Better Health, Brighter Future

SAS ユーザー総会 2017

Mantel-Haenszel 法により調整したリスク差の信頼区間に関する一考察

武田薬品工業株式会社　日本開発センター　生物統計室
佐々木 英麿、舟尾 暢男

Takeda Pharmaceutical Company Limited

要旨

Mantel-Haenszel 法により調整したリスク差に関する以下の信頼区間の算出方法を紹介し、各信頼区間の被覆確率をシミュレーションにより確認することで性能評価を行う。

- Greenland 信頼区間
- Sato 信頼区間
- Newcombe 信頼区間

キーワード

Mantel-Haenszel 法、調整済みリスク差、信頼区間、被覆確率、FREQ プロシジャ、STDRATE プロシジャ

発表概要

通常のリスク差に関する信頼区間

- 飯塚、浜田（2013）：
  SAS V9.3 で算出可能な信頼区間について性能評価
- 飯塚、魚住、浜田（2014）：
  SAS V9.4 で新たに追加された信頼区間についても性能評価

⇒ いずれも Newcombe スコアに基づく信頼区間を推奨

調整済みリスク差に関する信頼区間

- 本発表では、FREQ プロシジャと STDRATE プロシジャで算出可能な Greenland 信頼区間、Sato 信頼区間、Newcombe信頼区間を紹介し、これらの信頼区間について性能評価を行う

本日のメニュー

- **Mantel-Haenszel 法による調整済みリスク差とその信頼区間**
- FREQ/STDRATE プロシジャによる信頼区間算出時の注意点
- 被覆確率による各信頼区間の性能評価

通常のリスク差

投与群	イベントの有無		合計
	あり	なし	
群 1	x_1	$n_1 - x_1$	n_1
群 2	x_2	$n_2 - x_2$	n_2
			n_+

- 投与群が 2 群、二値応答を見る試験を考える
- 群 1 と群 2 のリスク差（群 1 ー 群 2 ）を求める
- 2 つの二項分布のモデル（unconditional）を仮定する

Takeda Pharmaceutical Company Limited

通常のリスク差

- 群 j の被験者のイベント発生確率を p_j とすると、

  $X_j \sim B(n_j, p_j) \sim N(n_j p_j, n_j p_j(1-p_j))$ $(j = 1, 2)$ となり、

  各群のリスクは $Y_j = \frac{X_j}{n_j} \sim N\left(p_j, \frac{p_j(1-p_j)}{n_j}\right)$ となるので、

  群 1 と群 2 のリスク差（群 1 － 群 2 ）は

  $$D = Y_1 - Y_2 \sim N\left(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right)$$ となり、

  リスク差の点推定値とその信頼区間が容易に求まる。

調整済みリスク差

層 i の分割表 $(i = 1, \cdots, k)$

投与群	イベントの有無		合計
	あり	なし	
群 1	x_{i1}	$n_{i1} - x_{i1}$	n_{i1}
群 2	x_{i2}	$n_{i2} - x_{i2}$	n_{i2}
			n_{i+}

- ある層別項目を考え、層が k 個あるとする
- Mantel-Haenszel 法により、層で調整した上でのリスク差（群 1 ― 群 2 ）を求める

Takeda Pharmaceutical Company Limited

調整済みリスク差

- 層 i 群 j の被験者のイベント発生確率を p_{ij} とすると、

$$X_{ij} \sim B(n_{ij}, p_{ij}) \sim N(n_{ij}p_{ij}, n_{ij}p_{ij}(1-p_{ij}))\ (i = 1, .., k;\ \ j = 1, 2)$$

となり、群 j のリスクは $Y_{ij} = \frac{X_{ij}}{n_{ij}} \sim N\left(p_{ij}, \frac{p_{ij}(1-p_{ij})}{n_{ij}}\right)$ となるので、

群 1 と群 2 のリスク差（群 1 － 群 2 ）は

$$D_i = Y_{i1} - Y_{i2} \sim N\left(p_{i1} - p_{i2}, \frac{p_{i1}(1-p_{i1})}{n_{i1}} + \frac{p_{i2}(1-p_{i2})}{n_{i2}}\right)$$ となる。

調整済みリスク差

- Mantel-Haenszel 流の方法を用いて、この層で調整した上でのリスク差（群 1 － 群 2 ）を求める
- 層 i の Mantel-Haenszel 型の重みは $\boldsymbol{w_i} = \frac{\boldsymbol{n_{i1} n_{i2}}}{\boldsymbol{n_{i1} + n_{i2}}}$
- 重み付き推定量は

$$\widehat{RD}_{MH} = \frac{\sum_{i=1}^{k} \boldsymbol{w_i} * D_i}{\sum_{i=1}^{k} \boldsymbol{w_i}}$$

$$= \frac{\boldsymbol{w_1}\left(\frac{x_{11}}{n_{11}} - \frac{x_{12}}{n_{12}}\right) + \boldsymbol{w_2}\left(\frac{x_{21}}{n_{21}} - \frac{x_{22}}{n_{22}}\right) + \cdots + \boldsymbol{w_k}\left(\frac{x_{k1}}{n_{k1}} - \frac{x_{k2}}{n_{k2}}\right)}{\boldsymbol{w_1} + \boldsymbol{w_2} + \cdots + \boldsymbol{w_k}}$$

$$= \frac{\sum_{i=1}^{k} (x_{i1} n_{i2} - x_{i2} n_{i1}) / n_{i+}}{\sum_{i=1}^{k} n_{i1} n_{i2} / n_{i+}}$$

Greenland (1985) の方法に基づいた信頼区間〔Greenland 信頼区間〕

- Mantel-Haenszel 流の方法を用いて、この層で調整した上でのリスク差（群 1 ― 群 2 ）を求める
- $\widehat{RD}_{MH}$の分散は

$$\widehat{\mathrm{var}}\left(\frac{\sum_{i=1}^{k} \boldsymbol{w_i} * D_i}{\sum_{i=1}^{k} \boldsymbol{w_i}}\right) = \frac{\sum_{i=1}^{k} \boldsymbol{w_i}^2 \left(\frac{x_{i1}(n_{i1} - x_{i1})}{n_{i1}^3} + \frac{x_{i2}(n_{i2} - x_{i2})}{n_{i2}^3}\right)}{(\sum_{i=1}^{k} \boldsymbol{w_i})^2}$$

- Greenland (1985) では、上記分散を用いて信頼区間を構成する方法を提案し、層の例数が多い場合に妥当な推定となる
- SAS V9.3 以降では STDRATE プロシジャにて計算可

Takeda Pharmaceutical Company Limited

例: Rothman (2012b)15章より

投与群	イベントの有無		合計
	あり	なし	
群 1	30	174	204
群 2	21	184	205
			409

- 投与群が 2 群、層の数が 2 個の場合を考える
- 群 1 と群 2 の crude なリスク差は 30/204 – 21/205 = 0.0446 ≒ 4.5% となる

例: Rothman (2012b)15章より

層 1 の分割表

投与群	イベントの有無		合計
	あり	なし	
群 1	8	98	106
群 2	5	115	120
			226

層 2 の分割表

投与群	イベントの有無		合計
	あり	なし	
群 1	22	76	98
群 2	16	69	85
			183

SAS Code (for SAS V9.4)

```
data SAMPLEDATA1 ;
  input TREAT STRATA X TOTAL ;
  cards ;
    1 1  8 106
    1 2 22  98
    2 1  5 120
    2 2 16  85
  ;
run ;

ods graphics on;
proc stdrate data=SAMPLEDATA1
  method=mh
  stat  =risk
  effect=diff
  plots =all ;
  population group=TREAT event=X total=TOTAL ;
  strata STRATA / order=internal stats(cl=normal) effect ;
run;
ods graphics off;
```

SAS Code (for SAS V9.4)

Mantel-Haenszel Standardized Risk Estimates

TREAT	Study Population: Observed Events	Study Population: Number of Observations	Study Population: Crude Risk	Mantel-Haenszel: Expected Events	Mantel-Haenszel: Weight	Standardized Risk: Estimate	Standardized Risk: Standard Error	Standardized Risk: 95% Normal Confidence Limits	
1	30	204	0.1471	14.4664	101.802	0.1421	0.0236	0.0959	0.1883
2	21	205	0.1024	10.9134	101.802	0.1072	0.0215	0.0651	0.1493

Risk Effect Estimates

TREAT 1	TREAT 2	Risk Difference	95% Normal Confidence Limits		Standard Error	Z	Pr > |Z|
0.1421	0.1072	**0.0349**	**-.027619**	**0.097419**	0.0319	1.09	0.2739

≒3.5%

Sato (1989) の方法に基づいた信頼区間〔Sato 信頼区間〕

- Sato (1989) では、Greenland (1985) の方法を修正し、層が多くデータが疎な場合でも妥当な推定となる様な $\widehat{RD}_{MH}$ の分散の推定方法を提案
  - SAS V9.4 では FREQ プロシジャにて計算可
  - 各層の各投与群の分母が全て 2 以上であれば、Greenland (1985) の方法で良い？（ Rothman (2012b) ）
- $\widehat{\mathrm{var}}^*\left(\widehat{RD}_{MH}\right) = \left(\widehat{RD}_{MH}\sum_{i=1}^{k}\boldsymbol{P_i} + \sum_{i=1}^{k}\boldsymbol{Q_i}\right)/\left(\sum_{i=1}^{k}\boldsymbol{w_i}\right)^2$
  - $\boldsymbol{P}_i = \left(n_{i1}^2 x_{i2} - n_{i2}^2 x_{i1} + \frac{n_{i1}n_{i2}(n_{i2}-n_{i1})}{2}\right)/(n_{i1}+n_{i2})^2$
  - $\boldsymbol{Q}_i = (x_{i1}(n_{i2}-x_{i2}) - x_{i2}(n_{i1}-x_{i1}))/2(n_{i1}+n_{i2})$

Takeda Pharmaceutical Company Limited

Stratified Wilson Score 信頼区間に基づいた信頼区間〔Newcombe 信頼区間〕

- Yan (2010) では、stratified Wilson score 信頼区間に基づいて信頼区間を算出する方法が提案されている
  - ✓ SAS V9.4 では FREQ プロシジャにて計算可
- $\widehat{RD}_{MH}$ の100(1-α)%信頼下限と上限(L, U)は以下
  - $L = \widehat{RD}_{MH} - z_{\alpha/2}\sqrt{\boldsymbol{\lambda_1} L_1(1-L_1) + \boldsymbol{\lambda_2} U_2(1-U_2)}$
  - $U = \widehat{RD}_{MH} + z_{\alpha/2}\sqrt{\boldsymbol{\lambda_1} U_1(1-U_1) + \boldsymbol{\lambda_2} L_2(1-L_2)}$
    - $\boldsymbol{\lambda_1} = \sum_{i=1}^{k} \boldsymbol{w_i^2} / n_{i1}$, $\boldsymbol{\lambda_2} = \sum_{i=1}^{k} \boldsymbol{w_i^2} / n_{i2}$
    - L_1, U_1 は群 1 の割合の stratified Wilson score 信頼上限と下限
    - L_2, U_2 は群 2 の割合の stratified Wilson score 信頼上限と下限

SAS Code (for SAS V9.4)

```
data SAMPLEDATA2 ;
  input LA TREAT X N ;
  cards ;
  1 1 1 8
  1 1 2 98
  1 2 1 5
  1 2 2 115
  2 1 1 22
  2 1 2 76
  2 2 1 16
  2 2 2 69
  ;
run ;

proc freq data=SAMPLEDATA2 ;
  table  LA*TREAT*X / riskdiff(common column=1 cl=newcombe)
                      nocol nopercent ;
  weight N ;
run ;
```

SAS Code (for SAS V9.4)

FREQ プロシジャ

TREAT * X の要約統計量
層別変数 : LA

手法	共通比率（リスク）差 値	標準誤差	95% 信頼限界	
Mantel-Haenszel	0.0349	0.0319	-0.0276	0.0974
Newcombe	0.0349		-0.0302	0.1000
要約スコア	0.0372	0.0299	-0.0215	0.0959

列 1 (X = 1)

≒3.5%

標本サイズの合計 = 409

本日のメニュー

- Mantel-Haenszel 法による調整済みリスク差とその信頼区間
- **FREQ/STDRATE プロシジャによる信頼区間算出時の注意点**
- 被覆確率による各信頼区間の性能評価

FREQ/STDRATE プロシジャ使用時の注意点

- 調整済みリスク差とその信頼区間を STDRATE プロシジャ又は FREQ プロシジャを用いて算出する例を紹介した
- データの状態によっては、いくつかの信頼区間が算出されない場合があるため注意が必要である。
  - Greenland 信頼区間
    全ての層において「全ての群が 0 %又は 100 %」⇒ ×
  - Sato 信頼区間
    全ての層において「全ての群が 0 %」又は「全ての群が 100 %」⇒ ×
  - Newcombe 信頼区間
    いずれかの群において「全ての層が 0 %又は 100 %」⇒ ×

全ての層において「全ての群が 0 %」又は「全ての群が 100 %」

- 全ての信頼区間が算出不可

層 1 の分割表

投与群	イベントの有無		合計
	あり	なし	
群 1	0	50	50
群 2	0	50	50
			100

層 2 の分割表

投与群	イベントの有無		合計
	あり	なし	
群 1	0	40	40
群 2	0	60	60
			100

全ての層において「全ての群が 0 %又は 100 %」

- Sato 信頼区間のみ算出可

層 1 の分割表

投与群	イベントの有無		合計
	あり	なし	
群 1	0	50	50
群 2	50	0	50
			100

層 2 の分割表

投与群	イベントの有無		合計
	あり	なし	
群 1	40	0	40
群 2	0	60	60
			100

いずれかの群において「全ての層が 0 %又は 100 %」

- Newcombe 信頼区間が算出不可

層 1 の分割表

投与群	イベントの有無		合計
	あり	なし	
群 1	0	50	50
群 2	20	30	50
			100

層 2 の分割表

投与群	イベントの有無		合計
	あり	なし	
群 1	40	0	40
群 2	30	30	60
			100

本日のメニュー

- Mantel-Haenszel 法による調整済みリスク差とその信頼区間
- FREQ/STDRATE プロシジャによる信頼区間算出時の注意点
- **被覆確率による各信頼区間の性能評価**

被覆確率の比較

- Greenland 信頼区間、Sato 信頼区間、Newcombe 信頼区間のそれぞれの被覆確率をシミュレーションにより算出し、性能評価を試みる
  - 信頼区間を算出するための信頼係数：95%
  - 層別項目の層の数：2 層、4 層、8 層
  - 例数：等例数と不等例数
  - 割合：等割合と不等割合
  - シミュレーション回数：10000回
  - 被覆確率： 94.5%未満 → XX.XX 、97.5%超 → XX.XX
  - 各層の群間差が全て 0 の設定：CMH 検定で「有意差なし」となる割合を併記

※ いずれかの群で「全ての層が 0 %」又は「全ての層が 100 %」となった場合を除外してシミュレーションを実施した。

各信頼区間の被覆確率〔２層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH検定
等例数、等割合	5.1.1.1	(20, 20) (20, 20)	(10%, 10%) (10%, 10%)	95.67	95.77	98.36	96.62
	5.1.1.2	(20, 20) (20, 20)	(30%, 30%) (30%, 30%)	94.06	94.45	95.40	95.31
	5.1.1.3	(20, 20) (20, 20)	(50%, 50%) (50%, 50%)	94.60	94.64	94.64	94.64
	5.1.1.4	(50, 50) (50, 50)	(10%, 10%) (10%, 10%)	94.82	94.85	95.78	94.91
	5.1.1.5	(50, 50) (50, 50)	(30%, 30%) (30%, 30%)	94.80	94.89	95.09	95.19
	5.1.1.6	(50, 50) (50, 50)	(50%, 50%) (50%, 50%)	94.55	94.55	94.55	94.55
	5.1.1.7	(100, 100) (100, 100)	(10%, 10%) (10%, 10%)	95.03	95.12	95.49	95.46
	5.1.1.8	(100, 100) (100, 100)	(30%, 30%) (30%, 30%)	95.02	95.06	95.20	95.25
	5.1.1.9	(100, 100) (100, 100)	(50%, 50%) (50%, 50%)	95.20	95.20	95.20	95.20

- Greenland 信頼区間：例数が小さい場合に被覆確率が 94.5% 未満になることが多い
- Newcombe 信頼区間：例数の影響はない

※ 各行の上段：(n_{11}, n_{12}) (p_{11}, p_{12})、各行の下段：(n_{21}, n_{22}) (p_{21}, p_{22})、α：Type I error rate

各信頼区間の被覆確率〔2 層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH検定
等例数、層内で等割合、層間で不等割合	5.1.2.1	(20, 20) (20, 20)	(50%, 5%) (50%, 5%)	93.95	94.74	98.17	95.07
	5.1.2.2	(20, 20) (20, 20)	(50%, 20%) (50%, 20%)	94.18	94.55	96.49	95.46
	5.1.2.3	(20, 20) (20, 20)	(50%, 40%) (50%, 40%)	94.61	94.78	94.87	94.86
	5.1.2.4	(50, 50) (50, 50)	(50%, 5%) (50%, 5%)	94.52	94.83	97.91	95.08
	5.1.2.5	(50, 50) (50, 50)	(50%, 20%) (50%, 20%)	94.87	94.96	96.43	95.17
	5.1.2.6	(50, 50) (50, 50)	(50%, 40%) (50%, 40%)	94.93	94.94	94.95	94.95
	5.1.2.7	(100, 100) (100, 100)	(50%, 5%) (50%, 5%)	94.88	94.99	97.77	95.04
	5.1.2.8	(100, 100) (100, 100)	(50%, 20%) (50%, 20%)	94.93	95.01	96.28	95.21
	5.1.2.9	(100, 100) (100, 100)	(50%, 40%) (50%, 40%)	95.34	95.38	95.43	95.42

- Newcombe 信頼区間: 層間の割合の差が大きい場合に被覆確率が 97.5% を超える傾向

※ 各行の上段:(n_{11}, n_{12}) (p_{11}, p_{12}) 、各行の下段:(n_{21}, n_{22}) (p_{21}, p_{22})、α:Type I error rate

各信頼区間の被覆確率〔2 層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH検定
等例数、層内及び層間で不等割合	5.1.3.1	(20, 20) (20, 20)	(50%, 5%) (5%, 50%)	93.92	98.34	97.95	-
	5.1.3.2	(20, 20) (20, 20)	(50%, 20%) (20%, 50%)	93.91	96.57	96.47	-
	5.1.3.3	(20, 20) (20, 20)	(50%, 40%) (40%, 50%)	94.70	94.94	94.93	-
	5.1.3.4	(50, 50) (50, 50)	(50%, 5%) (5%, 50%)	94.59	98.67	97.68	-
	5.1.3.5	(50, 50) (50, 50)	(50%, 20%) (20%, 50%)	94.82	96.84	96.38	-
	5.1.3.6	(50, 50) (50, 50)	(50%, 40%) (40%, 50%)	94.79	94.87	94.80	-
	5.1.3.7	(100, 100) (100, 100)	(50%, 5%) (5%, 50%)	94.88	98.83	97.82	-
	5.1.3.8	(100, 100) (100, 100)	(50%, 20%) (20%, 50%)	94.94	96.98	96.18	-
	5.1.3.9	(100, 100) (100, 100)	(50%, 40%) (40%, 50%)	95.36	95.46	95.45	-

- Sato 信頼区間: 層内の割合の差が大きい場合に被覆確率が 97.5% を超える傾向
- Newcombe 信頼区間: 層間の割合の差が大きい場合に被覆確率が 97.5% を超える傾向

※ 各行の上段:(n_{11}, n_{12}) (p_{11}, p_{12}) 、各行の下段:(n_{21}, n_{22}) (p_{21}, p_{22})、α:Type I error rate

各信頼区間の被覆確率〔２層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH検定
不等例数、等割合	5.1.4.1	(20, 10) (20, 10)	(10%, 10%) (10%, 10%)	97.19	97.26	98.78	97.33
	5.1.4.2	(20, 10) (20, 10)	(30%, 30%) (30%, 30%)	93.34	93.99	95.24	95.14
	5.1.4.3	(20, 10) (20, 10)	(50%, 50%) (50%, 50%)	94.57	94.98	95.15	95.12
	5.1.4.4	(50, 25) (50, 25)	(10%, 10%) (10%, 10%)	94.80	94.93	95.95	95.28
	5.1.4.5	(50, 25) (50, 25)	(30%, 30%) (30%, 30%)	94.60	94.84	95.34	95.31
	5.1.4.6	(50, 25) (50, 25)	(50%, 50%) (50%, 50%)	94.20	94.26	94.20	95.36
	5.1.4.7	(100, 50) (100, 50)	(10%, 10%) (10%, 10%)	94.95	95.01	95.66	95.15
	5.1.4.8	(100, 50) (100, 50)	(30%, 30%) (30%, 30%)	94.96	95.02	95.15	95.17
	5.1.4.9	(100, 50) (100, 50)	(50%, 50%) (50%, 50%)	94.41	94.45	94.41	94.97

※ 各行の上段：(n_{11}, n_{12}) (p_{11}, p_{12})、各行の下段：(n_{21}, n_{22}) (p_{21}, p_{22})、α：Type I error rate

各信頼区間の被覆確率〔２層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH検定
不等例数、層内で等割合、層間で不等割合	5.1.5.1	(20, 10) (20, 10)	(50%, 5%) (50%, 5%)	93.77	95.04	97.77	95.52
	5.1.5.2	(20, 10) (20, 10)	(50%, 20%) (50%, 20%)	93.69	94.18	96.20	95.42
	5.1.5.3	(20, 10) (20, 10)	(50%, 40%) (50%, 40%)	94.28	94.85	95.21	95.17
	5.1.5.4	(50, 25) (50, 25)	(50%, 5%) (50%, 5%)	94.63	94.95	97.24	95.32
	5.1.5.5	(50, 25) (50, 25)	(50%, 20%) (50%, 20%)	94.69	94.91	95.71	95.27
	5.1.5.6	(50, 25) (50, 25)	(50%, 40%) (50%, 40%)	94.45	94.50	94.45	94.98
	5.1.5.7	(100, 50) (100, 50)	(50%, 5%) (50%, 5%)	94.86	94.99	97.47	95.07
	5.1.5.8	(100, 50) (100, 50)	(50%, 20%) (50%, 20%)	94.72	94.82	95.87	95.06
	5.1.5.9	(100, 50) (100, 50)	(50%, 40%) (50%, 40%)	94.67	94.70	94.67	94.80

※ 各行の上段：(n_{11}, n_{12}) (p_{11}, p_{12})、各行の下段：(n_{21}, n_{22}) (p_{21}, p_{22})、α：Type I error rate

各信頼区間の被覆確率〔２層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH検定
不等例数、層内で等割合、層間で不等割合	5.1.6.1	(10, 20) (10, 20)	(50%, 5%) (50%, 5%)	93.01	95.03	98.56	95.37
	5.1.6.2	(10, 20) (10, 20)	(50%, 20%) (50%, 20%)	93.80	94.40	96.30	95.35
	5.1.6.3	(10, 20) (10, 20)	(50%, 40%) (50%, 40%)	94.13	94.78	95.27	95.21
	5.1.6.4	(25, 50) (25, 50)	(50%, 5%) (50%, 5%)	94.20	94.97	98.18	94.94
	5.1.6.5	(25, 50) (25, 50)	(50%, 20%) (50%, 20%)	94.54	94.73	96.15	95.16
	5.1.6.6	(25, 50) (25, 50)	(50%, 40%) (50%, 40%)	94.42	94.43	94.43	94.77
	5.1.6.7	(50, 100) (50, 100)	(50%, 5%) (50%, 5%)	94.75	95.10	97.99	95.12
	5.1.6.8	(50, 100) (50, 100)	(50%, 20%) (50%, 20%)	95.03	95.09	96.15	95.25
	5.1.6.9	(50, 100) (50, 100)	(50%, 40%) (50%, 40%)	94.85	94.85	94.85	94.86

- Greenland 信頼区間：例数が小さい場合に被覆確率が 94.5% 未満になることが多い
- Newcombe 信頼区間：層間の割合の差が大きい場合に被覆確率が97.5%を超える傾向

※ 各行の上段：(n_{11}, n_{12}) (p_{11}, p_{12})、各行の下段：(n_{21}, n_{22}) (p_{21}, p_{22})、α：Type I error rate

各信頼区間の被覆確率〔2 層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH検定
不等例数、層内及び層間で不等割合	5.1.7.1	(20, 10) (20, 10)	(50%, 5%) (5%, 50%)	93.36	98.04	97.99	-
	5.1.7.2	(20, 10) (20, 10)	(50%, 20%) (20%, 50%)	93.54	96.14	96.36	-
	5.1.7.3	(20, 10) (20, 10)	(50%, 40%) (40%, 50%)	94.29	95.11	95.28	-
	5.1.7.4	(50, 25) (50, 25)	(50%, 5%) (5%, 50%)	94.37	98.38	97.72	-
	5.1.7.5	(50, 25) (50, 25)	(50%, 20%) (20%, 50%)	94.67	96.62	96.12	-
	5.1.7.6	(50, 25) (50, 25)	(50%, 40%) (40%, 50%)	94.93	95.35	95.44	-
	5.1.7.7	(100, 50) (100, 50)	(50%, 5%) (5%, 50%)	94.80	98.56	97.52	-
	5.1.7.8	(100, 50) (100, 50)	(50%, 20%) (20%, 50%)	94.70	96.80	96.13	-
	5.1.7.9	(100, 50) (100, 50)	(50%, 40%) (40%, 50%)	94.65	94.82	94.65	-

- Greenland 信頼区間：例数が小さい場合に被覆確率が 94.5% 未満になることが多い
- Sato 信頼区間：層内の割合の差が大きい場合に被覆確率が 97.5% を超える傾向
- Newcombe 信頼区間：層間の割合の差が大きい場合に被覆確率が 97.5% を超える傾向

※ 各行の上段：(n_{11}, n_{12}) (p_{11}, p_{12})、各行の下段：(n_{21}, n_{22}) (p_{21}, p_{22})、α：Type I error rate

各信頼区間の被覆確率〔２層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH検定
追加検討	5.1.8.1	(10, 10) (10, 10)	(10%, 10%) (10%, 10%)	98.98	99.12	99.78	-
	5.1.8.2	(10, 10) (10, 10)	(50%, 50%) (50%, 50%)	92.29	93.16	96.00	-
	5.1.8.3	(10, 10) (10, 10)	(50%, 10%) (50%, 10%)	92.77	94.15	97.31	-
	5.1.8.4	(10, 10) (10, 10)	(50%, 10%) (10%, 50%)	92.83	97.39	97.43	-
	5.1.8.5	(50, 10) (50, 10)	(10%, 10%) (10%, 10%)	94.37	94.57	96.94	-
	5.1.8.6	(50, 10) (50, 10)	(50%, 50%) (50%, 50%)	94.48	94.50	94.50	-
	5.1.8.7	(50, 10) (50, 10)	(50%, 10%) (50%, 10%)	94.38	94.62	95.96	-
	5.1.8.8	(50, 10) (50, 10)	(10%, 50%) (10%, 50%)	94.22	94.85	96.89	-
	5.1.8.9	(50, 10) (50, 10)	(50%, 10%) (10%, 50%)	94.42	97.02	96.50	-
	5.1.8.10	(50, 10) (10, 50)	(10%, 10%) (10%, 10%)	94.75	95.12	98.37	-
	5.1.8.11	(50, 10) (10, 50)	(50%, 50%) (50%, 50%)	92.66	93.43	94.58	-
	5.1.8.12	(50, 10) (10, 50)	(50%, 10%) (50%, 10%)	92.22	93.12	97.30	-
	5.1.8.13	(50, 10) (10, 50)	(10%, 50%) (50%, 10%)	92.34	94.41	94.26	-
	5.1.8.14	(50, 10) (10, 50)	(50%, 10%) (10%, 50%)	94.48	99.37	99.41	-

※ 各行の上段：(n_{11}, n_{12}) (p_{11}, p_{12})、各行の下段：(n_{21}, n_{22}) (p_{21}, p_{22})、α：Type I error rate

各信頼区間の被覆確率〔4 層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH検定
等例数、等割合	5.2.1.1	(20,20,20,20) (20,20,20,20)	(10%,10%,10%,10%) (10%,10%,10%,10%)	94.24	94.89	96.71	95.10
	5.2.1.2	(20,20,20,20) (20,20,20,20)	(30%,30%,30%,30%) (30%,30%,30%,30%)	94.11	94.53	95.00	94.97
	5.2.1.3	(20,20,20,20) (20,20,20,20)	(50%,50%,50%,50%) (50%,50%,50%,50%)	94.13	94.81	94.92	94.91
	5.2.1.4	(50,50,50,50) (50,50,50,50)	(10%,10%,10%,10%) (10%,10%,10%,10%)	94.65	94.78	95.04	94.97
	5.2.1.5	(50,50,50,50) (50,50,50,50)	(30%,30%,30%,30%) (30%,30%,30%,30%)	94.54	94.69	94.78	94.85
	5.2.1.6	(50,50,50,50) (50,50,50,50)	(50%,50%,50%,50%) (50%,50%,50%,50%)	94.83	94.85	94.85	94.85
	5.2.1.7	(100,100,100,100) (100,100,100,100)	(10%,10%,10%,10%) (10%,10%,10%,10%)	94.58	94.75	94.83	94.81
	5.2.1.8	(100,100,100,100) (100,100,100,100)	(30%,30%,30%,30%) (30%,30%,30%,30%)	94.87	94.95	94.96	94.98
	5.2.1.9	(100,100,100,100) (100,100,100,100)	(50%,50%,50%,50%) (50%,50%,50%,50%)	94.70	94.70	94.70	94.70

- Greenland 信頼区間：例数が小さい場合に被覆確率が 94.5% 未満になることが多い
- Sato 信頼区間、Newcombe 信頼区間：概ね 94.5% ～ 97.5% の範囲内

※ 各行の上段：(n_{11}, n_{12}, n_{13}, n_{14}) (p_{11}, p_{12}, p_{13}, p_{14})、各行の下段：(n_{21}, n_{22}, n_{23}, n_{24}) (p_{21}, p_{22}, p_{23}, p_{24})、α：Type I error rate

各信頼区間の被覆確率〔４層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH検定
等例数、層内で等割合、層間で不等割合	5.2.2.1	(20,20,20,20) (20,20,20,20)	(50%,50%,50%,5%) (50%,50%,50%,5%)	94.50	94.99	97.11	95.49
	5.2.2.2	(20,20,20,20) (20,20,20,20)	(50%,50%,50%,20%) (50%,50%,50%,20%)	94.28	94.54	96.05	94.90
	5.2.2.3	(20,20,20,20) (20,20,20,20)	(50%,50%,50%,40%) (50%,50%,50%,40%)	93.98	94.87	95.09	95.06
	5.2.2.4	(50,50,50,50) (50,50,50,50)	(50%,50%,50%,5%) (50%,50%,50%,5%)	95.13	95.29	96.80	95.33
	5.2.2.5	(50,50,50,50) (50,50,50,50)	(50%,50%,50%,20%) (50%,50%,50%,20%)	94.81	94.88	95.93	95.02
	5.2.2.6	(50,50,50,50) (50,50,50,50)	(50%,50%,50%,40%) (50%,50%,50%,40%)	94.90	94.94	94.95	94.95
	5.2.2.7	(100,100,100,100) (100,100,100,100)	(50%,50%,50%,5%) (50%,50%,50%,5%)	94.96	95.05	96.84	95.17
	5.2.2.8	(100,100,100,100) (100,100,100,100)	(50%,50%,50%,20%) (50%,50%,50%,20%)	95.01	95.05	95.95	95.06
	5.2.2.9	(100,100,100,100) (100,100,100,100)	(50%,50%,50%,40%) (50%,50%,50%,40%)	94.89	94.89	94.89	94.89

※ 各行の上段：(n_{11}, n_{12}, n_{13}, n_{14}) (p_{11}, p_{12}, p_{13}, p_{14})、各行の下段：(n_{21}, n_{22}, n_{23}, n_{24}) (p_{21}, p_{22}, p_{23}, p_{24})、α：Type I error rate

各信頼区間の被覆確率〔４層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH検定
等例数、層内及び層間で不等割合	5.2.3.1	(20,20,20,20) (20,20,20,20)	(50%,50%,50%,5%) (5%,50%,50%,50%)	94.10	96.86	96.77	-
	5.2.3.2	(20,20,20,20) (20,20,20,20)	(50%,50%,50%,20%) (20%,50%,50%,50%)	94.13	95.83	96.06	-
	5.2.3.3	(20,20,20,20) (20,20,20,20)	(50%,50%,50%,40%) (40%,50%,50%,50%)	93.94	94.84	95.01	-
	5.2.3.4	(50,50,50,50) (50,50,50,50)	(50%,50%,50%,5%) (5%,50%,50%,50%)	94.85	96.91	96.76	-
	5.2.3.5	(50,50,50,50) (50,50,50,50)	(50%,50%,50%,20%) (20%,50%,50%,50%)	94.67	95.80	95.77	-
	5.2.3.6	(50,50,50,50) (50,50,50,50)	(50%,50%,50%,40%) (40%,50%,50%,50%)	94.81	94.90	94.90	-
	5.2.3.7	(100,100,100,100) (100,100,100,100)	(50%,50%,50%,5%) (5%,50%,50%,50%)	94.70	96.98	96.65	-
	5.2.3.8	(100,100,100,100) (100,100,100,100)	(50%,50%,50%,20%) (20%,50%,50%,50%)	94.74	95.72	95.67	-
	5.2.3.9	(100,100,100,100) (100,100,100,100)	(50%,50%,50%,40%) (40%,50%,50%,50%)	94.87	94.87	94.87	-

※ 各行の上段：(n_{11}, n_{12}, n_{13}, n_{14}) (p_{11}, p_{12}, p_{13}, p_{14})、各行の下段：(n_{21}, n_{22}, n_{23}, n_{24}) (p_{21}, p_{22}, p_{23}, p_{24})、α：Type I error rate

各信頼区間の被覆確率〔4 層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH 検定
不等例数、等割合	5.2.4.1	(20,20,20,10) (20,20,20,10)	(10%,10%,10%,10%) (10%,10%,10%,10%)	94.27	95.11	97.54	95.70
	5.2.4.2	(20,20,20,10) (20,20,20,10)	(30%,30%,30%,30%) (30%,30%,30%,30%)	94.39	94.73	94.94	95.09
	5.2.4.3	(20,20,20,10) (20,20,20,10)	(50%,50%,50%,50%) (50%,50%,50%,50%)	94.59	94.93	94.97	94.97
	5.2.4.4	(50,50,50,25) (50,50,50,25)	(10%,10%,10%,10%) (10%,10%,10%,10%)	94.78	94.92	95.47	95.15
	5.2.4.5	(50,50,50,25) (50,50,50,25)	(30%,30%,30%,30%) (30%,30%,30%,30%)	94.78	94.91	95.03	95.11
	5.2.4.6	(50,50,50,25) (50,50,50,25)	(50%,50%,50%,50%) (50%,50%,50%,50%)	94.94	95.21	95.27	95.27
	5.2.4.7	(100,100,100,50) (100,100,100,50)	(10%,10%,10%,10%) (10%,10%,10%,10%)	94.91	95.02	95.20	95.08
	5.2.4.8	(100,100,100,50) (100,100,100,50)	(30%,30%,30%,30%) (30%,30%,30%,30%)	94.91	95.00	95.08	95.10
	5.2.4.9	(100,100,100,50) (100,100,100,50)	(50%,50%,50%,50%) (50%,50%,50%,50%)	94.67	94.69	94.67	94.68

※ 各行の上段：(n_{11}, n_{12}, n_{13}, n_{14}) (p_{11}, p_{12}, p_{13}, p_{14})、各行の下段：(n_{21}, n_{22}, n_{23}, n_{24}) (p_{21}, p_{22}, p_{23}, p_{24})、α：Type I error rate

各信頼区間の被覆確率〔４層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH 検定
不等例数、層内で等割合、層間で不等割合	5.2.4.1	(20,20,20,10) (20,20,20,10)	(50%,50%,50%,5%) (50%,50%,50%,5%)	94.46	94.69	96.57	95.12
	5.2.5.2	(20,20,20,10) (20,20,20,10)	(50%,50%,50%,20%) (50%,50%,50%,20%)	94.26	94.98	95.56	95.44
	5.2.5.3	(20,20,20,10) (20,20,20,10)	(50%,50%,50%,40%) (50%,50%,50%,40%)	94.58	94.95	95.03	95.02
	5.2.5.4	(50,50,50,25) (50,50,50,25)	(50%,50%,50%,5%) (50%,50%,50%,5%)	94.84	95.21	96.54	95.54
	5.2.5.5	(50,50,50,25) (50,50,50,25)	(50%,50%,50%,20%) (50%,50%,50%,20%)	94.74	94.91	95.82	95.10
	5.2.5.6	(50,50,50,25) (50,50,50,25)	(50%,50%,50%,40%) (50%,50%,50%,40%)	94.77	95.20	95.32	95.29
	5.2.5.7	(100,100,100,50) (100,100,100,50)	(50%,50%,50%,5%) (50%,50%,50%,5%)	95.34	95.40	96.15	95.41
	5.2.5.8	(100,100,100,50) (100,100,100,50)	(50%,50%,50%,20%) (50%,50%,50%,20%)	94.97	95.16	95.34	95.23
	5.2.5.9	(100,100,100,50) (100,100,100,50)	(50%,50%,50%,40%) (50%,50%,50%,40%)	94.79	94.79	94.79	94.79

※ 各行の上段：(n_{11}, n_{12}, n_{13}, n_{14}) (p_{11}, p_{12}, p_{13}, p_{14})、各行の下段：(n_{21}, n_{22}, n_{23}, n_{24}) (p_{21}, p_{22}, p_{23}, p_{24})、α：Type I error rate

各信頼区間の被覆確率〔４層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH 検定
不等例数、層内で等割合、層間で不等割合	5.2.6.1	(10,20,20,20) (10,20,20,20)	(50%,50%,50%,5%) (50%,50%,50%,5%)	94.28	95.07	97.14	95.58
	5.2.6.2	(10,20,20,20) (10,20,20,20)	(50%,50%,50%,20%) (50%,50%,50%,20%)	94.15	94.69	96.06	95.16
	5.2.6.3	(10,20,20,20) (10,20,20,20)	(50%,50%,50%,40%) (50%,50%,50%,40%)	94.27	94.73	94.80	94.80
	5.2.6.4	(25,50,50,50) (25,50,50,50)	(50%,50%,50%,5%) (50%,50%,50%,5%)	94.74	95.10	96.89	95.31
	5.2.6.5	(25,50,50,50) (25,50,50,50)	(50%,50%,50%,20%) (50%,50%,50%,20%)	94.69	94.89	95.87	94.99
	5.2.6.6	(25,50,50,50) (25,50,50,50)	(50%,50%,50%,40%) (50%,50%,50%,40%)	94.43	94.91	95.19	95.12
	5.2.6.7	(50,100,100,100) (50,100,100,100)	(50%,50%,50%,5%) (50%,50%,50%,5%)	95.04	95.13	96.99	95.19
	5.2.6.8	(50,100,100,100) (50,100,100,100)	(50%,50%,50%,20%) (50%,50%,50%,20%)	94.84	94.88	95.75	94.95
	5.2.6.9	(50,100,100,100) (50,100,100,100)	(50%,50%,50%,40%) (50%,50%,50%,40%)	94.74	94.74	94.74	94.74

※ 各行の上段：(n_{11}, n_{12}, n_{13}, n_{14}) (p_{11}, p_{12}, p_{13}, p_{14})、各行の下段：(n_{21}, n_{22}, n_{23}, n_{24}) (p_{21}, p_{22}, p_{23}, p_{24})、α：Type I error rate

各信頼区間の被覆確率〔4 層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH 検定
不等例数、層内及び層間で不等割合	5.2.7.1	(20,20,20,10) (20,20,20,10)	(50%,50%,50%,5%) (5%,50%,50%,50%)	94.10	96.34	96.23	-
	5.2.7.2	(20,20,20,10) (20,20,20,10)	(50%,50%,50%,20%) (20%,50%,50%,50%)	94.14	95.57	95.71	-
	5.2.7.3	(20,20,20,10) (20,20,20,10)	(50%,50%,50%,40%) (40%,50%,50%,50%)	94.50	95.02	95.05	-
	5.2.7.4	(50,50,50,25) (50,50,50,25)	(50%,50%,50%,5%) (5%,50%,50%,50%)	94.67	96.66	96.55	-
	5.2.7.5	(50,50,50,25) (50,50,50,25)	(50%,50%,50%,20%) (20%,50%,50%,50%)	94.73	95.70	95.62	-
	5.2.7.6	(50,50,50,25) (50,50,50,25)	(50%,50%,50%,40%) (40%,50%,50%,50%)	94.81	94.83	94.81	-
	5.2.7.7	(100,100,100,50) (100,100,100,50)	(50%,50%,50%,5%) (5%,50%,50%,50%)	94.71	96.64	96.37	-
	5.2.7.8	(100,100,100,50) (100,100,100,50)	(50%,50%,50%,20%) (20%,50%,50%,50%)	94.85	95.77	95.60	-
	5.2.7.9	(100,100,100,50) (100,100,100,50)	(50%,50%,50%,40%) (40%,50%,50%,50%)	94.77	94.82	94.77	-

※ 各行の上段：(n_{11}, n_{12}, n_{13}, n_{14}) (p_{11}, p_{12}, p_{13}, p_{14})、各行の下段：(n_{21}, n_{22}, n_{23}, n_{24}) (p_{21}, p_{22}, p_{23}, p_{24})、α：Type I error rate

各信頼区間の被覆確率〔4 層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH 検定
追加検討	5.2.8.1	(10,10,10,10) (10,10,10,10)	(10%,10%,10%,10%) (10%,10%,10%,10%)	95.45	95.88	98.96	96.64
	5.2.8.2	(10,10,10,10) (10,10,10,10)	(50%,50%,50%,50%) (50%,50%,50%,50%)	93.66	94.28	94.30	94.30
	5.2.8.3	(10,10,10,10) (10,10,10,10)	(50%,50%,10%,10%) (50%,50%,10%,10%)	93.55	94.86	97.48	95.38
	5.2.8.4	(10,10,10,10) (10,10,10,10)	(50%,50%,10%,10%) (10%,10%,50%,50%)	93.08	97.69	97.37	-
	5.2.8.5	(20,20,10,10) (20,20,10,10)	(10%,10%,10%,10%) (10%,10%,10%,10%)	94.38	94.84	97.82	95.59
	5.2.8.6	(20,20,10,10) (20,20,10,10)	(50%,50%,50%,50%) (50%,50%,50%,50%)	94.35	94.64	94.61	94.61
	5.2.8.7	(20,20,10,10) (20,20,10,10)	(50%,50%,10%,10%) (50%,50%,10%,10%)	94.39	94.85	96.99	95.20
	5.2.8.8	(20,20,10,10) (20,20,10,10)	(10%,10%,50%,50%) (10%,10%,50%,50%)	93.78	94.65	97.30	95.03
	5.2.8.9	(20,20,10,10) (20,20,10,10)	(50%,50%,10%,10%) (10%,10%,50%,50%)	94.07	97.72	97.02	-

- Greenland 信頼区間：例数が小さい場合に被覆確率が 94.5% 未満になることが多い
- Sato 信頼区間：層内の割合の差が大きい場合に被覆確率が大きくなる傾向
- Newcombe 信頼区間：層間の割合の差が大きい場合に被覆確率が大きくなる傾向

※ 各行の上段：(n_{11}, n_{12}, n_{13}, n_{14}) (p_{11}, p_{12}, p_{13}, p_{14})、各行の下段：(n_{21}, n_{22}, n_{23}, n_{24}) (p_{21}, p_{22}, p_{23}, p_{24})、α：Type I error rate

各信頼区間の被覆確率〔4 層の場合〕

	No.	設定※		被覆確率			1-α
		例数	割合	Greenland	Sato	Newcombe	CMH 検定
追加検討	5.2.8.10	(20,20,10,10) (20,20,10,10)	(50%,10%,50%,10%) (50%,10%,50%,10%)	94.11	95.01	97.42	95.41
	5.2.8.11	(20,20,10,10) (20,20,10,10)	(50%,10%,50%,10%) (10%,50%,10%,50%)	93.76	97.88	97.08	-
	5.2.8.12	(20,20,10,10) (10,10,20,20)	(10%,10%,10%,10%) (10%,10%,10%,10%)	93.84	94.87	98.09	95.67
	5.2.8.13	(20,20,10,10) (10,10,20,20)	(50%,50%,50%,50%) (50%,50%,50%,50%)	93.66	94.30	95.26	95.20
	5.2.8.14	(20,20,10,10) (10,10,20,20)	(50%,50%,10%,10%) (50%,50%,10%,10%)	93.67	94.57	97.39	95.17
	5.2.8.15	(20,20,10,10) (10,10,20,20)	(50%,50%,10%,10%) (10%,10%,50%,50%)	93.73	98.83	98.58	-
	5.2.8.16	(20,20,10,10) (10,10,20,20)	(10%,10%,50%,50%) (50%,50%,10%,10%)	93.19	96.49	95.73	-
	5.2.8.17	(20,20,10,10) (10,10,20,20)	(50%,10%,50%,10%) (50%,10%,50%,10%)	93.61	94.59	97.41	95.29
	5.2.8.18	(20,20,10,10) (10,10,20,20)	(50%,10%,50%,10%) (10%,50%,10%,50%)	93.37	97.84	97.21	-

- Greenland 信頼区間：例数が小さい場合に被覆確率が 94.5% 未満になることが多い
- Sato 信頼区間：層内の割合の差が大きい場合に被覆確率が大きくなる傾向
- Newcombe 信頼区間：層間の割合の差が大きい場合に被覆確率が大きくなる傾向

※ 各行の上段：(n_{11}, n_{12}, n_{13}, n_{14}) (p_{11}, p_{12}, p_{13}, p_{14})、各行の下段：(n_{21}, n_{22}, n_{23}, n_{24}) (p_{21}, p_{22}, p_{23}, p_{24})、α：Type I error rate

各信頼区間の被覆確率〔 8 層の場合〕

No.	設定※		被覆確率			1-α
	例数	割合	Greenland	Sato	Newcombe	CMH 検定
5.3.1	(10,…,10, 10,…,10) (10,…,10, 10,…,10)	(10%,…,10%, 10%,…,10%) (10%,…,10%, 10%,…,10%)	94.05	95.55	98.16	95.90
5.3.2	(10,…,10, 10,…,10) (10,…,10, 10,…,10)	(50%,…,50%, 50%,…,50%) (50%,…,50%, 50%,…,50%)	92.93	94.90	95.25	95.23
5.3.3	(10,…,10, 10,…,10) (10,…,10, 10,…,10)	(50%,…,50%, 10%,…,10%) (50%,…,50%, 10%,…,10%)	93.44	94.73	97.29	95.21
5.3.4	(10,…,10, 10,…,10) (10,…,10, 10,…,10)	(50%,…,50%, 10%,…,10%) (10%,…,10%, 50%,…,50%)	93.84	98.27	97.41	-
5.3.5	(20,…,20, 20,…,20) (20,…,20, 20,…,20)	(10%,…,10%, 10%,…,10%) (10%,…,10%, 10%,…,10%)	94.49	95.06	95.94	95.20
5.3.6	(20,…,20, 20,…,20) (20,…,20, 20,…,20)	(50%,…,50%, 50%,…,50%) (50%,…,50%, 50%,…,50%)	94.48	95.11	95.11	95.11
5.3.7	(20,…,20, 20,…,20) (20,…,20, 20,…,20)	(50%,…,50%, 10%,…,10%) (50%,…,50%, 10%,…,10%)	94.67	95.18	97.17	95.31
5.3.8	(20,…,20, 20,…,20) (20,…,20, 20,…,20)	(50%,…,50%, 10%,…,10%) (10%,…,10%, 50%,…,50%)	94.62	98.36	97.47	-
5.3.9	(20,…,20, 10,…,10) (20,…,20, 10,…,10)	(10%,…,10%, 10%,…,10%) (10%,…,10%, 10%,…,10%)	94.17	94.91	96.60	95.31
5.3.10	(20,…,20, 10,…,10) (20,…,20, 10,…,10)	(50%,…,50%, 50%,…,50%) (50%,…,50%, 50%,…,50%)	94.05	94.99	95.66	95.51
5.3.11	(20,…,20, 10,…,10) (20,…,20, 10,…,10)	(50%,…,50%, 10%,…,10%) (50%,…,50%, 10%,…,10%)	94.31	95.01	96.62	95.14

- 2 層と 4 層のシミュレーション結果と同様の傾向

※ 各行の上段：(n_{11}, …, n_{14}, n_{15}, …, n_{18}) (p_{11}, …, p_{14}, p_{15}, …, p_{18})、各行の下段：(n_{21}, …, n_{24}, n_{25}, …, n_{28}) (p_{21}, …, p_{24}, p_{25}, …, p_{28})
α：Type I error rate

各信頼区間の被覆確率〔8 層の場合〕

No.	設定※		被覆確率			1-α
	例数	割合	Greenland	Sato	Newcombe	CMH 検定
5.3.12	(20,…,20, 10,…,10) (20,…,20, 10,…,10)	(10%,…,10%, 50%,…,50%) (10%,…,10%, 50%,…,50%)	93.69	94.85	97.42	95.10
5.3.13	(20,…,20, 10,…,10) (20,…,20, 10,…,10)	(50%,…,50%, 10%,…,10%) (10%,…,10%, 50%,…,50%)	94.43	98.08	97.14	-
5.3.14	(20,…,20, 10,…,10) (10,…,10, 20,…,20)	(10%,…,10%, 10%,…,10%) (10%,…,10%, 10%,…,10%)	93.85	95.01	96.84	95.14
5.3.15	(20,…,20, 10,…,10) (10,…,10, 20,…,20)	(50%,…,50%, 50%,…,50%) (50%,…,50%, 50%,…,50%)	93.64	94.68	94.99	94.96
5.3.16	(20,…,20, 10,…,10) (10,…,10, 20,…,20)	(50%,…,50%, 10%,…,10%) (50%,…,50%, 10%,…,10%)	93.43	94.49	97.30	94.84
5.3.17	(20,…,20, 10,…,10) (10,…,10, 20,…,20)	(10%,…,10%, 50%,…,50%) (10%,…,10%, 50%,…,50%)	93.61	94.68	97.24	95.12
5.3.18	(20,…,20, 10,…,10) (10,…,10, 20,…,20)	(50%,…,50%, 10%,…,10%) (10%,…,10%, 50%,…,50%)	94.40	99.08	98.62	-
5.3.19	(20,…,20, 10,…,10) (20,…,20, 10,…,10)	(40%,…,40%, 10%,…,10%) (10%,…,10%, 40%,…,40%)	94.35	97.16	96.70	-
5.3.20	(30,…,30, 10,…,10) (30,…,30, 10,…,10)	(50%,…,50%, 10%,…,10%) (10%,…,10%, 50%,…,50%)	94.74	97.96	97.00	-
5.3.21	(40,…,40, 10,…,10) (40,…,40, 10,…,10)	(50%,…,50%, 10%,…,10%) (10%,…,10%, 50%,…,50%)	94.67	97.60	96.68	-
5.3.22	(20,…,20, 10,…,10) (20,…,20, 10,…,10)	(50%,…,50%, 5%,…,5%) (5%,…,5%, 50%,…,50%)	94.13	98.84	97.93	-

- 2 層と 4 層のシミュレーション結果と同様の傾向

※ 各行の上段：(n_{11}, …, n_{14}, n_{15}, …, n_{18}) (p_{11}, …, p_{14}, p_{15}, …, p_{18})、各行の下段：(n_{21}, …, n_{24}, n_{25}, …, n_{28}) (p_{21}, …, p_{24}, p_{25}, …, p_{28})
α：Type I error rate

まとめ

Greenland 信頼区間

- 各群の各層の例数が小さい場合に被覆確率が名義水準を下回る可能性が高くなる
- Rothman（2012b）では、各層の各群の分母が全て 2 以上であれば、Greenland 信頼区間を用いることが可能だと述べているが、シミュレーションの結果からは必ずしもそうではないため、例数が少ない場合においては Greenland 信頼区間の適用は注意が必要であるといえる

Sato 信頼区間

- Greenland 信頼区間と比べて層の例数が小さいことによる影響は受けにくい
- 層内の割合の差が大きい場合に名義水準を上回る傾向がみられた

Newcombe 信頼区間

- 層の数や例数によらず被覆確率が名義水準を維持
- 層間の割合の差が大きい場合には名義水準を上回る傾向がみられた

本日のメニュー

- Mantel-Haenszel 法による調整済みリスク差とその信頼区間
- FREQ/STDRATE プロシジャによる信頼区間算出時の注意点
- 被覆確率による各信頼区間の性能評価

参考文献

- Greenland S. and Robins J. M. (1985), “Estimation of a common effect parameter from sparse follow-up data,” Biometrics, 41(1): 55-68.
- Sato T. (1989), “On variance estimator for the Mantel-Haenszel risk difference. (Letter),” Biometrics; 45: 1323-1324.
- Yan X. and Su X. G. (2010), "Stratified Wilson and Newcombe Confidence Intervals for Multiple Binomial Proportions," Statistics in Biopharmaceutical Research; Vol. 2, No. 3: 329-335.
- Rothman K. J. (2012a), “Epidemiology: An Introduction, 2nd Ed.,” Oxford University Press.
- Rothman K. J., Lash T. L., and Greenland S. (2012b), “Modern Epidemiology, 3rd Ed.,” Lippincott Williams & Wilki.
- SAS/STAT(R) 14.1 User's Guide.
- 飯塚 政人, 浜田 知久馬（2013）, "2群の割合の差における信頼区間の構成法の比較," SASユーザー総会.
- 飯塚 政人, 魚住 龍史, 浜田 知久馬（2014）, "FREQプロシジャによる割合の差の信頼区間," SASユーザー総会.

Takeda Pharmaceutical Company Limited