

ワクチン試験での必要イベント数の設定について

○浦狩保則¹

(¹シミック (株) データサイエンス事業本部)

On Determination of Required Number of Events in Vaccine Study

Yasunori,Uragari

Data Science Div., CMIC Co.,Ltd.

要旨

本稿では、ワクチン群とプラセボ群の2群比較試験の計画において、2群のイベント発現数合計に条件付きでワクチン群のイベント発現数が従う二項分布に基づいて、ワクチン有効性についての仮説検定に必要なイベント発現数を算出する方法（条件付き法）に関する検討結果を記述した。本稿では、条件付き法の適用が妥当であるための条件を、両群の試験期間中の発現割合が0.3以下であることと設定した。必要被験者数ではなく、必要イベント数に注目したのは、イベント数による試験の進捗管理が行いやすいためである。そのことを、イベント発現数合計についての条件付けを行わず必要被験者数を算出する方法との比較に基づいて示した。さらに、ワクチン有効性の推定において、Senn(2022)によるケース・プロポーションの95%信頼区間算出法が優れた性質をもつことを示したうえで、ワクチン有効性の推定精度の観点から必要イベント数を算出するための一案を提示した。

キーワード：ワクチン試験、必要イベント数、ワクチン有効性、リスク比、ケース・プロポーション

1. ワクチン試験の特徴

ワクチン群とプラセボ群のイベント（例えば、感染、発症）の発現割合（あるいは発現率）の比較試験を考える。最初に、そのようなワクチン試験について、3つの特徴を踏まえておきたい。第1の特徴は、有効性の評価指標として、ワクチン有効性が用いられることである。ワクチン有効性は、 $1 - \text{リスク比}$ （ $= 1 - \text{相対リスク}$ ）であり、リスク比は、 $(\text{ワクチン群のイベント発現割合}) / (\text{プラセボ群のイベント発現割合})$ で定義する。イベント発現割合の大小に関わらず、イベント発現割合の比によって、リスク比、そしてワクチン有効性が決定されることに注意したい。そして、リスク比が1より小さな値であるほど、ワクチン有効性が高くなる。第2の特徴は、ワクチンの有効性の証明にあたって、対プラセボ優越性を示すこと、具体的にワクチン有効性=0の帰無仮説を棄却し、ワクチン有効性>0の対立仮説を採択すること（あるいは、リスク比=1の帰無仮説を棄却し、リスク比<1の対立仮説を採択すること）だけでは不十分であると見なされるのが通常になっていることである。第3の特徴は、ワクチン試験の検出力に対し、両群で発現するイベント数が直接的な影響力をもつことである。本稿のタイトルの通り、必要イベント数を設定すること、および両群で生起

しているイベント総数を用いて試験の進捗管理を行うことが有用である(Senn,2022)。イベント発現割合（あるいは発現率）の大小が影響するのは、必要被験者数あるいは1被験者あたりの観察期間である。イベント発現割合（あるいは発現率）が小さいほど、より多くの被験者あるいはより長期の観察が必要となる。

本稿の以降の構成は次の通りである。2節では、ワクチン試験における必要イベント数の設定法として、両群の発現総イベント数に条件付きで、ワクチン群の発現イベント数が従う条件付き二項分布に基づく方法（条件付き法と略す）に着目するとともに、その方法の適用が妥当であるための条件について述べる。3節では、試験の進捗管理において、条件付き法により算出される必要総イベント数を用いる方法が、必要総被験者数を用いる方法よりも優れたものであることを示す。4節では、ワクチン有効性の推定方法に目を向け、ワクチン有効性の推定精度の観点からの必要イベント数の設定法の一案を提示する。最後の5節で、上記事項のまとめを示すとともに、関連するいくつかの事項について触れる。

2. ワクチン有効性についての仮説検定と必要イベント数

2.1 ワクチン試験のデザイン時の諸特性についての規定

ワクチン群を v 、プラセボ群を p で表し、両群への被験者数の割付け比率は1対1であるとし、そのもとでの両群の被験者数を $n_v = n_p = n$ で表す。これらの群の発現率（件/人年）をそれぞれ Λ_v 、 Λ_p で表す。試験の追跡期間が D （年）の場合、試験期間中の発現割合はワクチン群、プラセボ群でそれぞれ $P_v = \Lambda_v D$ 、 $P_p = \Lambda_p D$ と表せる。（ここでは、すべての被験者について追跡期間 D の全体にわたってイベントが観測されること、追跡期間中にイベント未発現の状態です試験から脱落することがないこと、各被験者におけるイベント発現数は0回または1回であることを仮定している。）

このとき、両群のイベント発現数 Y_v 、 Y_p の確率分布はそれぞれ二項分布 $Y_v \sim \text{Bin}(n, P_v)$ 、 $Y_p \sim \text{Bin}(n, P_p)$ で表される。 n が大きく、 P_v 、 P_p が適度に小さいとき、 Y_v 、 Y_p の確率分布がPoisson分布 $Po(\lambda_v)$ 、 $Po(\lambda_p)$ で近似されることがよく知られている。ここで、 $\lambda_v = nP_v$ 、 $\lambda_p = nP_p$ である。

今、リスク比（あるいは相対リスク）を $RR = P_v/P_p$ で表すと、ワクチン有効性（Vaccine Effect）は $\pi = 1 - RR$ で与えられる。これらと密接な関係をもつ指標として、ケース・プロポーション

$$\theta = \frac{nP_v}{nP_v + nP_p} = \frac{\lambda_v}{\lambda_v + \lambda_p} = \frac{1 - \pi}{2 - \pi} \quad (1)$$

がある(Chan & Bohider,1998 ; Loiacono,2018)。この割合は、(ワクチン群の期待イベント数)/(両群の期待総イベント数)であり、このことが θ をケース・プロポーションと呼ぶ理由になっている(Senn,2022)。(1)から $\pi = (2\theta - 1)/(\theta - 1)$ の関係が成立する。

今、両群のイベント数がそれぞれ独立にPoisson分布に従う、すなわち $Y_v \sim Po(\lambda_v)$ 、 $Y_p \sim Po(\lambda_p)$ である場合、それらの同時分布は、Poisson分布の再生性を通して

$$p(Y_v, Y_p | \lambda, \theta) = \frac{\lambda^T e^{-\lambda}}{T!} \times \frac{T!}{Y_v!(T - Y_v)!} \theta^{Y_v} (1 - \theta)^{T - Y_v} \quad (2)$$

で示される(Przyborowski & Wilenski,1940)。ここで、 $T = Y_v + Y_p$ は両群の総イベント数、および $\lambda = \lambda_v + \lambda_p$ である。(2)式右辺の第1因数は、パラメータ λ のPoisson分布のもとで、 T が起こる確率であり、第2因数は T を与えたもとでのワクチン群のイベント数 Y_v の確率分布を表し、それがパラメータ θ の二項分布であることがわかる。すなわち、 T に条件付きで、 Y_v の確率分布は

$$Y_v | T \sim \text{Bin}(T, \theta) = \binom{T}{k} \theta^k (1 - \theta)^{T - k} \quad (3)$$

で表される(Przyborowski & Wilenski,1940; Chan & Bohidar, 1998; Senn,2022)。前節で述べた条件付き法は、(3)式の条件付き二項分布に基づく方法を指している。

2.2 ワクチン有効性の仮説検定で必要なイベント数

ワクチン有効性についての検定仮説

$$H_0: \pi \leq \pi_0 \text{ vs } H_1: \pi > \pi_0$$

を考える (π_0 は0とは限らない)。これは、ケース・プロポーションについての検定仮説

$$H_0: \theta \geq \theta_0 \text{ vs } H_1: \theta < \theta_0, \text{ ただし } \theta_0 = (1 - \pi_0)/(2 - \pi_0)$$

で置き換えることができる。そして、両群の総イベント数 T およびワクチン群のイベント数 $y_{v,obs}$ が得られた場合、検定の p 値は(3)式を用いて

$$p = \Pr(Y_v \leq y_{v,obs} | \theta_0) = \sum_{k=0}^{y_{v,obs}} \binom{T}{k} \theta_0^k (1 - \theta_0)^{T-k} \quad (4)$$

で算出される。検定の片側有意水準を α で表すと、検定の棄却限界値 Y_c は

$$Y_c = \text{Max}_Y [\Pr(Y_v \leq Y | \theta_0) < \alpha] \quad (5)$$

となり、このときの実際の α レベルは

$$\Pr(Y_v \leq Y_c | \theta_0) = \sum_{k=0}^{Y_c} \binom{T}{k} \theta_0^k (1 - \theta_0)^{T-k} \quad (6)$$

となる。ワクチン有効性についての期待値 (対立仮説の期待値) を π_1 とすると、それに対応するケース・プロポーション値は $\theta_1 = (1 - \pi_1)/(2 - \pi_1)$ であり、検定の検出力は

$$1 - \beta = \Pr(Y_v \leq Y_c | \theta_1) = \sum_{k=0}^{Y_c} \binom{T}{k} \theta_1^k (1 - \theta_1)^{T-k} \quad (7)$$

で評価することができる。

検定仮説を、ワクチン有効性 (π_1 :期待値、 π_0 : 帰無仮説値) と対応するケース・プロポーション (θ_1 :期待値、 θ_0 : 帰無仮説値) で記述したもとの、片側有意水準 0.025 で 0.9 以上の検出力を得るために必要な総イベント数 (T_{req}) と棄却限界値 (Y_c : 有意差が得られるワクチン群のイベント数の最大値) を表 1 に示す。

表 1. ワクチン有効性の仮説検定ための必要総イベント数 T_{req} と棄却限界値 Y_c

(片側有意水準:0.025、必要な検出力:0.90)

ワクチン有効性による記述		ケース・プロポーションによる記述		必要総イベント数: T_{req}	棄却限界値: Y_c	α レベル: (6)式	検出力: (7)式
π_1	π_0	θ_1	θ_0				
0.1	0.0	0.474	0.500	3,845	1,861	0.025	0.903
0.2	0.0	0.444	0.500	880	410	0.023	0.906
	0.1	0.444	0.474	3,115	1,420	0.024	0.903
0.3	0.0	0.412	0.500	350	156	0.024	0.910
	0.1	0.412	0.474	700	305	0.024	0.907
	0.2	0.412	0.444	2,457	1,043	0.024	0.904
0.4	0.0	0.375	0.500	178	75	0.021	0.911
	0.1	0.375	0.474	279	115	0.023	0.910
	0.2	0.375	0.444	548	220	0.023	0.907
	0.3	0.375	0.412	1,901	740	0.024	0.904
0.5	0.0	0.333	0.500	99	39	0.022	0.916
	0.1	0.333	0.474	138	53	0.021	0.911
	0.2	0.333	0.444	216	81	0.023	0.914
	0.3	0.333	0.412	419	152	0.023	0.908
	0.4	0.333	0.375	1,431	500	0.024	0.906

<必要イベント数算出時の留意点>

必要イベント数の算出場面では、両群の総イベント数 T を、例えば1から5000まで1刻みで動かし、(6)の実際のレベルが α を超えず、(7)の検出力が目標検出力を上回るように、 T とそれもとの棄却限界値 Y_c を見出すことになる。ただし、そのような総イベント数 T を見出すことができたとしても、それ以上の総イベント数のもとで検出力が目標値を下回ることが起こり得る。例えば、表1において $\pi_1 = 0.4, \pi_0 = 0.0$ の場合について、

α レベルは常に名目有意水準以下に保たれる。しかし、検出力は、 $T = 169$ のとき初めて 0.9 を超えるが、それ以上のいくつかの T の値 (170、172、177) で 0.9 を下回り、 $T = 178$ 以上で常に 0.9 を超えることが観察される。これは、二項分布の離散的性質 (discrete nature) によるものである。そのため、十分に大きい総イベント数のもとでの検出力も考慮して、必要総イベント数を設定するべきである。

<必要被験者数>

必要総イベント数 T_{req} が算出されれば、当該試験における必要被験者数の算出へ移る。その際、試験期間 D (年)、ワクチン群およびプラセボ群の期待イベント発現率 Λ_v 、 Λ_p (件/人年) を用いることが通常である。(なお、期待イベント発現率の単位を件/人年としているが、同一被験者において初発のイベントに注目し、初発イベントの発現までの期間で人年をカウントする場合は、期待イベント発現率の単位は人/人年となり、すなわち期待イベント発現率は試験期間当たりのイベント発現割合と見なすことができる。) また、期待発現率の比がリスク比に等しいこと、すなわち $\Lambda_v/\Lambda_p=RR$ であることを仮定する。このとき、 $\Lambda_v=\Lambda_p RR = \Lambda_p(1 - \pi_1)$ である。 D (年)の試験期間中に、2 群合計で T_{req} 件のイベントを観察するための 1 群あたりの必要被験者 N_{req} は

$$N_{req} = \frac{T_{req}}{\{D(\Lambda_p + \Lambda_v)\}} = \frac{T_{req}}{\{D(2 - \pi_1)\Lambda_p\}} \quad (8)$$

で算出される。割付比が 1 対 1 であるので、必要被験者数は 2 群合計で $2N_{req}$ となる。

表 1 で示した各ケースについて、プラセボ群の期待イベント発現率を $\Lambda_p=0.2$ 件/人年と想定した場合で、 $D=0.5$ 年、 $D=1$ 年とする場合の必要被験者数 (2 群合計) を表 2 に示す。

表 2. 表 1 の各ケースの必要被験者数:プラセボ群の期待イベント発現率を $\Lambda_p=0.2$ 件/人年と想定

ワクチン有効性		プラセボ群	ワクチン群	必要被験者数 (2 群合計)	
π_1	π_0	発現率 (件/人年)	発現率 (件/人年)	試験期間 0.5 年	試験期間 1 年
0.1	0.0	0.2	0.18	40,474	20,238
0.2	0.0	0.2	0.16	9,778	4,890
	0.1	0.2	0.16	34,612	17,306
0.3	0.0	0.2	0.14	4,118	2,060
	0.1	0.2	0.14	8,236	4,118
	0.2	0.2	0.14	28,906	14,454
0.4	0.0	0.2	0.12	2,226	1,114
	0.1	0.2	0.12	3,488	1,744
	0.2	0.2	0.12	6,850	3,426
	0.3	0.2	0.12	23,764	11,882
0.5	0.0	0.2	0.1	1,320	660
	0.1	0.2	0.1	1,840	920
	0.2	0.2	0.1	2,880	1,440
	0.3	0.2	0.1	5,588	2,794
	0.4	0.2	0.1	19,080	9,540

試験期間を長く設定すれば、必要イベント数を達成するために必要な被験者数は少なくなることが伺える。

2.3 条件付き法の適用条件

条件付き法は(2)式に基づいているが、(2)式が妥当であるためには、両群のイベント発現数 Y_v 、 Y_p の確率分布が Poisson 分布で十分に近似できることが前提であった。2.1 節では、Poisson 近似のための条件を「 n が大きく、 P_v 、 P_p が適度に小さい」と記載したが、ここではさらに詳しく調べることにした。具体的に、被験者数 n を 20~500 まで (20 刻み) の範囲、イベント発現割合 P を 0.1~0.8 まで (0.1 刻み) の範囲、で動かしたときの各組合せで、イベント発現数 Y についての二項分布上の理論確率と Poisson 分布上の理論確率の相違を、カルバック・ライブラー距離 (Kullback-Leibler Distance)

$$KLD = \sum_{x=0}^n [p(x|n, P) \log_2 \{p(x|n, P) / q(x|nP)\}], \quad p(x|n, P) : \text{二項確率}, \quad q(x|nP) : \text{Poisson 確率}$$

で評価した(杉山,2013)。(n,P)=(200,0.1)、(200,0.3)、(200,0.5)、(200,0.7)の各場合における二項分布と Poisson 分布を重ねて描写するとともに、KLD 値(単位は bit)を図 1 に示した。また、0.1~0.8 まで P のそれぞれで、n の変化に伴う KLD 値の挙動を図 2 に示した。図 2 より、二項分布と Poisson 分布の相違を KLD 値で計量する場合、KLD 値に大きく影響するのは P の値であり、n の大小の影響は小さいことが見て取れる。KLD 値がどの程度に小さい値であれば、二項分布への Poisson 分布近似による情報損失が十分に小さいかという明確な基準は存在しないと思われる。ただ、図 1 において P = 0.3 の場合に KLD 値が 0.05bit 未満で、Poisson 近似がほぼ有効であると見なすことができるであろう。したがって、条件付き法の適用が妥当であるための条件として、イベント発現割合が 0.3 以下であることを提案したい。

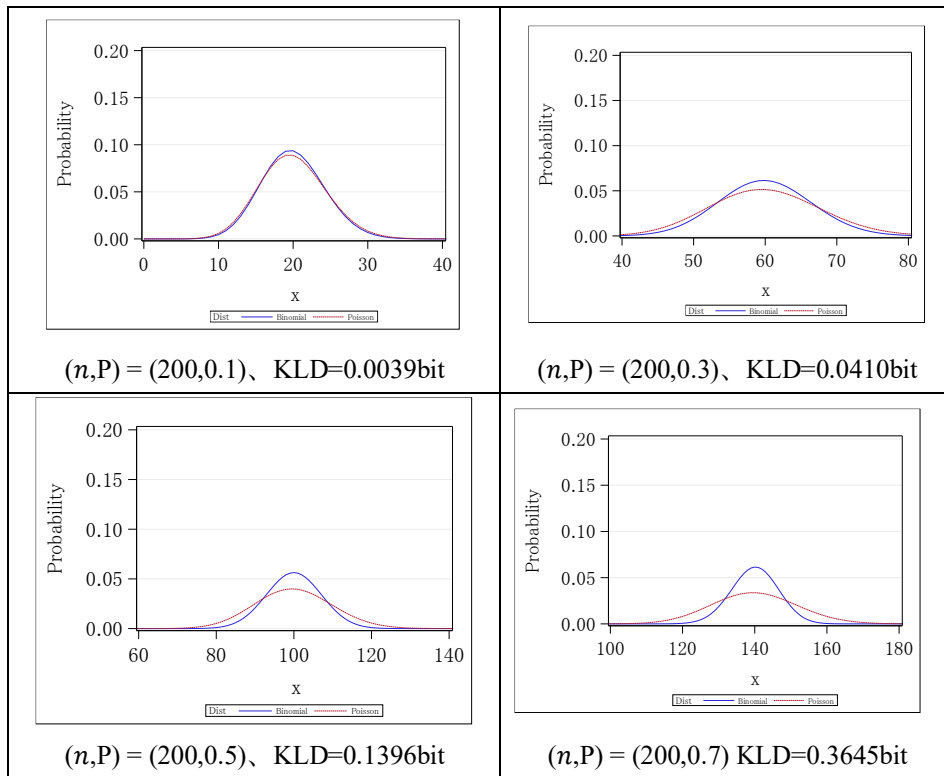


図 1. 二項分布と Poisson 分布および両者間の KLD 値

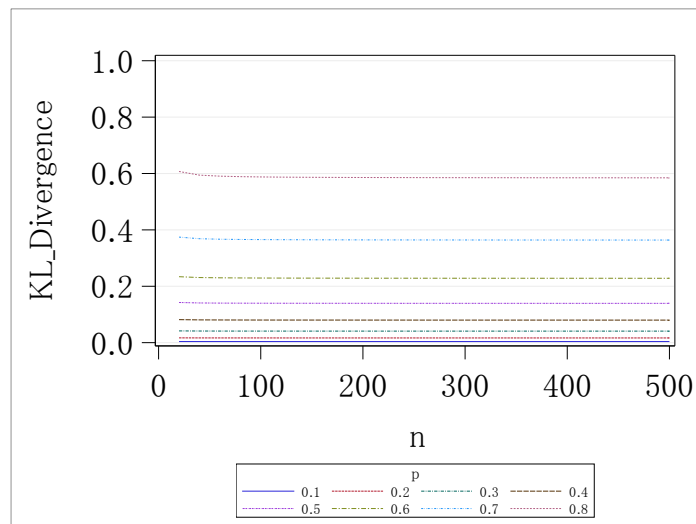


図 2. n の変化に伴う KLD 値の挙動 : P=0.1~0.8

3. 条件付き法と無条件法の比較

前節で述べた条件付き法は、両群の総イベント数 T への条件付けのもとでワクチン群のイベント発現数が従う条件付き二項分布に基づいて、必要イベント数を算出するものであった。一方、そのような条件付けを行わず、必要被験者数を算出する方法も多く用いられている。このように、総イベント数 T についての条件付けを行わず、(必要イベント数ではなく)必要被験者数を算出する方法を「無条件法」と呼ぶことにする。ここでは、無条件法の一例として、Farrington & Manning(1990)のスコア検定に基づく、リスク比についての仮説検定のための必要標本サイズの算出法をとりあげる。この方法は、SAS PROC POWER(V15.1)を用いて、以下のコードにより実行でき、使用にあたってはリスク比についての帰無仮説値、対立仮説の期待値、およびプラセボ群のイベント発現割合の想定値を指定することが必要である。

```
proc power;
  twosamplefreq test=fm_rr
  nullrelativerisk = 0.9 /* リスク比についての帰無仮説値 :  $RR_0$  */
  relativerisk = 0.6 /* リスク比についての対立仮説の期待値 :  $RR_1$  */
  refproportion = 0.05 /* プラセボ群のイベント発現割合の想定値 :  $P_p$  */
  sides = L ntotal = . alpha=0.025 power = 0.9;
run;
```

ここでは、リスク比について、帰無仮説値が $RR_0=0.9$ 、対立仮説の期待値が $RR_1=0.6$ である仮説検定で、2群の発現割合が、リスク比についての対立仮説の期待値を満たす3通りのケースを取り上げ、条件付き法で①必要総イベント数、無条件法で③必要総被験者数を算出した結果を表3に示した。この表では、条件付き法による②期待必要総被験者数(①必要総イベント数から算出)、無条件法による④期待総イベント数(③必要総被験者数から算出)も示している。ここで、「期待」の語を冠している理由は、2群の発現割合が正しく想定されていることを仮定したもとの期待値であるためである。表3から、条件付き法による②期待総被験者数が無条件法による③必要総被験者数よりも大きな例数で算出されることがわかる。このことは、Loiacono (2018)の Discussion での記述、すなわち「条件付き法で算出される必要標本サイズが大きなものとなり易い」と一致している。

表3. 条件付き法による①必要総イベント数、無条件法による③必要総被験者の算出結果

		帰無仮説値	対立仮説の 期待値	条件付き法		無条件法	
				①必要総 イベント 数	②期待必 要総被験 者数	③必要総 被験者数	④期待総 イベント 数
リスク比 (ワクチン有効性)		$RR_0=0.9$ ($\pi_0=0.1$)	$RR_1=0.6$ ($\pi_1=0.4$)	279			
発 現 割 合 の 想 定 値	ケース 1 ワクチン群 プラセボ群	$P_{v0}=0.045$ $P_p=0.050$	$P_{v1}=0.030$ $P_p=0.050$	279	6,975	6,402	257
	ケース 2 ワクチン群 プラセボ群	$P_{v0}=0.0225$ $P_p=0.025$	$P_{v1}=0.015$ $P_p=0.025$	279	13,950	13,070	262
	ケース 3 ワクチン群 プラセボ群	$P_{v0}=0.009$ $P_p=0.010$	$P_{v1}=0.006$ $P_p=0.010$	279	34,875	33,078	265

ところで、より重要なことは、条件付き法では①必要総イベント数が、リスク比についての帰無仮説値と対立仮説の期待値によって一意に算出され、3通りのケースのすべてで279件となっていることである。一方、無条件法による③必要総被験者数の算出では、リスク比についての帰無仮説値、対立仮説の期待値に加えて、プラセボ群の発現割合を指定することが必要であり、その発現割合の想定値が実際の値から乖離すれば、得られる③必要総被験者数が大きく変動するという難点が伴う。具体的に、3通りのケースで③必要総被験者数が大きく相違することがわかる。同様の不確定性は、条件付き法による②期待総被験者数、無条件法による④期待総イベント数の算出においても起こり得る。

以上のことから、必要総イベント数は、プラセボ群の発現割合の想定値の妥当性に左右されないという意味で確定的な目標値となり得る。そして、試験実施中にそれまでの時点で起こった総イベント数の、目標値に対する到達度によって試験の進捗管理を行えることが、条件付き法で必要総イベント数を算出することの意義であるといえる。

このことを表す例示として、2つの状況についてのシミュレーション結果を示す。ここでは、表3におけるリスク比についての対立仮説の期待値の通りのデータが得られる状況を設定した。

1つ目は、表3におけるケース1での発現割合の想定値（すなわちワクチン群の発現割合が $P_{v1}=0.030$ 、プラセボ群の発現割合が $P_p=0.050$ ）の通りの状況である。リスク比は0.6であり、条件付き法での必要総イベント数は279件である。このもとで2群合計6,976例（1群あたり3,488例）について、イベントの有無の二項乱数を生成し、得られたデータについて、リスク比についての帰無仮説 $RR \geq 0.9$ 、対立仮説 $RR < 0.9$ の片側有意水準0.025の検定を、Farrington & Manning(1990)のスコア検定で実施した。このことを100,000回反復したときの2群のイベント数合計の割合のヒストグラムと、検定結果が有意であった割合、すなわち検出力の折れ線グラフを図3-aに示した。なお、グラフの横軸の1つの級の幅を10としている。例えば、総イベント数 T_{class} が270の級は $265 < \text{総イベント数} \leq 275$ である。この図から、総イベント数が270以上の T_{class} に含まれる割合が大きく（正確には80.0%）、その場合に90%以上の検出力が得られていることがわかる。

2つ目は、両群の発現割合がケース1における想定値よりも小さい状況であり、ワクチン群の発現割合が $P_{v1}=0.024$ 、プラセボ群の発現割合が $P_p=0.040$ である。ただし、リスク比は0.6で上記と同じとした。そして、図3-aの場合と同じ2群合計6,976例（1群あたり3,488例）のデータを生成した。その結果を図3-bに示す。この図から、総イベント数が270以上の T_{class} に含まれる割合が相当に小さいこと（正確には1.95%）、その場合でも90%以上の検出力が得られていることがわかる。

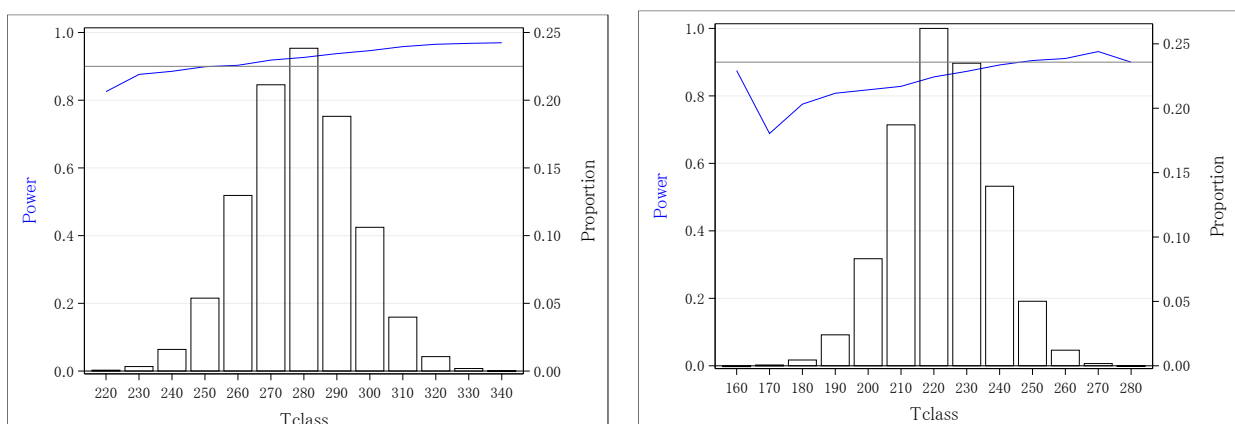


図3. イベント数合計のヒストグラムとリスク比の検定の検出力

- (左図)両群の発現割合が想定通りの場合： $P_{v1}=0.030$ 、 $P_p=0.050$ 、リスク比=0.6
- (右図)両群の発現割合が想定より小さい場合： $P_{v1}=0.024$ 、 $P_p=0.040$ 、リスク比=0.6

以上より、目標総イベント数（上記シミュレーションでは279）を目標値として試験の進捗管理を行うこと
 によって、2群のリスク比に関する仮説検定で目標検出力の達成可能性を高い確度で予測することが可能で
 あることがうかがえる。

ただし、2群の実際のリスク比が対立仮説の期待値とは相違する場合には、目標総イベント数を頼りにす
 る試験の進捗管理は不確かなものになることには注意が要る。図 3-b で示した 2 群の発現割合が想定よりも
 小さい場合も含め、開鍵を伴う中間解析によって 2 群の実際の発現割合を確認し、目標総イベント数の見直
 しを含めた試験のアダプテーションを念頭に置くことが必要であろう。

4. ワクチン有効性の推定の観点からの必要イベント数の決定

4.1 ワクチン有効性の推定方法

2群の総イベント数 T のうち、ワクチン群のイベント数 Y_v の観測値が $y_{v,obs}$ であるとき、ケース・プロポーシ
 ョン θ の推定値は $\hat{\theta} = y_{v,obs}/T$ であるが、Senn(2022)は θ の95%信頼区間を次のように与えている。

$$95\% \text{ 信頼下限値 } \theta_{LCL} : \Pr(Y_v \leq y_{v,obs} | \theta) = 0.025 \text{ を満たす } \theta$$

$$95\% \text{ 信頼上限値 } \theta_{UCL} : \Pr(Y_v \geq y_{v,obs} | \theta) = 0.025 \text{ を満たす } \theta$$

この信頼区間を数値的に算出するには、 $\theta = 0.0001, 0.0002, \dots, 0.9999$ (0.0001 刻み)で、二項変数 $Y_v \sim \text{Bin}(T, \theta)$
 を用いて上記の確率を評価すればよい。 $\hat{\theta}$ 、 θ_{LCL} および θ_{UCL} のそれぞれからワクチン有効性 π の推定値 $\hat{\pi}$ 、上限
 値 π_{UCL} および下限値 π_{LCL} を次のように算出する。

$$\hat{\pi} = (2\hat{\theta} - 1) / (\hat{\theta} - 1)$$

$$\pi_{UCL} = (2\theta_{LCL} - 1) / (\theta_{LCL} - 1)$$

$$\pi_{LCL} = (2\theta_{UCL} - 1) / (\theta_{UCL} - 1)$$

表 4. 総イベント数が $T = 3000$ の場合のワクチン群のイベント数 $y_{v,obs}$ の各値に対応する
 ケース・プロポーシオンおよびワクチン有効性の95%信頼区間

$y_{v,obs}$	$\hat{\theta}$	θ_{LCL}	θ_{UCL}	$\hat{\pi}$	π_{LCL}	π_{UCL}	$H_0: \pi \leq \pi_0$ に対する p 値		
							$\pi_0 = 0$	$\pi_0 = 0.1$	$\pi_0 = 0.2$
1,278	0.4260	0.4082	0.4439	0.2578	0.2018	0.3102	0.0000	0.0000	0.0218
1,279	0.4263	0.4085	0.4443	0.2568	0.2005	0.3094	0.0000	0.0000	0.0238
1,280	0.4267	0.4089	0.4446	0.2558	0.1995	0.3082	0.0000	0.0000	0.0260
:	:	:	:	:	:	:	:	:	:
1,365	0.4550	0.4371	0.4730	0.1651	0.1025	0.2235	0.0000	0.0211	0.8813
1,366	0.4553	0.4374	0.4734	0.1640	0.1010	0.2225	0.0000	0.0230	0.8884
1,367	0.4557	0.4377	0.4737	0.1629	0.0999	0.2216	0.0000	0.0250	0.8953
:	:	:	:	:	:	:	:	:	:
1,444	0.4813	0.4633	0.4994	0.0720	0.0024	0.1368	0.0213	0.8044	1.0000
1,445	0.4817	0.4636	0.4997	0.0707	0.0012	0.1357	0.0233	0.8143	1.0000
1,446	0.4820	0.4640	0.5001	0.0695	-0.0004	0.1343	0.0254	0.8240	1.0000
:	:	:	:	:	:	:	:	:	:

$y_{v,obs}$: ワクチン群イベント数、 $\hat{\theta}, \theta_{LCL}, \theta_{UCL}$: ケース・プロポーシオンの推定値、95%信頼下限値、
 上限値、 $\hat{\pi}, \pi_{LCL}, \pi_{UCL}$: ワクチン有効性の推定値、95%信頼下限値、上限値

例えば、総イベント数が $T = 3,000$ の場合のワクチン群のイベント数 $y_{v,obs}$ の各値に対応するケース・プロポ
 ーションおよびワクチン有効性の95%信頼区間を表 4 に示す。この表には、ワクチン有効性に関する仮説
 $H_0: \pi \leq \pi_0$ vs $H_1: \pi > \pi_0$ について、 $\pi_0 = 0$ 、 $\pi_0 = 0.1$ および $\pi_0 = 0.2$ の 3 通りの検定における p 値も示してい
 る。検定の p 値が片側有意水準である $\alpha (= 0.025)$ 未満か α 以上かの識別と、ワクチン有効性の95%信頼下限値

の π_{LCL} が帰無仮説値である π_0 を超えるか π_0 以下かの識別とが、寸分違わず一致していることを確認できる。したがって、Senn(2022)が提示したケース・プロポーシヨンの95%信頼区間は、信頼区間が具備すべき要件を備えていると判断できる。

4.2 ワクチン有効性の推定精度に基づく必要イベント数の設定方法の一案

ここでは、所定のワクチン有効性が期待される状況で、ワクチン有効性の95%信頼区間長を所定の限界値未満とするのに必要な総イベント数の算出を考える。ワクチン有効性の期待値を π_1 とすると、対応するケース・プロポーシヨンは、 $\theta_1 = (1 - \pi_1)/(2 - \pi_1)$ である。総イベント数 T のもとで、ワクチン群のイベント数 Y_v が二項分布 $\text{Bin}(T, \theta_1)$ に従って変動し、それに伴ってワクチン有効性の95%信頼区間も変動するため、その信頼区間を $[\pi_{LCL}(Y_v|T, \theta_1), \pi_{UCL}(Y_v|T, \theta_1)]$ で表す。このとき95%信頼区間の半分幅の期待値が目標値(h とする)未満であること、すなわち

$$\{E[\pi_{UCL}(T, \theta_1)] - E[\pi_{LCL}(T, \theta_1)]\}/2 < h \quad (9)$$

となる T の最小値を必要な総イベント数 T^* とする。 T^* の算出手順は次の通りである。

- ① T の初期値を T_0 とする。
- ② 起こり得る Y_v の値 y_v ($0, 1, \dots, T$)のそれぞれで、ケース・プロポーシヨンの95%信頼区間の下限値 $\theta_{LCL}(y_v)$ と上限値 $\theta_{UCL}(y_v)$ を算出する。前項のSenn(2022)の方法を用いる。
- ③ ケース・プロポーシヨンの95%信頼区間の下限値 $\theta_{LCL}(Y_v)$ と上限値 $\theta_{UCL}(Y_v)$ の各期待値を算出する。
 - ・ $E[\theta_{LCL}|T, \theta_1] = \sum_{y_v=0}^T \theta_{LCL}(y_v) \binom{T}{y_v} \theta_1^{y_v} (1 - \theta_1)^{T-y_v}$
 - ・ $E[\theta_{UCL}|T, \theta_1] = \sum_{y_v=0}^T \theta_{UCL}(y_v) \binom{T}{y_v} \theta_1^{y_v} (1 - \theta_1)^{T-y_v}$
- ④ ワクチン有効性の95%信頼区間の下限値と上限値のそれぞれの期待値を算出する。
 - ・ $E[\pi_{LCL}|T, \theta_1] = \frac{2 \times E[\theta_{UCL}|T, \theta_1] - 1}{E[\theta_{UCL}|T, \theta_1] - 1}$
 - ・ $E[\pi_{UCL}|T, \theta_1] = \frac{2 \times E[\theta_{LCL}|T, \theta_1] - 1}{E[\theta_{LCL}|T, \theta_1] - 1}$
- ⑤ 判定： $\{E[\pi_{UCL}(T, \theta_1)] - E[\pi_{LCL}(T, \theta_1)]\}/2 < h$ であれば、 T を必要総イベント数 T^* とし、手順を終了する。そうでなければ、 $T = T + 1$ へ更新し、手順(2)へ戻る。

上記手順を用い、0.1 から 0.5 まで (0.1 刻み) の π_1 について、95%信頼区間の半分幅の期待値を 0.1 未満とするための必要総イベント数を算出し、表 5 に示した。

表 5. ワクチン有効性の95%信頼区間の半分幅の期待値を 0.1 未満とするための必要総イベント数

ワクチン有効性の期待値 π_1	必要総イベント数 T^*	ワクチン有効性の95%信頼区間の期待値	信頼区間の半分幅
0.1	1,289	[-0.005427, 0.19451]	0.09997
0.2	1,032	[0.093953, 0.29394]	0.09999
0.3	810	[0.19324, 0.39319]	0.09998
0.4	620	[0.29234, 0.49218]	0.09992
0.5	458	[0.39093, 0.5908]	0.09998

5. 最後に

本稿では、ワクチン群とプラセボ群の2群比較試験の計画において、2群のイベント発現総数に条件付きでワクチン群のイベント発現数が従う二項分布に基づいて、ワクチン有効性についての仮説検定で必要なイベント発現数を算出する方法(条件付き法)に関する検討結果を記述した。本稿では、条件付き法の適用が妥当であるための条件を、両群の試験期間中の発現割合が0.3以下であることと設定した。必要被験者数ではなく、必要イベント数に注目したのは、イベント数による試験の進捗管理が行いやすいためである。そのこ

とを、総イベント数についての条件付けを行わず必要被験者数を算出する方法との比較に基づいて示した。さらに、ワクチン有効性の推定において、Senn(2022)によるケース・プロポーションの95%信頼区間算出法が優れた性質をもつことを示したうえで、ワクチン有効性の推定精度の観点から必要イベント数を算出するための一案を提示した。

本稿では、1対1の割付比の2群試験を対象としたが、1対1以外の割付比の2群試験における条件付き法の定式化がChan & Bohidar(1988)で与えられている。また、条件付き法は、プラセボ対照試験だけではなく、実薬対照の非劣性試験でも利用可能である(江口・浦狩,2025)。具体的に、対照薬を基準とするリスク比を $RR'=(\text{被験薬での発現割合})/(\text{対照薬での発現割合})$ と定義し、非劣性検定の帰無仮説を $RR' \geq r_0$ 、対立仮説を $RR' < r_0$ と与える。ここで、 r_0 はリスク比の尺度上の非劣性マージンである。対照薬を基準とするワクチン有効性を $\pi'=(1-RR')$ とおけば、非劣性検定の帰無仮説が $\pi' \leq \pi'_0$ 、対立仮説が $\pi' > \pi'_0$ で表せる($\pi'_0=1-r_0$ はワクチン有効性の尺度上の非劣性マージン)。例えば、リスク比 RR' について非劣性マージンが2.0と設定され、被験薬と対照薬の発現割合が等しいことを仮定して行う非劣性検定は、帰無仮説値を $r_0 = 2.0$ 、対立仮説の期待値を $r_1 = 1.0$ とする検定となる。そして、この検定は対照薬を基準とするワクチン有効性 π' について、帰無仮説値を $\pi'_0 = -1.0$ 、対立仮説の期待値を $\pi'_1 = 0$ とする検定と同等である。この非劣性検定において、片側有意水準0.025で90%の検出力を確保するのに必要な総イベント数は275件と算出される。

参考文献

- Chan, I.S.F. and Bohidar, N.R. (1998). Exact power and sample size for vaccine efficacy studies. *Communications in Statistics-Theory and Methods*, 27(6), 1305-1322.
- Gu, K., Ng, H.K.T., Tang, M.L., and Schucany, W. (2008) Testing the ratio of two Poisson rates. *Biometrical Journal*, 50(2), 283-298.
- Farrington, C. P., and Manning, G. (1990). Test statistics and sample size formulae for comparative Binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. *Statistics in Medicine*, 9, 1447-1454.
- Loiacono, M.L. (2018). Sample size estimation and power calculations for vaccine efficacy trials for exceedingly rare diseases. (This thesis was approved by Hanna Bandos, PhD, Robert Krafty, PhD and Willem G. van Panhuis, MD PhD in University of Pittsburgh).
- Senn, S. (2022). The design and analysis of vaccine trials for COVID-19 for the purpose of estimating efficacy. *Pharmaceutical Statistics*, 21, 790-807.
- Przyborowski, J. and Wilenski, H. (1940). Homogeneity of results in testing samples from Poisson series: with an application to testing clover seed for dodder. *Biometrika*, 31, 313-323.
- 伊藤要二 (2016). ポアソン分布に従うと仮定された総ケース数が固定された条件の下での rate ratio の条件付き信頼区間. SAS ユーザー総会・報告資料.
- 江口凌平・浦狩保則 (2025). ワクチン開発における非劣性試験のデザインについて. SAS ユーザー総会・報告資料.
- 杉山 将 (2013). 確率分布間の距離推定：機械学習分野における最新動向. *日本応用数理学会論文誌*, 23(3), 439-452.