構造化仮説群の多重性制御法 2 (Graphical Approach, 固定順序法, Fallback 法) Multiplicity Adjustment Methods Based on Structured Hypotheses (Graphical Approach, Fixed Sequence Procedure, Fallback Procedure)

藤原正和
Masakazu Fujiwara

塩野義製薬株式会社 解析センター
Biostatistics Department, Shionogi & Co., Ltd.
e-mail:masakazu.fujiwara@shionogi.co.jp

Multiple test procedures are often used in the analysis of clinical trials including multiple objectives. To control the familywise error rate in the strong sense, various powerful procedures for structured hypotheses have been developed. In this paper, we introduce the graphical approach based on Bonferroni-based closed test procedures. Using the graphical approach, one can discuss different test strategies with the clinical team and build the appropriate multiple test procedure for the given study objectives. We illustrate not only the property of the graphical approach but also how to perform it by using SAS or R. Moreover, fixed sequence procedure and fallback procedure are presented within the framework of the graphical approach.

Key words: Bonferroni, clinical trial, closed testing, familywise error rate, gatekeeping procedure, multiple comparison procedure.

1. はじめに

多重性制御の方法は、複数の目的を含む臨床試験においてしばしば用いられる。例えば、複数の実薬用量群と対照薬群との群間比較、または複数の評価項目に対して新薬のベネフィットを検証したい場合などが挙げられる。Type 1 FWER (familywise error rate) を強く制御する方法として、様々な方法が開発されてきている。最近では、試験の複数の目的に対して、それらの相対的な重要性を反映させた構造化仮説群に対する多重性制御法が発展してきている。具体的な方法としては gatekeeping 法や、本稿で取り上げる固定順序法や fallback 法が該当する。

多重性制御法は、発展してきている一方で課題も幾つかあるのが現状である。ここでは3つの課題を挙げる。1つ目は、閉検定手順を考えると、帰無仮説が増えるにつれて、積仮説が増えいき、計算が煩雑になる点が挙げられる。例として、帰無仮説 $H_1, \dots, H_4$ がある状況を考える。試験の主要目的に関する検定の帰無仮説を含む仮説族を $F_1 = \{H_1, H_2\}$ 、副次目的に関する仮説

族を $F_2 = \{H_3, H_4\}$ とする。Dmitrienko et al. (2003) 等で取り上げられている gatekeeping 法は、このような構造化仮説群に対する多重性制御の方法として有用であるが、gatekeeping 法では 15 個の積仮説 ($2^4 - 1 = 15$) を含む決定行列を考える必要がある。仮説の数がさらに増えると、積仮説も膨大に膨れ上がってしまう。そのため、積仮説を考慮しつつも、効率的に検定を実施する多重性制御法が望まれる。

次に 2 つ目として、多重性制御法が発展する一方で、それ自体が複雑になってきており、臨床チーム(プロジェクトチーム)に対して内在する検定の戦略を説明することが困難さを増してきている点が挙げられる。複雑な多重性制御法を統計家ではない方々へ説明するためのツールが必要である。

最後に 3 つ目として、Holm 法や gatekeeping 法等の検定手順で考慮できる検定の枠組みは限定的である点が挙げられる。複雑な検定の枠組みを持つ臨床試験では、柔軟に検定の枠組みを構築する必要がある。複雑な検定の枠組みとは、例えば前述の帰無仮説 4 つの例において、 $F_1 = \{H_1, H_2\}$ 内の H_1 が棄却できた場合のみ、 F_2 内の H_3 の検定を実施し、 F_1 内のいずれかの仮説が棄却できた場合に H_4 の検定を実施するといった場合などが挙げられる。試験に必要な検定の枠組みに応じて、幅広く適用可能な多重性制御法が望まれる。

本稿で紹介する graphical approach はこれらの課題に対する解決策となる可能性を秘めている。Graphical approach は決定行列を用いることなく検定が可能であり、また graphical approach ではグラフを通して検定の枠組みを視覚的に見ることができると、決定行列を用いる場合に比べて、臨床チームと検定の戦略を議論しやすくなると考えられる。さらに、ルールを守って適切なグラフを描くことで、graphical approach による検定では、柔軟に様々な検定の枠組みを構築可能であり、また、多くの既存の多重性制御法に対応した検定の枠組みも表現可能である。

本稿の構成を述べる。2 章で Bretz et al. (2009) で提案された graphical approach を紹介する。3 章では構造化仮説群の多重性制御法に属する固定順序法(fixed sequence procedure)、及び Wiens and Dmitrienko (2005) で提案された fallback 法(fallback procedure)を紹介し、4 章で graphical approach による検定が閉検定手順であることを具体例を通して述べる。複雑な検定の枠組みに対応した graphical approach の例を 5 章で取り上げ、6 章で SAS と R による graphical approach の実装方法を簡単に紹介する。最後に 7 章で graphical approach に対する考察や最近の話題について述べる。

2. Graphical approach

2.1 Graphical approach とは

本章では Bretz et al. (2009) で提案された graphical approach を紹介する。説明のため、帰無仮説が m 個ある状況($H_1, \dots, H_m$)を考える。Graphical approach を利用するためには、各帰無仮説に対する有意水準と transition matrix を規定することが必要である。

まず、各帰無仮説に対する有意水準についてである。各帰無仮説($H_1, \dots, H_m$)に対して割り当てられている有意水準をそれぞれ $\alpha_1, \dots, \alpha_m$ と表記する。Graphical approach を適用するためには、各帰無仮説に対する有意水準の和が予め規定された有意水準 α 以下($\sum_{i=1}^m \alpha_i \leq \alpha$)であるよ

うに設定する必要がある。

次に, transition matrix を規定する. 本稿では transition matrix を G と表記する. G は各成分が (g_{ij}) である, $m \times m$ の行列である. Transition matrix の各成分 (g_{ij}) は以下の条件を満たすように定義する必要がある.

$$0 \leq g_{ij} \leq 1, g_{ii} = 0 \text{ and } \sum_{k=1}^m g_{ik} \leq 1 \text{ for all } i, j = 1, \dots, m$$

g_{ij} は H_i が棄却されるときに, H_j へ割り当てられる有意水準 α_i の割合とする. 各帰無仮説に対する有意水準と transition matrix を規定することで, 帰無仮説間の関係を反映したグラフを構築することが可能となる.

ここまでの内容について, 具体例の一つ挙げて説明する. 帰無仮説の数を 3 つ (H_1, H_2, H_3), 各帰無仮説に対する有意水準 $\alpha_1, \alpha_2, \alpha_3$ をそれぞれ $\alpha/3$ とし, G を以下のように規定する.

$$G = \begin{pmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \end{pmatrix}$$

このとき, 対応するグラフは図 1 で与えられる. 図 1 内の H_1, H_2, H_3 は各帰無仮説を表し, 各仮説内の $1/3$ とは, α に掛けられる値を意味しており, 各仮説に対して $\alpha/3$ が割り当てられていることに対応する. H_1 から H_2 へ出ている矢印が $g_{12} = 1/2$ を意味する. 他の矢印も同様である. このグラフが表現していることは, 例えば H_1 が棄却されるときに, $g_{12} = 1/2$ であるため, H_1 に割り当てられている有意水準 $\alpha/3$ が H_2 に対して $g_{12} \times \alpha/3 = 1/2 \times \alpha/3 = \alpha/6$ だけ割り当てられることを意味している. つまり, H_1 が棄却されたときに, H_2 に対する有意水準が $\alpha/3 + \alpha/6 = \alpha/2$ に更新される.

図 1. グラフの例

Graphical approach とは, 予め規定した各帰無仮説に対する有意水準と transition matrix のもとで, 以下のアルゴリズムに基づき逐次検定を行っていく方法である. アルゴリズム説明のため, 各帰無仮説に対する未調整 p 値を $p_i, i \in M = \{1, \dots, m\}$ とする.

アルゴリズム

1. $I = M$ とする.
2. $j = \operatorname{argmin}_{i \in I} p_i / \alpha_i$ とする.
3. もし $p_j \leq \alpha_j$ の場合, H_j を棄却して 4 へ進む. もしそうでない場合, この帰無仮説を含む他の全ての帰無仮説を採択して, 検定を終了する.
4. 以下のルールでグラフ (I, G, α) を更新する.

$$I \rightarrow I \setminus \{j\} /$$

$$\alpha_l \rightarrow \begin{cases} \alpha_l + \alpha_j g_{jl} & l \in I \\ 0 & \text{otherwise} \end{cases}$$

$$g_{lk} \rightarrow \begin{cases} \frac{g_{lk} + g_{lj} g_{jk}}{1 - g_{lj} g_{jl}} & l, k \in I, l \neq k \\ 0 & \text{otherwise} \end{cases}$$

5. もし $|I| \geq 1$ の場合, 2 へ戻る. もしそうでない場合は, 検定を終了する.

2.2 Graphical approach の例

前章で述べたアルゴリズムを Holm 法を例にして, 詳細に見ていく. まず Holm 法の手順を復習する. m 個の帰無仮説に対する p 値を小さい順に並び替えた際の記載を $p_{(1)}, \dots, p_{(m)}$, 対応する帰無仮説を $H_{(1)}, \dots, H_{(m)}$ で表す. Holm 法は最も小さな p 値に対する仮説 $H_{(1)}$ から仮説検定が始まり, 以下の手順 l ($l = 1, \dots, m$) で検定を進めていく.

Holm 法の手順

- 手順 1
 $p_{(1)} \leq \alpha/m$ ならば, 仮説 $H_{(1)}$ を棄却し, 手順 2 へ進む.
 それ以外では, 仮説 $H_{(1)}, \dots, H_{(m)}$ を採択し, 検定を終える.
- 手順 2
 $p_{(2)} \leq \alpha/(m-1)$ ならば, 仮説 $H_{(2)}$ を棄却し, 手順 3 へ進む.
 それ以外では, 仮説 $H_{(2)}, \dots, H_{(m)}$ を採択し, 検定を終える.
- 手順 l ($l = 3, \dots, m-1$)
 $p_{(l)} \leq \alpha/(m-l+1)$ ならば, 仮説 $H_{(l)}$ を棄却し, 手順 $l+1$ へ進む.
 それ以外では, 仮説 $H_{(l)}, \dots, H_{(m)}$ を採択し, 検定を終える.
- 手順 m
 $p_{(m)} \leq \alpha$ ならば, 仮説 $H_{(m)}$ を棄却し, 検定を終える.
 それ以外では, 仮説 $H_{(m)}$ を採択し, 検定を終える.

説明のため, 帰無仮説を 3 つとする. この場合, Holm 法では p 値が小さい帰無仮説から順に有意水準 $\alpha/3, \alpha/2, \alpha$ でそれぞれ検定することになる. 例として, 各仮説 H_1, H_2, H_3 に対する p 値を $p_1 = 0.020, p_2 = 0.055, p_3 = 0.012$ とする. これ以降, $\alpha = 0.05$ とする.

図 2 が graphical approach で Holm 法を表現した場合に対応する. 各仮説に対して有意水準 $\alpha/3$

図 2. Holm 法

が割り当てられており, transition matrix は以下になる.

$$G = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

このグラフに対して, 2.1 節で述べた graphical approach のアルゴリズムをあてはめる. 以下の各ステップの番号は, 2.1 節のアルゴリズムに合わせて記載している. Graphical approach のアルゴリズムにおける 5 つのステップを繰り返していることを確認されたい.

- 1-1. $I = \{1, 2, 3\}$ とする.
- 1-2. $j = \operatorname{argmin}_{i \in I} p_i / \alpha_i = 3$ が選ばれる.
- 1-3. $p_3 = 0.012 \leq \alpha/3 = 0.0166$ であるため, H_3 を棄却する.
- 1-4. 以下のルールでグラフ (I, G, α) を更新する.

$$\begin{aligned} I &= I \setminus \{3\} \neq \{1, 2\} / \\ \alpha_1 &= \alpha_1 + \alpha_3 g_{31} = \alpha/2 \\ \alpha_2 &= \alpha_2 + \alpha_3 g_{32} = \alpha/2 \\ g_{12} &= \frac{g_{12} + g_{13}g_{32}}{1 - g_{13}g_{31}} = 1 \\ g_{21} &= \frac{g_{21} + g_{23}g_{31}}{1 - g_{23}g_{32}} = 1 \end{aligned}$$

更新を終えたグラフが図 3 である. なお, G の他成分は 0 である.

- 1-5. $|I| \geq 1$ であるため, 再びアルゴリズムをあてはめる.
- 2-1. $I = \{1, 2\}$ とする.
- 2-2. $j = \operatorname{argmin}_{i \in I} p_i / \alpha_i = 1$ が選ばれる.
- 2-3. $p_1 = 0.020 \leq \alpha/2 = 0.025$ であるため, H_1 を棄却する.

図 3. Holm 法 H_3 棄却後

2-4. 以下のルールでグラフ (I, G, α) を更新する.

$$I = I \setminus \{1\} \neq \{2\} /$$

$$\alpha_2 = \alpha_2 + \alpha_2 g_{12} = \alpha$$

$$g_{22} = 0$$

更新を終えたグラフが図 4 である.

図 4. Holm 法 H_1 棄却後

2-5. $|I| \neq 1$ であるため, 再びアルゴリズムをあてはめる.

3-1. $I = \{2\}$ とする.

3-2. $j = \operatorname{argmin}_{i \in I} p_i / \alpha_i = 2$ が選ばれる.

3-3. $p_2 = 0.055 \not\leq \alpha$ より, H_2 は棄却できず, H_2 を採択して検定は終了となる.

p 値の小さい帰無仮説から順に $\alpha/3, \alpha/2, \alpha$ で検定が実施されており, 図 2 のグラフに対して前述したアルゴリズムを適用することは Holm 法に対応することが確認できる. このように, graphical approach は各帰無仮説に対する有意水準と transition matrix を事前に規定しておき, p 値が得られた後, アルゴリズムを通して, 各帰無仮説に対する有意水準と transition matrix を逐次更新しながら, 検定を行う方法である. この例から分かるように, 各帰無仮説に対する有意水準と transition matrix を変えることで, 様々な多重性制御法を構築することが可能となる. 次章では 2 つの新しい多重性制御法を紹介する.

3. 固定順序法, Fallback 法

固定順序法, fallback 法は構造化仮説群に対する多重性制御法であり, 複数ある帰無仮説に対して, 検定順序を予め規定しておく必要がある. ここでは説明を容易にするため, 帰無仮説が 3 つ (H_1, H_2, H_3) の場合で述べる. また, 各帰無仮説に対する有意水準を α_i とする.

3.1 固定順序法

固定順序法は、各帰無仮説の検定が α で実施されるが、その帰無仮説が棄却されない限り、次の仮説の検定が実施されない手法である。つまり、検定順序が最初の帰無仮説を α で検定して、棄却されれば、2 番目の帰無仮説を α で検定する。また、最初の帰無仮説が棄却されない場合は全ての帰無仮説を採択する。検定順序が 2 番目以降の仮説の検定に対しても同様である。

グラフで表現すると図 5 が固定順序法である。各帰無仮説に対する有意水準 $\alpha_1, \alpha_2, \alpha_3$ はそれぞれ $\alpha, 0, 0$ であり、 G を以下の通り規定した場合が固定順序法となる。

$$G = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

図 5. 固定順序法

図 5 の H_1 内に 1 と記載されているのは、有意水準 α が割り当てられていることを意味する。図 5 から分かるように、もし H_1 が棄却された場合、 H_2 が α で検定され、また H_1 が棄却できなかった場合、 H_1, H_2, H_3 を採択することになる。本手法の長所は、他の多重性制御法と比較して、検定順序が最初の仮説に対する検出力が最も高いことが挙げられる。一方で、ある仮説が棄却できなかった場合にそれ以降の帰無仮説に対して検定が実施されないことが短所として挙げられる。

今は仮説が 3 つの場合であったが、仮説の数が増えても最初の仮説に対する有意水準を α とし、それ以降の仮説に対して有意水準 0 を割り当て、また、仮説間の G の成分を 1 と規定すれば、それが固定順序法に該当する。

3.2 Fallback 法

検定を開始する前に、事前に規定した各帰無仮説 i に対する所与の有意水準を α'_i とし、検定を行う段階で各帰無仮説 i に対して割り当てられている有意水準を α_i で表記する。つまり検定を開始する前は、 $\alpha'_i = \alpha_i$ である。Fallback 法は、 $i-1$ 番目の帰無仮説が棄却されれば、 i 番目の帰無仮説に対する有意水準を $\alpha_i = \alpha'_i + \alpha_{i-1}$ と更新して、 $i-1$ 番目の仮説が棄却されなければ、 $\alpha_i = \alpha'_i$ として検定を行う方法である。

例を図 6 で示す。この例では、 $\alpha'_1 = \alpha/2, \alpha'_2 = \alpha/4, \alpha'_3 = \alpha/4$ である。このとき、例えば H_1

図 6. Fallback 法

が棄却された場合, $\alpha_2 = \alpha'_2 + \alpha_1 = \alpha/4 + \alpha/2 = 3 \times \alpha/4$ と更新される. H_1 が棄却されなかった場合, $\alpha_2 = \alpha'_2 = \alpha/4$ である. 固定順序法と異なり, H_1 が棄却されなかった場合にも H_2 の検定が実施される.

Fallback 法の長所は, 各帰無仮説に対する有意水準を事前に適切に割り当てておくことで, 固定順序法と異なり, 複数ある検定を途中で止めることなく, 最後まで実施できる点である. 反対に, 各帰無仮説に対する検出力が, 割り当てた有意水準に依存するため, その有意水準の設定が困難であることが短所として挙げられる.

今は仮説が 3 つの場合であったが, 仮説の数が増えても全ての帰無仮説に対して α'_i を割り当て, 仮説間の G の成分を 1 と規定すれば, それが fallback 法に該当する. また, fallback 法は様々な改良版が提案されている. 詳細は Wiens and Dmitrienko (2005) を参照されたい.

4. Graphical approach が閉検定手順であること

各帰無仮説に対する有意水準と transition matrix を規定しておき, 得られた p 値に応じて, 2.1 節で述べたアルゴリズムで逐次検定を行っていくことは, 各帰無仮説に対して, 重みが monotonicity condition を満たす, 重み付き Bonferroni 法に基づく閉検定手順を行っていることに対応する. つまり, graphical approach は閉検定手順である. 本章ではこのことを Holm 法といった具体例を通して, 説明する. 一般的な場合における証明は Bretz et al. (2009) に記載されており, 重み付き Bonferroni 法や monotonicity condition については Hommel et al. (2007) に詳細が記載されている. 必要に応じて参照されたい.

まず重み付き Bonferroni 法, 及び monotonicity condition についてそれぞれ説明して, その後, monotonicity condition を満たす重み付き Bonferroni 法について述べる. 最後に, graphical approach が閉検定手順に対応することを述べる.

4.1 Bonferroni 法に基づく閉検定手順

$H_1, \dots, H_m$ を m 個の帰無仮説とする. $I \subseteq \{1, \dots, m\}$ として, 積仮説 H_I を $H_I = \cap \{H_i : i \in I\}$ と表記する. 帰無仮説 i を含む積仮説 H_I に対する重みを $w_i(I)$ とする. $w_i(I)$ は以下の性質を満たすとする.

$$0 \leq w_i(I) \leq 1 \text{ and } \sum_{i \in I} w_i(I) \leq 1$$

帰無仮説 i に対する未調整 p 値を p_i , 積仮説 H_I に対する調整 p 値を p_I^* としたとき, 重み付き Bonferroni 法は, 以下のように積仮説 H_I に対する調整 p 値を求める方法である. あくまでも積仮説に対する調整 p 値であり, 各仮説に対する調整 p 値ではないことに注意する.

$$p_I^* = \min\{q_i(I) : i \in I\}$$

ここで, $q_i(I)$ は次のように求められる.

$$q_i(I) = \begin{cases} p_i/w_i(I) & \text{if } w_i(I) > 0 \\ 1 & \text{if } w_i(I) = 0 \end{cases}$$

H_i を含む全ての積仮説が棄却されるときに, H_i を棄却するとした場合が, 重み付き Bonferroni

表 1. 帰無仮説 i を含む積仮説 I に対する $q_i(I)$

積仮説	含意される帰無仮説		
	H_1	H_2	H_3
H_{123}	$p_1/w_1(1, 2, 3)$	$p_2/w_2(1, 2, 3)$	$p_3/w_3(1, 2, 3)$
H_{12}	$p_1/w_1(1, 2)$	$p_2/w_2(1, 2)$	
H_{13}	$p_1/w_1(1, 3)$		$p_3/w_3(1, 3)$
H_{23}		$p_2/w_2(2, 3)$	$p_3/w_3(2, 3)$
H_1	$p_1/w_1(1)$		
H_2		$p_2/w_2(2)$	
H_3			$p_3/w_3(3)$

法に基づく閉検定手順になる.

例を挙げる. 帰無仮説が 3 つ (H_1, H_2, H_3), 各帰無仮説に対する未調整 p 値を p_1, p_2, p_3 としたとき, 帰無仮説 i を含む積仮説 H_I に対する $q_i(I)$ は表 1 のように定義される. 空欄となっている箇所は, $q_i(I)$ が定義されないことを意味している.

4.2 Monotonicity condition

帰無仮説 i を含む積仮説に対する重みが以下の性質を満たすとき, monotonicity condition が成立するという.

$$w_i(I) \leq w_i(J) \text{ for all } i, I, J \text{ with } i \in I, J \subseteq I$$

つまり, J より I の方が大きい集合であるとき, H_J 内で H_i に与えられる重みの方が, H_I 内で H_i に与えられる重みより大きいことを意味する.

表 2 に例を挙げる. これは monotonicity condition を満たす. 表 2 と前節の内容から, 重み付き Bonferroni 法に対する決定行列が表 3 で得られる. この決定行列は Holm 法の決定行列に該当する. このことから, Holm 法は monotonicity condition を満たす重み付き Bonferroni 法の一種であることが確認できる.

表 2. Monotonicity condition を満たす重みの例 (Holm 法)

積仮説	含意される帰無仮説		
	H_1	H_2	H_3
H_{123}	$w_1(1, 2, 3) = 1/3$	$w_2(1, 2, 3) = 1/3$	$w_3(1, 2, 3) = 1/3$
H_{12}	$w_1(1, 2) = 1/2$	$w_2(1, 2) = 1/2$	
H_{13}	$w_1(1, 3) = 1/2$		$w_3(1, 3) = 1/2$
H_{23}		$w_2(2, 3) = 1/2$	$w_3(2, 3) = 1/2$
H_1	$w_1(1) = 1$		
H_2		$w_2(2) = 1$	
H_3			$w_3(3) = 1$

表 3. 重み付き Bonferroni 法に対する決定行列の例 (Holm 法)

積仮説	p 値	含意される帰無仮説		
		H_1	H_2	H_3
H_{123}	$p_{123}^* = \min\{3p_1, 3p_2, 3p_3\}$	p_{123}^*	p_{123}^*	p_{123}^*
H_{12}	$p_{12}^* = \min\{2p_1, 2p_2\}$	p_{12}^*	p_{12}^*	
H_{13}	$p_{13}^* = \min\{2p_1, 2p_3\}$	p_{13}^*		p_{13}^*
H_{23}	$p_{23}^* = \min\{2p_2, 2p_3\}$		p_{23}^*	p_{23}^*
H_1	$p_1^* = p_1$	p_1		
H_2	$p_2^* = p_2$		p_2	
H_3	$p_3^* = p_3$			p_3

4.3 Monotonicity condition を満たす重み付き Bonferroni 法

Monotonicity condition を満たす重み付き Bonferroni 法を考える. いま, monotonicity condition を満たすため, $i \in \mathcal{J}$, $J \subseteq I$ であるとき, 以下のようになる.

$$p_i/w_i(J) \leq p_i/w_i(I)$$

帰無仮説 H_i を含む積仮説を考えた際, 積仮説 H_J に対する $p_i/w_i(J)$ より積仮説 H_I に対する $p_i/w_i(I)$ が大きな値となる. このことは, 積仮説 H_I が棄却されるとき, 積仮説 H_J も棄却できることを意味している.

帰無仮説が m 個ある場合に, 閉検定手順では通常 $2^m - 1$ 個の積仮説の検定を考慮する必要があるところを, この monotonicity condition を満たす重み付き Bonferroni 法であれば, 本節で述べた性質から m 回の検定で済ませることが可能となる. ここでは次節で具体例を通してこのことを見ていく.

4.4 Graphical approach による検定が閉検定手順であること

2 章の内容に従って各帰無仮説に対する有意水準と transition matrix を規定すること, 及びアルゴリズムを適用することが, 閉検定手順に対応した monotonicity condition を満たす重み付き Bonferroni 法であることを, 2.2 節の例を用いて説明する. 帰無仮説は 3 つ (H_1, H_2, H_3), 各帰無仮説に対する有意水準は $\alpha_1 = \alpha_2 = \alpha_3 = \alpha/3$, transition matrix は以下とする.

$$G = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

このとき, graphical approach のアルゴリズムを思い出すと, 最初に検定が実施される帰無仮説は, $j = \operatorname{argmin}_{i \in I} p_i/\alpha_i$ を満たす H_j である. いまは $\alpha_1 = \alpha_2 = \alpha_3 = \alpha/3$ であるため, $3 \times p_i$ が最小となる帰無仮説が選ばれる. $\min\{3p_1, 3p_2, 3p_3\}$ で選ばれる p 値は, 実は重み付き Bonferroni 法を利用した場合の積仮説 H_{123} に対する p 値と同値である. この説明のため, 4.1 節で紹介した重み付き Bonferroni 法を考えよう. 各帰無仮説に対して割り当てられている有意水準の α の係数を重みとする. 最初に検定が実施される帰無仮説に対して, 重みは全て $1/3$ であり, 4.1 節で

挙げた重みの性質を満たす. また, 重み付き Bonferroni 法の決定行列である表 3 において積仮説 H_{123} における p 値は $\min\{3p_1, 3p_2, 3p_3\}$ であることから, graphical approach のアルゴリズム $\min\{3p_1, 3p_2, 3p_3\}$ で選ばれる p 値は, 重み付き Bonferroni 法を利用した場合の積仮説 H_{123} に対する p 値と同値になる.

次に, ある帰無仮説, 例えば H_3 が棄却された状況を考える. つまり図 3 の状況である. 検定が実施される帰無仮説は, これまでと同様に, $j = \operatorname{argmin}_{i \in I} p_i/\alpha_i$ を満たす H_j である. いま $\alpha_1 = \alpha_2 = \alpha/2$ であるため, $2 \times p_i$ が最小となる帰無仮説が選ばれる. $\min\{2p_1, 2p_2\}$ で選ばれる p 値は, 重み付き Bonferroni 法を利用した場合の積仮説 H_{12} における p 値と同値になる. 各帰無仮説に対して割り当てられている有意水準の α の係数を重みとする. 重みは全て $1/2$ であり, 4.1 節で挙げた重みの性質を満たす. また, 重み付き Bonferroni 法の決定行列である表 3 において積仮説 H_{12} における p 値は $\min\{2p_1, 2p_2\}$ であることから, graphical approach のアルゴリズム $\min\{2p_1, 2p_2\}$ で選ばれる p 値は, 重み付き Bonferroni 法を利用した場合の積仮説 H_{12} に対する p 値と同値になる.

同様に考え, 整理すると, graphical approach における検定対象の各帰無仮説と, 重み付き Bonferroni 法における積仮説との対応関係は表 4 のようになる. 例えば, $I = \{1, 2, 3\}$ のとき(帰無仮説として H_1, H_2, H_3 がある場合), 重み付き Bonferroni 法における積仮説 H_{123} の検定だけに注目すれば良いことになる.

表 4. Graphical approach による検定に対応した重み付き Bonferroni 法の積仮説の検定

graphical approach 検定対象の帰無仮説	重み付き Bonferroni 法 対応する検定の積仮説
H_1, H_2, H_3 (i.e. $I = \{1, 2, 3\}$)	H_{123}
H_1, H_2 (i.e. $I = \{1, 2\}$)	H_{12}
H_1, H_3 (i.e. $I = \{1, 3\}$)	H_{13}
H_2, H_3 (i.e. $I = \{2, 3\}$)	H_{23}
H_1 (i.e. $I = \{1\}$)	H_1
H_2 (i.e. $I = \{2\}$)	H_2
H_3 (i.e. $I = \{3\}$)	H_3

これらを踏まえて各帰無仮説に対して得られた未調整 p 値を $p_1 = 0.020$, $p_2 = 0.055$, $p_3 = 0.012$ とした場合における graphical approach のアルゴリズムの意味合いを説明する. つまり, 閉検定手順になっていることを確認する. 2.2 節のアルゴリズムの順に述べる.

- 1-1. $I = \{1, 2, 3\}$ とする.
- 1-2. $j = \operatorname{argmin}_{i \in I} p_i/\alpha_i = 3$ が選ばれる.
- 1-3. 閉検定手順であるためには, H_3 を含む全ての積仮説の検定を行う必要がある. そのため $H_{123}, H_{13}, H_{23}, H_3$ を全て検定する必要があるが, monotonicity condition を満たしている
と仮定した場合, H_{123} が棄却できれば, $H_{123}, H_{13}, H_{23}, H_3$ が全て棄却されることになる.

また、表 4 で見てきたように、graphical approach のアルゴリズムにおける H_3 の検定は、重み付き Bonferroni 法における H_{123} の検定に対応する。表 3 と表 4 より $p_{123}^* = 3p_3 = 0.036 \leq \alpha$ であることから、 H_{123} を棄却する。

したがって、積仮説 $H_{123}, H_{13}, H_{23}, H_3$ が全て棄却される。

- 1-4. 以下のルールでグラフ (I, G, α) を更新する。

$$\begin{aligned} I &= I \setminus \{3\} \neq \{1, 2\} / \\ \alpha_1 &= \alpha_1 + \alpha_3 g_{31} = \alpha/2 \\ \alpha_2 &= \alpha_2 + \alpha_3 g_{32} = \alpha/2 \\ g_{12} &= \frac{g_{12} + g_{13}g_{32}}{1 - g_{13}g_{31}} = 1 \\ g_{21} &= \frac{g_{21} + g_{23}g_{31}}{1 - g_{23}g_{32}} = 1 \end{aligned}$$

更新を終えたグラフが図 3 である。

- 1-5. $|I| \geq 1$ であるため、再びアルゴリズムをあてはめる。

- 2-1. $I = \{1, 2\}$ とする。

- 2-2. $j = \operatorname{argmin}_{i \in I} p_i / \alpha_i = 1$ が選ばれる。

- 2-3. 閉検定手順であるためには、 H_1 を含む全ての積仮説の検定を行う必要がある。 H_{123}, H_{13} はステップ 1-3 で棄却されているため、 H_{12}, H_1 を全て検定する必要がある。Monotonicity condition を満たしていると仮定した場合、 H_{12} が棄却できれば、 H_{12}, H_1 が棄却されることになる。

また、表 4 で見てきたように、graphical approach のアルゴリズムにおける H_1 の検定は、重み付き Bonferroni 法における H_{12} の検定に対応する。表 3 と表 4 より $p_{12}^* = 2p_1 = 0.04 \leq \alpha$ であるため、 H_{12} を棄却する。

したがって、積仮説 H_{12}, H_1 が全て棄却される。

- 2-4. 以下のルールでグラフ (I, G, α) を更新する。

$$\begin{aligned} I &= I \setminus \{1\} \neq \{2\} / \\ \alpha_2 &= \alpha_2 + \alpha_1 g_{12} = \alpha \\ g_{22} &= 0 \end{aligned}$$

更新を終えたグラフが図 4 である。

- 2-5. $|I| \neq 1$ であるため、再びアルゴリズムをあてはめる。

- 3-1. $I = \{2\}$ とする。

- 3-2. $j = \operatorname{argmin}_{i \in I} p_i / \alpha_i = 2$ が選ばれる。

3-3. 閉検定手順であるためには, H_2 を含む全ての積仮説の検定を行う必要がある. H_{123} , H_{12} , H_{23} はステップ 1-3, 及び 2-3 で棄却されているため, H_2 を棄却する必要がある. $p_2 = 0.055 \not\leq \alpha$ より, H_2 は棄却できず, 検定は終了となる.

このように, 各帰無仮説に対する有意水準と transition matrix を規定して, 2.1 節のアルゴリズムを適用すると, 各帰無仮説 H_i が棄却されるとき, H_i を含むすべての積仮説が棄却されていることがわかる. このことは, この事例だけでなく一般的に成立する. つまり, graphical approach による検定は閉検定手順であり, type 1 FWER を強い意味で制御する. Type 1 FWER を強い意味で制御するため, 各帰無仮説に対する有意水準の合計は, 試験で規定している有意水準 α 以下に設定する必要があることを再度強調しておく.

本節では, 上記ステップ 1-3 や 2-3 において, monotonicity condition を満たすと仮定して話を進めてきたが, 各帰無仮説に対する有意水準と transition matrix を規定して, 2.1 節のアルゴリズムを適用することで, monotonicity condition の仮定は必ず満たされる. 証明は Bretz et al. (2009) に記載されているため, 参照されたい.

5. Graphical approach の臨床試験への適用

Graphical approach で表現可能な多重性制御法として, ここまで Holm 法, 固定順序法, 及び fallback 法を取り上げてきた. これらの手法だけでなく, 臨床試験の目的に合わせてグラフを構築することで, 様々な検定の枠組みが実装可能である. ここでは以下に記した 2 つの検定の枠組みを graphical approach で実装する事例を紹介する. 例では, 2 群の比較試験(実薬群, 対照薬群)を想定する.

1. 主要評価項目に関する帰無仮説のいずれか 1 つが棄却できた際に, 副次評価項目の検定に移る検定の枠組み
2. 主要評価項目, 及び副次評価項目に対して非劣性検定を行い, 非劣性が示せた場合に優越性検定に移る検定の枠組み

5.1 2 つの主要評価項目と 2 つの副次評価項目の検定

主要評価項目を 2 つ (E_1, E_2), 副次評価項目を 2 つ (E_3, E_4) とする. 評価項目 ($E_1, \dots, E_4$) について, 実薬群と対照薬群の比較に関する帰無仮説を順に $H_1, \dots, H_4$ とする.

試験の主要目的はいずれかの主要評価項目に対して, 対照薬群に対する実薬群の有効性を示すこととする. 加えて, 主要目的が達成できた場合に, 副次評価項目に対しても, 対照薬群に対する実薬群の有効性を示すことを副次目的とする. この場合, 主要評価項目に対してまず検定を実施して, いずれかの主要評価項目が棄却された場合に, 副次評価項目の検定を実施するといった検定の枠組みが必要である.

幾つかのグラフを構築することが可能であるが, ここでは Bretz et al. (2009) のグラフを紹介する. それが図 7 である. H_1, H_2 に $\alpha/2$ が割り当てられており, H_3, H_4 に対しては有意水準 0 が割り当てられている. このグラフは, 主要評価項目 H_1, H_2 のいずれかが棄却されない限り, 副次評価項目の検定が実施されないといった検定の枠組みが適切に表現されている. これは parallel

図 7. 2つの主要評価項目と2つの副次評価項目の検定 (Bretz et al. (2009))

gatekeeping 法の一つである.

また, 試験によっては, 主要評価項目をひとつでも多く棄却したい場合も考えられる. その場合は H_1, H_2 を繋ぐようにグラフを構築すれば良い. つまり G を変更することで, 別の検定の枠組みを構築可能である.

5.2 主要評価項目と副次評価項目をそれぞれ1つ含む非劣性検定と優越性検定

$E1, E2$ をそれぞれ主要評価項目, 副次評価項目とする. 試験の主要目的は, 主要評価項目の対照薬群に対する非劣性検証であるとする. また, これが示せた場合に, 主要評価項目に対する優越性検証や, 副次評価項目に対しても同様に非劣性検証, 優越性検証を考える. このとき, 各評価項目に対して, 非劣性が示せた場合に限り, 優越性の検定に移ることが可能な検定の枠組みが必要である.

H_1 を主要評価項目に対する非劣性, H_2 を主要評価項目に対する優越性, H_3 を副次評価項目に対する非劣性, H_4 を副次評価項目に対する優越性検定に対する帰無仮説とする.

この検定の枠組みに合致したグラフの1つを図8に示す. 非劣性が示せた場合に限り, 優越性の検定に移ることといった検定の枠組みが表現されている. 例えば, H_1 が棄却されて, 主要評価項目に対する非劣性が示せた場合に, 初めて, H_2 (主要評価項目に対する優越性)の検定が実施される. また, H_1 が棄却されて, 次に H_3 が棄却された場合, graphical approach の更新アルゴリズムに基づけば, 図8で出ていない H_2 から H_4 への重みの矢印が出ることになる. これはもともと H_2 から H_3 へ, H_3 から H_4 への矢印が出ていて, H_3 が棄却されたため, H_2 から H_4 への重みの矢印が出ることになる. このように, graphical approach を用いる場合は, 各仮説が棄却された場合に, どのようにグラフが更新されて検定が実施されていくかを把握しておくことが重要である.

この例では, G は以下であるが, 実適用の際には各仮説の検出力や重要度等に応じて, 変更すべきである.

図 8. 主要評価項目と副次評価項目をそれぞれ 1 つ含む非劣性検定と優越性検定

$$G = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

なお、この検定の枠組みは、Dmitrienko et al. (2007) で提案された tree-structured gatekeeping 法やこれを一般的に拡張した Dmitrienko and Tamhane (2013) で提案された mixture procedure でも対応可能である。具体的には、第 1 仮説族に H_1 、第 2 仮説族に H_2, H_3 、第 3 仮説族に H_4 が含まれるとして、 H_2, H_3 に対する serial rejection set を H_1 として、 H_4 に対する serial rejection set を H_3 とすると、本節に必要な検定の枠組みに合致する。ただし、各仮説族に対する多重性制御法の選択によって各仮説に対する検出力は変わり、図 8 に基づく graphical approach の検定の検出力とも異なってくる。試験に必要な検定の枠組みと、それに適した多重性制御法を把握したもとの、仮説の重要度や各仮説の想定される検出力等に応じて、適切な多重性制御法を選択していく必要がある。

6. Graphical approach の実装

Graphical approach による検定は SAS と R でそれぞれ実施することが可能である。SAS, R の順で見ていく。

6.1 SAS による実装

Bretz et al. (2011a) の Appendix に graphical approach による検定に対応した SAS プログラムが記載されている。このプログラムには IML プロシジャが利用されているため、SAS の IML プロシジャが利用可能であることが利用のための必要条件である。図 2 のグラフに基づく検定を SAS で実行することを例にして、説明を行う。SAS プログラム実行のために指定する必要がある情報は、各帰無仮説に対する最初の有意水準、transition matrix, h 、及び各仮説に対する未調整

p 値の計 4 つである.

各帰無仮説に対する最初の有意水準と transition matrix はそれぞれ SAS プログラム中では a と w で表される. 図 2 では帰無仮説が 3 つ (H_1, H_2, H_3) あり, 各帰無仮説に対する有意水準を $a = (0.0166 \ 0.0166 \ 0.0166)$ として, transition matrix を以下のように指定すれば良い.

$$w = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

h は各帰無仮説に対する検定結果が有意であるかどうかを表す指示変数である. プログラム実行時は全ての帰無仮説に対して 0 を与えておき, 実行後に 1 となっている箇所に対応する帰無仮説が棄却されたことを意味する. つまり, 実行時には $h = (0 \ 0 \ 0)$ で与えて, 仮に実行後に $h = (1 \ 0 \ 1)$ であった場合には, H_1, H_3 が棄却されたことを意味する.

あとは, 各仮説に対する未調整 p 値を与えることで graphical approach による検定が実行可能である. なお, SAS では graphical approach による検定で, 各帰無仮説が棄却されたかどうかの情報が得られるだけであり, グラフは表示されない.

6.2 R による実装

R では gMCP パッケージを用いることで graphical approach による検定が実施可能である. SAS と同様に, graphical approach による検定で, 各帰無仮説が棄却されたかどうかの情報が得られるだけでなく, R ではユーザーが操作してグラフを構築して, 各仮説に対する p 値を与えて, 検定を順に行い, グラフが更新されていく様子を視覚的に確認することができる. なお, 本稿の図は全て gMCP パッケージを用いて描いたものである. 詳細は R の gMCP パッケージのマニュアルを参照されたい.

7. おわりに

本稿では graphical approach による検定を紹介した. 各帰無仮説に対する有意水準と transition matrix を事前に規定しておき, 得られた p 値に基づき, 2.1 節のアルゴリズムを適用することが graphical approach による検定となる. 検定結果に依存して, グラフが更新されていく様子が視覚的に確認できるため, 複雑な検定の枠組みに関する臨床チームとの議論を助けるツールとなる可能性があると考えられる. FDA の公式見解ではないが, Hung and Wang (2010) では, graphical approach による検定のことを, 用量と評価項目を組み合わせた仮説が複数あるような場合に, 実行可能な手法の一つであると紹介されている. また, 統計家以外からも注目され始めていて, 既存の多重性制御法の理解を促進する役割も果たす可能性があると考えられている.

4 章で見てきたように, graphical approach による検定は, 各帰無仮説に対して, 重みが monotonicity condition を満たす, 閉検定手順に基づいた重み付き Bonferroni 法を行っていることに対応する. なお, アルゴリズムにおいて, $j = \operatorname{argmin}_{i \in I} p_i / \alpha_i$ で選ばれる j が複数ある場合, どれを選択しても最終的に棄却できる帰無仮説には違いがないため, 特に悩む必要はない. なお, 本稿で取り上げた graphical approach はあくまでも重み付き Bonferroni 法に基づくことに注意す

る。Bretz et al. (2011b) では、Simes 法やパラメトリック検定を基にした graphical approach が提案されている。なお、これらの方法はグラフで表現することができない。

また、5 章で述べたように、試験の目的に応じて柔軟に検定の枠組みを構築することが可能な点が graphical approach のメリットとして挙げられる。ただし、柔軟であるがゆえに、graphical approach の使用には注意が必要な点もある。試験の目的に適合した有意水準の割り当てになっているか、また、臨床的に、統計的にも妥当な transition matrix になっているかといった点に注意が必要である。例えば、主要評価項目に対する帰無仮説 A と副次評価項目に対する帰無仮説 B があった場合に、双方向に繋がったグラフを構築したとする。このとき、場合によっては仮説 B が棄却できた後に仮説 A を検定する可能性がある。このことは、重要度の低い副次評価項目に関する仮説の検定結果が、重要度の高い主要評価項目に関する仮説の検定結果へ影響を与えることを意味しており、適切ではないと考えられることが多い。これは 3 章で紹介した固定順序法や fallback 法にも共通する話であり、構造化仮説群の多重性制御法では、十分に注意しておく必要がある。

最後に、最近の話題として、群逐次デザインの臨床試験へ適用可能な graphical approach が提案されたことが挙げられる。これは中間解析を伴う臨床試験において、各解析時点における α 消費関数を利用する graphical approach である。詳細は Maurer and Bretz (2013) を参照されたい。今後も graphical approach は、様々なデザインの臨床試験に対して、適用方法が研究されていくと考えられる。それと共に、臨床試験における実適用が広まっていくことが期待される。

謝 辞

セミナーの講師、及びオーガナイザーの方々から多くの貴重で有益なご意見を頂きました。心から謝意を表します。

参考文献

- Bretz, F., Maurer, W., Brannath, W., and Posch, M. (2009). A Graphical approach to sequentially rejective multiple test procedures. *Statistics in medicine*, **28**, 586–604.
- Bretz, F., Maurer, W., and Hommel, G. (2011a). Test and power considerations for multiple endpoint analyses using sequentially rejective graphical procedures. *Statistics in medicine*, **30**, 1489–1501.
- Bretz, F., Posch, M., Glimm, E., Klinglmueller, F., Maurer, W., and Rohmeyer, K. (2011b). Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biometrical Journal*, **53**, 894–913.
- Dmitrienko, A., Offen, W. W., and Westfall, P. H. (2003). Gatekeeping strategies for clinical trials that do not require all primary effects to be significant. *Statistics in medicine*, **22**, 2387–2400.
- Dmitrienko, A., and Tamhane, A. C. (2013). General theory of mixture procedures for gatekeeping. *Biometrical Journal*, **55**, 402–419.

- Dmitrienko, A., Wiens, B. L., Tamhane, A. C., and Wang, X. (2007). Tree-structured gatekeeping tests in clinical trials with hierarchically ordered multiple objectives. *Statistics in medicine*, **26**, 2465–2478.
- Hommel, G., Bretz, F., and Maurer, W. (2007). Powerful short-cuts for multiple testing procedures with special reference to gatekeeping strategies. *Statistics in Medicine*, **26**, 4063–4073.
- James Hung, H. M., and Wang, S. J. (2010). Challenges to multiple testing in clinical trials. *Biometrical Journal*, **52**, 747–756.
- Maurer, W., and Bretz, F. (2013). Multiple testing in group sequential trials using graphical approaches. *Statistics in Biopharmaceutical Research*, **5**, 311–320.
- Rohmeyer, K., and Klinglmueller, F. (2014). Package ‘gMCP’.
- Wiens, B. L., and Dmitrienko, A. (2005). The fallback procedure for evaluating a single family of hypotheses. *Journal of Biopharmaceutical Statistics*, **15**, 929–942.