

ON SAMPLE SIZE DETERMINATION BASED ON A BAYESIAN INDEX OF NON-INFERIORITY FOR TWO INDEPENDENT BINOMIAL PROPORTIONS

YOHEI KAWASAKI and HIROSHI YAMADA

Department of Drug Evaluation and Informatics

School of Pharmaceutical Sciences

University of Shizuoka

52-1 Yada, Suruga-ku

Shizuoka, 422-8526

Japan

e-mail: ykawasaki@u-shizuoka-ken.ac.jp

Abstract

In drug development, non-inferiority tests are often employed to determine the difference between two independent binomial proportions. Many test statistics for non-inferiority are based on the frequentist framework. In this paper, we propose a method for the determination of sample sizes using the index $\tau = P(\pi_1 > \pi_2 - \Delta_0 | X_1, X_2)$, where X_1 and X_2 denote binomial random variables for trials n_1 and n_2 , and parameters π_1 and π_2 , respectively, and the non-inferiority margin is $\Delta_0 > 0$. In this paper, we propose a new method of calculation, by which we determine in advance the estimated difference between the posterior proportions and the value to be derived at the end of the test to obtain the necessary sample size. We provide the lists of necessary sample sizes with various assumptions.

2010 Mathematics Subject Classification: 62F15.

Keywords and phrases: independent binomial proportions, Bayesian inference, non-inferiority, sample size.

Received November 19, 2015

© 2015 Scientific Advances Publishers

1. Introduction

Non-inferiority tests are often conducted in clinical trials, primarily to determine whether the response to the study drugs is clinically not much worse than the response to the reference (control) drugs. In particular, non-inferiority tests are employed to derive the difference between two binomial proportions if the response is an independent binomial. The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH-E9) guidelines [1] and the European Medicines Agency guidelines [2] provide the framework for setting non-inferiority comparisons between treatment groups.

Biostatisticians determine the main analysis method, the estimated difference between binomial proportions, the significance level, non-inferiority margin, and the power before clinical study. This is the method of calculation for sample size using the frequentist framework.

A large body of literature exists on the method of calculation for sample size, including both frequentist (Lachin [12]; Lemeshow et al. [13]) and Bayesian approaches. Several criteria have been proposed for Bayesian sample size determination. The particular application to the binomial parameter was examined in detail in Pham-Gia and Turkkan [16], while Adcock [3] considered multinomial experiments, which include the binomial as a special case. Adcock [4] compared the various approaches presented in the above two papers. Joseph et al. [8] proposed three different Bayesian approaches to calculating sample size. These proposed methods are based on Highest Posterior Density (HPD)-credible intervals are discussed and illustrated in the context of a binomial experiment. Decision-theoretic criteria (Lindley [14]) and sample sizes based on average power of hypothesis tests (Spiegelhalter and Freedman [17]) have also been considered. See Chaloner and Verdinelli [7] for a recent review of Bayesian optimal design and Adcock [5] for a review of both frequentist and Bayesian sample size criteria. More recently, M'Lan

et al. [15] investigated the binomial sample size problem by using generalized versions of the Average Length and Average Coverage Criteria, the Median Length and Median Coverage Criteria, and the Worst Outcome Criterion and its modified version. These methods of calculation for sample size are based on HPD-credible interval.

Berry [6] detailed a Bayesian approach to comparing the binomial proportions of two groups using some examples. Zaslavsky [18] proposed a one sided hypothesis based on a one sample situation. Kawasaki and Yamada [11] proposed sample size calculation method for a superiority situation of binomial proportions. Kawasaki and Miyaoka [10] proposed an index

$$\tau = P(\pi_1 > \pi_2 - \Delta_0 | X_1, X_2),$$

where X_1 and X_2 denote binomial random variables for trials n_1 and n_2 , and parameters π_1 and π_2 , respectively, and the non-inferiority margin is $\Delta_0 > 0$. They provided approximate and exact expressions for index τ by using the beta prior and presented the results of actual clinical trials to show the utility of this index.

In this paper, we propose a new method of calculation for sample size using index τ . For this method, the difference between two independent posterior binomial proportions and index τ are decided in advance of clinical research, and the calculation result of the sample size required in that case is shown. Moreover, we compare the sample size using approximate and exact methods for index τ .

The remainder of this paper is organized as follows. We show the approximate and exact expressions for index τ in Section 2. In Section 3, we propose a method of calculation for sample size using τ , and we provide some lists of necessary sample sizes through simulation results with various assumptions in Section 4. In Section 5, we consider an imaginary clinical trial of a new treatment and calculate the value of sample sizes using the lists of necessary sample size. Finally, we conclude the paper with a brief summary in Section 6.

2. Methodology

Let X_1 and X_2 denote binomial random variables for n_1 and n_2 trials with parameters π_1 and π_2 , respectively. The conjugate prior density for π_i is a beta distribution with parameters α_i and β_i , where $\alpha_i > 0$, $\beta_i > 0$, and $i = 1, 2$. The proposed posterior density for π_i is

$$g_i(\pi_i|x_i) = \frac{1}{B(a_i, b_i)} \pi_i^{a_i-1} (1 - \pi_i)^{b_i-1},$$

where $a_i = \alpha_i + x_i$, $b_i = n_i - x_i + \beta_i$, and $B(a, b)$ is the proposed beta function.

2.1. Approximate expression for τ

One method of calculation for the index τ is an approximation using the standard normal table. The new index τ can be calculated via an approximation using the standard normal table. We assume that a_i and b_i of the posterior density are large. The approximate expression for τ is given by

$$\begin{aligned} \tau &= P(\pi_1 > \pi_2 - \Delta_0 | X_1, X_2) \\ &\approx 1 - \Phi \left(\frac{-\Delta_0 - (\mu_{1,post} - \mu_{2,post})}{\sqrt{\frac{\mu_{1,post}(1 - \mu_{1,post})}{a_1 + b_1 + 1} + \frac{\mu_{2,post}(1 - \mu_{2,post})}{a_2 + b_2 + 1}}} \right), \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, and $\mu_{i,post} = a_i / (a_i + b_i)$ denotes the posterior mean of π_i .

2.2. Exact expression for τ

Additionally, we can calculate the new index τ by using the exact posterior PDF. The exact expression for τ is given by

$$\tau = \int_{-\Delta_0}^1 f(\delta) d\delta,$$

where $f(\delta)$ for δ is the exact posterior PDF for $\pi_1 - \pi_2$. The exact posterior PDF is expressed as Kawasaki and Miyaoka [9].

3. Method of Calculation for Sample Size

We propose a method of calculation for sample size using τ in this section. The detailed flow of the method is as follows:

(1) Determine the value of hyperparameters $\alpha_1, \beta_1, \alpha_2$, and β_2 in beta distributions that are prior distributions and non-inferiority-margin is $\Delta_0 > 0$.

(2) Determine the value of τ , which is the difference in the posterior densities of the two groups

$$\delta = E(\pi_{1,post} - \pi_{2,post} + \Delta_0) = \frac{a_1}{a_1 + b_1} - \frac{a_2}{a_2 + b_2} + \Delta_0.$$

(3) Determine the value of τ_{\min} , which is the lower limit of τ .

(4) Determine the ratio r between the sample size of the two groups (where $n_1 = r * n_2$).

(5) Based on the above assumptions (1) to (4), set to $n_2 = 1$ and calculate the following:

(6) Find the combinations of the realized values x_1 and x_2 of the random variables X_1 and X_2 , which are satisfied by the condition of δ in step (2).

- (7) Calculate the value of τ for all combinations found in step (6).
- (8) Update the value of n_2 by $n_2 + 1$ and go back to step (6) (iterate this calculation an appropriate number of times).
- (9) Find the value of the minimum n_1 for which all obtained values of τ in step (7) exceed τ_{\min} , and determine that the necessary sample size for the treatment group is this value. The necessary sample size for the control group can be defined by the relative expression $n_1 = r * n_2$.

4. Simulation Results

In this section, we simulate the necessary sample size using τ , which was defined in the previous section. Each assumption and setting is as follows: the non-informative prior is $(\alpha_1 = \beta_1 = \alpha_2 = \beta_2 = 1)$, the expected difference between the posterior densities of the two groups is $\delta = 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.11, 0.12, 0.13, 0.14, 0.15$, the lower limit of τ is $\tau_{\min} = 0.7, 0.75, 0.8, 0.85, 0.9, 0.95$, and the ratio of sample sizes between the treatment and control groups is $r = 1, 2, 3$. In addition, the necessary sample size under other conditions is easily calculated by simulation, although we do not include this in the present paper.

4.1. Simulation results when $n_1 = n_2$ and $\Delta_0 = 0.1$

In this subsection, we displayed two tables in Table 1; one is calculated based on the exact probability of τ and the other, on the approximate probability.

We display the detailed results for sample size in Table 1. The vertical column shows the value of τ , and the horizontal row shows the difference between posterior proportions. We can see at several points that the necessary sample size for exact calculation is slightly higher than for approximation.

Table 1. The result of necessary sample size n_1 when $n_1 = n_2$

τ	The difference between posterior proportions ($\Delta_0 = 0.1$)										
	0.05	0.06	0.07	0.08	0.09	0.1	0.11	0.12	0.13	0.14	0.15
70%	65	45	25	22	20	18	16	6	13	5	11
75%	108	75	52	34	30	27	25	14	13	12	11
80%	155	108	78	56	51	37	33	23	21	20	18
85%	265	163	123	93	72	56	42	38	36	26	24
90%	397	253	185	132	103	83	68	54	50	40	37
95%	665	435	306	239	175	146	113	94	80	67	63

(a) Calculation based on exact method.

τ	The difference between posterior proportions ($\Delta_0 = 0.1$)										
	0.05	0.06	0.07	0.08	0.09	0.1	0.11	0.12	0.13	0.14	0.15
70%	65	45	25	22	20	18	16	6	13	5	11
75%	108	75	52	34	30	27	25	14	13	12	11
80%	155	108	78	56	51	37	33	23	21	20	18
85%	265	163	123	93	72	56	42	38	36	26	24
90%	397	253	185	132	103	83	68	54	50	40	37
95%	665	435	306	239	175	146	113	94	80	67	63

(b) Calculation based on approximate method.

4.2. Simulation results when $n_1 \neq n_2$ and $\Delta_0 = 0.1$

Table 2 shows the detailed results. These tables show that the necessary sample size $n_1 + n_2$ is larger in this case than when sample sizes are equal among groups. We find from the following result that this trend increases as the imbalance between n_1 and n_2 becomes larger. In addition, Table 3 show the result when $n_1 = 3 * n_2$. Those also show that when sample sizes among groups are imbalanced the necessary sample size becomes larger. We also find that as the imbalance between n_1 and n_2 becomes larger, the difference between the necessary sample sizes for exact and approximate calculation tends to become larger.

Table 2. The result of necessary sample size n_1 when $n_1 = 2 * n_2$

τ	The difference between posterior proportions ($\Delta_0 = 0.1$)										
	0.05	0.06	0.07	0.08	0.09	0.1	0.11	0.12	0.13	0.14	0.15
70%	92	58	46	30	24	20	10	10	8	12	6
75%	158	108	76	54	36	32	22	18	16	14	10
80%	246	166	122	82	68	54	40	34	24	22	20
85%	398	256	184	134	106	84	66	52	46	36	34
90%	600	400	278	212	164	126	105	86	72	58	54
95%	1000	656	476	346	270	220	180	146	120	104	88

(a) Calculation based on exact method.

τ	The difference between posterior proportions ($\Delta_0 = 0.1$)										
	0.05	0.06	0.07	0.08	0.09	0.1	0.11	0.12	0.13	0.14	0.15
70%	92	58	44	26	20	14	10	10	8	12	6
75%	158	108	76	54	36	32	18	16	14	12	10
80%	246	166	110	82	60	52	40	34	24	22	20
85%	398	256	184	134	106	84	66	52	42	36	32
90%	600	400	278	212	164	126	104	86	72	58	52
95%	1000	656	476	348	270	220	172	146	120	102	88

(b) Calculation based on approximate method.

Table 3. The result of necessary sample size n_1 when $n_1 = 3 * n_2$

τ	The difference between posterior proportions ($\Delta_0 = 0.1$)										
	0.05	0.06	0.07	0.08	0.09	0.1	0.11	0.12	0.13	0.14	0.15
70%	129	81	54	42	30	24	9	12	15	12	9
75%	222	144	96	72	57	45	30	27	21	21	15
80%	339	222	156	120	87	72	57	42	36	33	21
85%	516	348	246	186	141	114	87	72	63	51	45
90%	804	531	384	282	222	177	141	114	96	81	69
95%	1335	888	630	468	366	294	237	198	168	141	123

(a) Calculation based on exact method.

τ	The difference between posterior proportions ($\Delta_0 = 0.1$)										
	0.05	0.06	0.07	0.08	0.09	0.1	0.11	0.12	0.13	0.14	0.15
70%	126	75	54	39	15	24	9	12	9	12	9
75%	207	144	96	69	48	39	30	27	21	21	15
80%	336	222	156	120	84	72	54	42	36	27	21
85%	516	348	246	177	141	111	87	69	57	51	39
90%	801	528	372	282	213	168	141	114	96	81	69
95%	1332	876	630	468	366	285	237	192	159	141	117

(b) Calculation based on approximate method.

5. Example

In this section, we provide examples of the setting method for sample size for a case in which we consider an imaginary clinical trial of a new treatment based on the chart in Section 3. Suppose the following situation. Assume a non-informative prior as the prior distribution ($\alpha_1 = \beta_1 = \alpha_2 = \beta_2 = 1$) and non-inferiority margin is $\Delta_0 = 0.1$. Set the same sample size for the treatment and control groups ($n_1 = n_2$). Assume that the expected value of difference of the posterior density between the two groups is 0.06. Under the above condition, from Table 1, if the lower limit of τ is set to 0.90, the necessary sample size for the

treatment and new groups is 400 by both exact and approximate calculation. If the lower limit is set to 0.95, the necessary sample size becomes 656 by exact calculation, but 654 by approximation.

6. Conclusion

In this paper, we suggested a new calculation method to obtain the sample size for the index of binomial proportion $\tau = P(\pi_1 > \pi_2 - \Delta_0 | X_1, X_2)$ suggested by Kawasaki and Miyaoka [10] and simulated the actual value. We also defined a method of calculation for τ based on both approximation and exact methods and compared the results.

The method to define sample size suggested in this paper begins with setting the estimated difference between the posterior proportions and the lower limit of τ . In addition, by setting the proportion of the sample sizes of the treatment and control groups and adding the information (if it is available), the sample size can be derived.

We found three important results regarding τ derived by simulation for both exact calculation and approximation. First, when the sample size becomes larger, τ increases, but not monotonously. Second, as the estimated difference between the posterior proportions becomes larger, the necessary sample size decreases. Third, when the difference between the proportions of the treatment and control groups becomes larger, the necessary sample size for both groups becomes larger.

In addition, comparison of the exact and approximate values in the simulation yields the following result. Exact calculation and approximation require almost the same sample size, and the trend for τ to increase when the sample size becomes larger is almost the same. The sample size for exact calculation is slightly larger than that for approximation. This difference tends to increase as the imbalance between the treatment group and the control group increases.

Finally, index τ , which has been extensively investigated in this research, is derived using the Bayesian framework rather than the frequentist framework as before. One of the merits of using the Bayesian framework is that it is easy to understand; another is that calculation with advanced information is possible. Therefore, the index is expected to deliver very important information on clinical trials, where setting the sample size is considered very important. The setting method suggested in this paper may serve as a useful guideline in this regard.

Acknowledgements

We are very grateful to the editor and two referees for their invaluable comments.

References

- [1] ICH E9 expert Working Group, ICH harmonized tripartite guideline, statistical principles for clinical trials, *Statistics in Medicine* 18 (1999), 1905-1942.
- [2] Efficacy Working Party, European Medicines Agency Committee for Medicinal Products for Human Use (CHMP) guideline on the choice of the non-inferiority margin, *Statistics in Medicine* 25 (2006), 1628-1638.
- [3] C. J. Adcock, A Bayesian approach to calculating sample sizes for multinomial sampling, *The Statistician* 36 (1987), 155-159.
- [4] C. J. Adcock, Bayesian approaches to the determination of sample sizes for binomial and multinomial sampling- some comments on the paper by Pham-Gia and Turkkan, *The Statistician* 41 (1992), 399-404.
- [5] C. J. Adcock, Sample size determination: A review, *The Statistician* 46(2) (1997), 261-283.
- [6] D. A. Berry, Bayesian approaches for comparative effectiveness research, *Clinical Trials* 9 (2012), 37-47.
- [7] K. Chaloner and I. Verdinelli, Bayesian experimental design: A review, *Statistical Science* 10(3) (1995), 273-304.
- [8] L. Joseph, D. B. Wolfson and R. Berger, Sample size calculations for binomial proportions via highest posterior density intervals, *The Statistician* 44(2) (1995), 143-154.
- [9] Y. Kawasaki and E. Miyaoka, A posterior density for the difference between two binomial proportions and the highest posterior density credible interval, *Journal of the Japan Statistical Society* 40(2) (2010), 265-275.

- [10] Y. Kawasaki and E. Miyaoka, A Bayesian non-inferiority test for two independent binomial proportions, *Pharmaceutical Statistics* 12(4) (2013), 201-206.
- [11] Y. Kawasaki and H. Yamada, On sample size determination based on a Bayesian index of superiority for two independent binomial proportions, *Journal of Modern Applied Statistical Methods* (2015) (in press).
- [12] J. M. Lachin, Introduction to sample size determination and power analysis for clinical trials, *Controlled Clinical Trials* 2 (1981), 93-113.
- [13] S. Lemeshow, Jr., D. W. Hosmer, J. Klar and S. K. Lwanga, *Adequacy of Sample Size in Health Studies*, Wiley, New York, 1990.
- [14] D. V. Lindley, The choice of sample size, *The Statistician* 46 (1997), 129-138.
- [15] C. E. M'LAN, L. Joseph and D. B. Wolfson, Bayesian sample size determination for binomial proportions, *Bayesian Analysis* 3(2) (2008), 269-296.
- [16] T. Pham-Gia and N. Turkkan, Sample size determination in Bayesian analysis, *The Statistician* 41 (1992), 389-397.
- [17] D. J. Spiegelhalter and L. S. Freedman, A predictive approach to selecting the size of a clinical trial, based on subjective clinical, *Statistics in Medicine* 5 (1986), 1-13.
- [18] B. G. Zaslavsky, Empirical Bayes models of Poisson clinical trials and sample size determination, *Pharmaceutical Statistics* 9 (2009), 133-141.

