett et. al. (1994), p.255
浜田・藤井(2003), 5枚目

非劣性(Schoenfeld)

$$D = \frac{4(Z_{\alpha} + Z_{\beta})^2}{(\log(\theta_1) + \log(\delta_0))^2}$$

ADDPLAN® 5 p.538

D : 両群での必要なイベント数, Z : 標準正規分布の上側パーセント点, θ : ハザード比(Hazard ratio), δ_0 : 非劣性マージン(non-inferiority margin)

非劣性と優越性 (2/2)

- 非劣性を優越性に変換

$$\theta = \theta_1 \times \delta_0$$

$$D = \frac{4(Z_\alpha + Z_\beta)^2}{(\log(\theta))^2}$$

非劣性: 対照薬と比較

- 実対照薬と比較
T: 治験薬 vs C: 対照薬
- 治験薬の対照薬に対する非劣性

$$\theta = \frac{\lambda_t}{\lambda_c}$$

- λ : ハザード関数
- θ : ハザード比
- $\theta < 1$ は治験薬が対照薬より優れると表す

非劣性：対照薬の効果

- 対照薬の効果 (active control effect)
C: 対照薬 vs P: プラセボ (placebo)

$$\delta = \frac{\lambda_c}{\lambda_p}$$

- $\delta < 1$ は対照薬がプラセボより優れると表す

非劣性：非劣性マージン

- 非劣性マージン (δ_0) は対照薬の効果から決められる (森川, 2000)
- 対照薬の効果

$$\delta_0 = \delta \text{ の } 95\% \text{ 信頼区間の上限}$$

- 対照薬の効果の半分

$$\delta_0 = \sqrt{\delta \text{ の } 95\% \text{ 信頼区間の上限}}$$

- $\log(\delta_0) = \log(\delta)$ の 95% 信頼区間の上限の半分, $\delta_0 < 1$

非劣性：統計仮説 (hypotheses)

- δ_0 を非劣性マージンとする .

$$H_0 : \theta \geq \frac{1}{\delta_0} \text{ vs } H_1 : \theta < \frac{1}{\delta_0}$$

$$H_0 : \log(\theta) \geq -\log(\delta_0) \text{ vs } H_1 : \log(\theta) < -\log(\delta_0)$$

$$\theta = \frac{\lambda_t}{\lambda_c}$$

Note: $\delta_0 < 1$ が小さい場合、非劣性が証明しやすくなる

非劣性：治験薬効果の確認

- T薬のC薬に対する非劣性が確立した場合

$$\theta < \frac{1}{\delta_0} \quad \delta \leq \delta_0 \quad (\text{マージンの選択より})$$

$$\theta = \frac{\lambda_t}{\lambda_c} \quad \delta = \frac{\lambda_c}{\lambda_p}$$

$$\theta\delta = \frac{\lambda_t}{\lambda_p} < 1$$

- 非劣性が確立したとき、少なくとも治験薬がプラセボより優れることを保証する .

非劣性マージ: 説明例

- δ の95%信頼上限: 0.63

$$\delta_0 = 0.63$$

$$\delta_0 = \sqrt{0.63} = 0.8$$



$$\theta < \frac{1}{0.63} = 1.59$$

$$\theta < \frac{1}{0.8} = 1.25$$

- いずれも, 非劣性が確立したとき試験薬がプラセボより優れることを保証する.

非劣性と優越性(再掲)

- 非劣性を優越性に変換

$$\theta = \theta_1 \times \delta_0$$

$$D = \frac{4(Z_\alpha + Z_\beta)^2}{(\log(\theta))^2}$$

非劣性と優越性：比較の変換

- 対照薬との比較からプラセボとの比較へ

$$H_0 : \theta = \frac{1}{\delta_0} \text{ vs } H_1 : \theta = \theta_1$$

↓

$$H_0 : \theta\delta = 1 \text{ vs } H_1 : \theta\delta = \theta_1\delta_0$$

$$D = \frac{4(Z_\alpha + Z_\beta)^2}{(\log(\theta_1\delta_0))^2}$$

非劣性：ハザード比

- ハザード比に基づく方法 (Schoenfeld)

$$D = \frac{4(Z_\alpha + Z_\beta)^2}{(\log(\theta_1) + \log(\delta_0))^2} \quad N = \frac{D}{\frac{\pi_t + \pi_c}{2}}$$

D : 両群での必要なイベント数
 Z : 標準正規分布の上側パーセント点
 θ_1 : 対立仮説での(真の)ハザード比
 δ_0 : 非劣性マージン
 N : 両群での症例数
 π_t, π_c : 治験薬群, 対照薬群における死亡率

非劣性：ハザードの差 (1/2)

- ハザードの差 ($\lambda_t - \lambda_c$) に基づく場合
- $\Delta_0 < 0$: 非劣性マージン

$$D = (Z_\alpha + Z_\beta)^2 \frac{(\lambda_t + \lambda_c + \Delta_0)^2}{(\lambda_t - \lambda_c + \Delta_0)^2} \quad (\text{Freedman1})$$

$$N = \frac{D}{\frac{\pi_t + \pi_c}{2}} \quad \text{浜田・藤井 (2003), 77 枚目}$$

非劣性：ハザードの差 (2/2)

- ハザードの差 ($\lambda_t - \lambda_c$) に基づく場合
- N : 両群での症例数
- $\Delta_0 < 0$: 非劣性マージン
- π_t, π_c : 治験薬群, 対照薬群における死亡率

$$N = \frac{2(Z_\alpha + Z_\beta)^2}{(\lambda_t - \lambda_c + \Delta_0)^2} \left(\frac{\lambda_t^2}{\pi_t} + \frac{\lambda_c^2}{\pi_c} \right) \quad (\text{Freedman2})$$

Chow, Shao and Wang (2003), p.169

ハザードの比と差のどっち？

- ハザード比の方が望ましいと思われる：
 - 非劣性マージンの大きさが把握しやすい。例えば $\delta_0=0.8$, $1/\delta_0=1.25$ の場合, 治験薬に25%のハンディキャップを与えると解釈できる。
 - 必要な死亡数 D はハザード比と非劣性マージンだけで計算できる (対照薬のハザード不要) ので, 対照薬の情報が少ないとき, 試験デザインの検討に便利。
 - 検定統計量の分散が仮説の影響を受けにくい。

検定統計量

- X を検定統計量とする。
- μ_0 と μ_1 をそれぞれ, 帰無仮説と対立仮説の下での X の期待値とする。
- σ_0^2 と σ_1^2 をそれぞれ, 帰無仮説と対立仮説の下で X の分散とする。
- 正規分布以外は検定統計量の分散が仮説によって変わる。

Sample Size Equation

- 以下の式をSample size equationと呼ぶ。

$$|\mu_0 - \mu_1| = Z_\alpha \sigma_0 + Z_\beta \sigma_1 \quad (\text{Lachin})$$

Lachin (1981)

浜田・藤井(2003), 18枚目

- 今までの症例数算出式は全てSample size equation (Lachin)の近似である。

ハザード比: Schoenfeld式の導出1/2

- 生存時間が指数分布に従う, 検定統計量を対数ハザード比の最尤推定量とする。

$$X = \log(\hat{\lambda}_t) - \log(\hat{\lambda}_c)$$

$$\mu_0 = \log(\lambda_t) - \log(\lambda_c) = -\log \delta_0$$

$$\mu_1 = \log(\lambda_t) - \log(\lambda_c) = \log \theta_1$$

ハザード比:(Schoenfeld)式の導出2/2

- ハザードの最尤推定の漸近分散を観察Fisher情報量の逆数で計算する場合, 検定統計量Xの分散がハザードに対する仮説と関係しない.

$$\sigma_0^2 = \sigma_1^2 = \frac{1}{d_t} + \frac{1}{d_c} \stackrel{d_t=d_c}{=} \frac{4}{D}$$

- μ_0, μ_1 と σ_0^2, σ_1^2 をSample size equationに代入すると(Schoenfeld)式が得られる.

ハザード差:(Freedman2)の導出

- 検定統計量Xの分散がハザードに対する仮説と関係する.

$$X = \hat{\lambda}_t - \hat{\lambda}_c \quad \mu_0 = \lambda_t - \lambda_c = -\Delta_0, \quad \mu_1 = \lambda_t - \lambda_c$$

$$VX = \frac{\lambda_t^2}{d_t} + \frac{\lambda_c^2}{d_c} \stackrel{n_c=n_t}{=} \frac{\lambda_t^2}{\frac{N}{2}\pi_t} + \frac{\lambda_c^2}{\frac{N}{2}\pi_c}$$

- σ_0^2 と σ_1^2 をVXで近似して, μ_0, μ_1 と一緒にSample size equationに代入すると(Freedman2)式が得られる.

まとめ

- 非劣性の症例数設計は優越性で使われている2つの方法, ハザード比に基づく方法 (Schoenfeld) とハザードの差に基づく方法 (Freedman) が利用できる.
- 使い勝手からハザード比に基づく方法 (Schoenfeld) はより良いと思われる.

補足

- 本報告では, 指数分布のパラメータのハザードに対するWald検定に基づいている.
- 登録, 打ち切りなどを考慮していない.
- 登録, 打ち切りなどを考慮した方法はCrisp and Curtis (2008)を参照する.
- 非劣性マージンを既知の定数としている.

課題

- 紹介した2つの方法ともWald検定統計量に基づき、特に症例数が小さいとき、名目有意水準を保っていないという問題がある。非劣性に適するScore検定統計量に基づく症例数設計方法を開発する必要がある。
- Wald検定に基づく症例数設計方法にハザード関数を定数とする指数分布の仮定が必要である。この仮定を満たすのは困難である。区分ごとの非劣性の方法論研究も必要と思われる。

参考文献 (1/2)

- 浜田知久馬, 藤井陽介(2003), 生存時間解析の症例数設計. 日本SASユーザー会論文集, 73-98.
- 浜田知久馬, 安藤英一(2005), POWERプロシジャによる症例数設計. 日本SASユーザー会論文集, 127-151.
- 中西豊支, 五条正彦, 菅波秀規(2006), POWERプロシジャを用いた生存時間解析による症例数設計の統計学一考察. 日本SASユーザー会論文集, 161-169.
- 森川(2000), 統計ガイドラインと対照薬の選択. EBMジャーナル, Vol.1 5, 130-135.
- 水澤純基・浜田知久馬(2008), 生存時間解析による症例数設計方法の性能比較. 日本SASユーザー会論文集, 19-28.

参考文献 (2/2)

- ADDPLAN Adaptive Designs-Plans and Analyses® User's Guide, Release 5
- Collett D. (1994), Modelling Survival Data in Medical Research. Chapman & Hall: London.
- Chow S.C., Shao J. and Wang H.(2003), Sample Size Calculations In Clinical Research. Marcel Dekker. New York.
- Crisp A. and Curtis P. (2008), Sample size estimation for non-inferiority trials of time-to-event data. Pharmaceutical Statistics, 7: 236–244.
- Lachin J.M. (1981), Introduction to ample size determination and power analysis for clinical trail. Controlled Clinical Trial, 2:93-113.