Comparison of risk assessment approaches for adverse drug reactions in the post-marketing study.

2019年6月5日　統計数理研究所　オープンハウス

伊藤　陽一　　データ科学研究系　教授

INTRODUCTION:
In GPSP (Good Post-marketing Study Practice), the investigator must conduct drug use-results surveys according to the post-marketing survey protocol with the pre-planned sample size. In almost all post-marketing survey protocols in Japan, the sample size was determined by n=3/λ for an anticipated risk λ (Machin D. et. al 2009). However, this approach accommodates for detection of unknown ADR (Adverse Drug Reaction). In general, there are three objectives for conducting a drug use-results survey,

1) To detect unknown ADR
2) To evaluate the magnitude of risk of known ADRs
3) To evaluate the factors which affect efficacy and safety of the durg

In this study, we propose an sample size estimation method for determining the risk of known ADRs.

Objectives:
1) Propose an alternative approach of sample size estimation for determining the risk of ADR in the post-marketing study.
2) Investigate the required sample sizes in typical cases.

Proposed Sample Size Estimation Approach:
On the label of drugs, the risk of ADR was often classified into three categories; (a) <1%, 1% - 5%, >= 5% or (b) <0.1%, 0.1% - 1%, >=1%. It is clinically useful if the risk of ADR was clearly classified according to the above categories.
In our approach, when we use the category (a), the sample size is determined based on binomial distributions to meet the following conditions:
Pr(r < 1% | λ=5%) < 1% and Pr(r >= 5% | λ=1%) < 1%
where r is the observed risk, and λ is the true risk.
Pr(r < 1% | λ=5%) and Pr(r >= 5% | λ=1%) were called the classification error probability in our approach. A scheme of the classification error probability was shown in Figure 1.

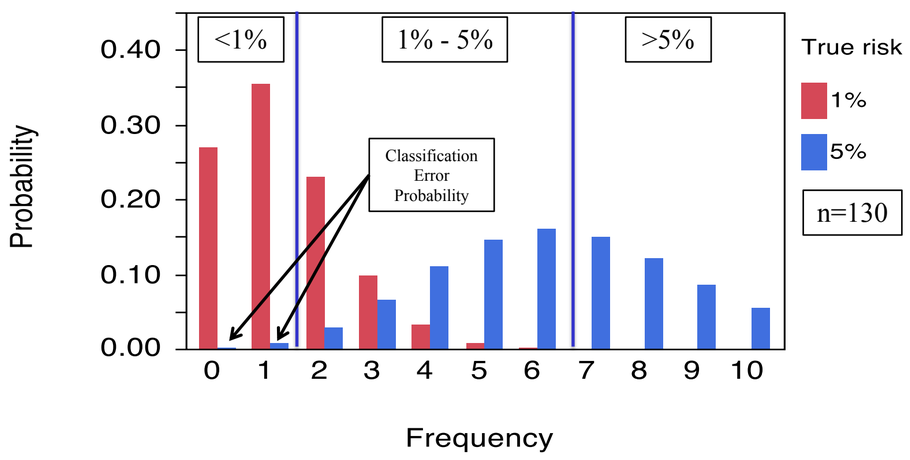

Figure 1: A Schema of the classification error probability

Pr(r < 1% | λ=5%) and Pr(r >= 5% | λ=1%) were calculated by the following formulas.

$$\Pr(r<1\%\mid\lambda=5\%)=\sum_{x=0}^{m}\binom{n}{x}0.05^{x}0.95^{n-x} \qquad m=INT(0.01\times n)$$

$$\Pr(r>5\%\mid\lambda=1\%)=1-\sum_{x=0}^{m}\binom{n}{x}0.01^{x}0.99^{n-x} \qquad m=INT(0.05\times n)$$

Applied example:
We applied our approach to the two cases of the category of (a) <1%, 1% - 5%, >= 5% and (b) <0.1%, 0.1% - 1%, >=1%.
For the category (a), CEP and required sample size n was shown in Figure 2. Pr(r < 1% | λ=5%) = 0.00997 and the required sample size was 130. And the results of the category (b) was shown in Figure 3. Pr(r < 0.1% | λ=1%) = 0.00992 and the sample size was 459.
Table 1 showed comparison of required sample sizes of our approach with that of other risk assessment approaches; to detect unknown ADR approach and test-based approach. On the test-based approach, we assumed a test of a binomial proportion using the normal approximation and a variance estimate based on the null proportion with a one-sided significance level of 0.05.
If true risk was 1%, required sample size of to detect unknown ADR approach was 300 and that of the test-based approach with power 95% was 562. On the other hand, our proposed approach with 1% upper threshold and 0.1% lower threshold was 459.

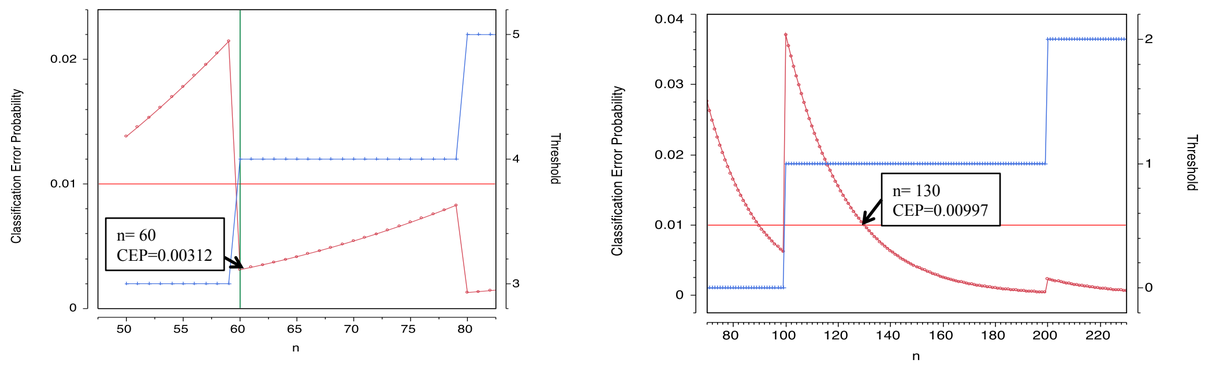

Figure 2:
Classification error probability (CEP) and threshold of number of ADR in the category of <1%, 1%-5%, >5%; Pr(r > 5% | λ=1%) = 0.00312, n=60 (left), and Pr(r < 1% | λ=5%) = 0.00997, n=130 (right). The required sample size was estimated 130.

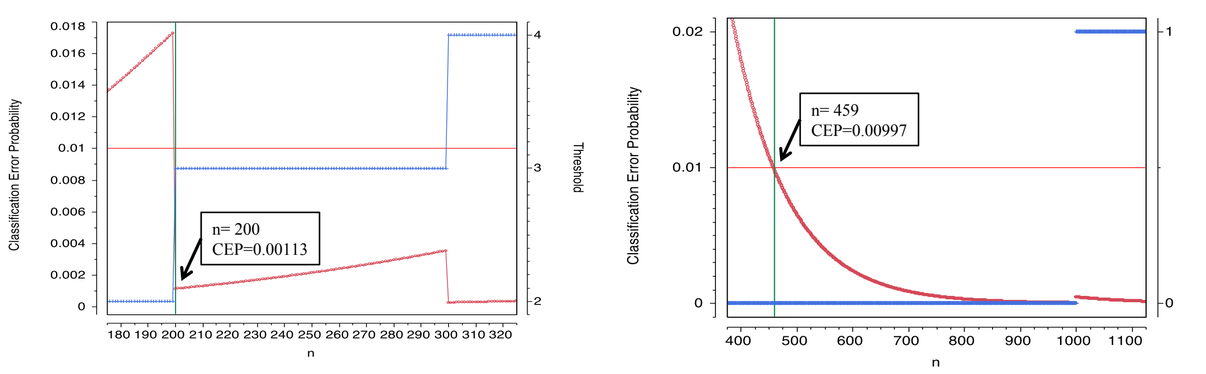

Figure 3:
Classification error probability (CEP) and threshold of number of ADR in the category of <0.1%, 0.1%-1%, >1%; Pr(r > 1% | λ=0.1%) = 0.00113, n=200 (left), and Pr(r < 0.1% | λ=1%) = 0.00997, n=459 (right). The required sample size was estimated 459.

Table 1: Comparison of risk assessment approaches

	Upper threshold	Lower threshold		Required Sample Size
Proposed approach	5%	1%		130
	1%	0.1%		459
	True risk			Required Sample Size
To detect unknown ADR approach	5%			60
	1%			300
	0.1%			3000
	True risk	Null Risk	Power	Required Sample Size
Test-based approach One-sided, α=0.05	5%	1%	80%	76
	5%	1%	95%	171
	1%	0.1%	80%	228
	1%	0.1%	95%	575

Discussion:
Our approach was focused on the control of classification error probability. This feature was clinically useful for discrimination of high risk and low risk.
To discriminate the risk of larger than 5% and the risk of less than 1%, the sample size of 130 was adequate. On the other hand, to discriminate the risk of larger than 1% and the risk of less than 0.1%, the sample size of 459 is needed. In the rare ADRs, the more sample size is needed to discrimination.
The required sample sizes of test-based approach with 1% true risk and 95% power was 575. And our approach with 1% upper threshold and 0.1% lower threshold was 459. These figures were larger than that of to detect unknown ADR approach with 1% true risk, and smaller than that with 0.1% true risk. This means that very large sample size is needed to detect low risk, but moderate sample size is adequate to discriminate high risk and low risk.
In general, it is difficult to evaluate the risk of ADR in a specific subgroup for example patients with severe liver disease etc. Our approach was useful for evaluating the risk of ADR in such subgroups because the required sample size was small.

Referece:
Machin D, Campbell MJ, Fayers PM, Pinol APY. (2009). Sample Size Tables for Clinical Studies, Second Edition. Wiley-Blackwell.