

## Analysis of Corrections Methods in Genome-Wide Association Studies

Ming Zheng<sup>1, a</sup>, and Mugui Zhuo<sup>1, b\*</sup>

1Guangxi Colleges and Universities Key Laboratory of Professional Software Technology, Wuzhou University, Wuzhou, China

a370505375@qq.com, b756456050@qq.com

**Keywords:** Genome wide association study; Multiple correction; FWER; FDR; Simulation

**Abstract.** Genome wide association study (GWAS) can directly study the relationship between human behavior and genotype, which provides a new way for researchers to explore the genetic basis of human behavior from the whole genome level. GWAS involves a large number of sites and behavior of the association test, so we must use multiple correction to control the overall false. Although there are a variety of correction methods to choose from, but the applicability of different correction methods in GWAS research is still lack of systematic research, which makes the choice of multiple correction methods in GWAS lack of theoretical and empirical basis. GWAS commonly used in the correction method is based on the family-wise error rate (FWER) standard Bonferroni correction method, Holm regression adjustment method, permutation test method and false detection rate (FDR) standard BH method. In this paper, the principle and process of the 4 kinds of multiple correction methods are described in detail. A simulation method of GWAS data is presented. The results show that the first 3 methods based on FWER are very small, they are the most stringent control of false, but the number of bits of the real association is significantly lower than the FDR based BH method. Independent data, the BH reported by the SNPs method has the highest rate of interpretation, that is, compared with other methods, the BH method better balance the false and hit. The BH method can be used to correct the results in future studies.

### Introduction

Human behavior is the result of the interaction between genetic and acquired environment. Heredity determines the individual behavior pattern to a great extent. From ancient Greek philosophy to behavioral genetics, human thinking and exploration of behavior has never stopped. At present, based on molecular biology technology, researchers are gradually unraveling the biological basis of human behavior genetics. In the 90s of last century, researchers have used candidate gene approaches to identify a number of genes that are associated with abnormal behavior. For example, IT15 gene mutations can lead to Huntington's disease<sup>[1]</sup>; the DSCAM gene is associated with down syndrome<sup>[2]</sup>. However, the candidate gene approach to selection dependent protein gene expression to explore the relationship of knowledge, only the gene sequence and encoding region of the phenotype, to study on the relationship between the majority of non-encoding fragment and phenotype.

Since 2005, with the development of molecular biology and gene detection technology, researchers can have sequenced single nucleotide polymorphism in the human genome, opened the behavior genetic studies from the whole genome level of the gate. Genome wide association studies of SNPs and specific behavioral phenotypes, such as intelligence, emerge as the times require.

The GWAS study, researchers usually based on the contribution of single SNP on behavioral phenotypes (explicit and implicit assumptions or additive), relevance examines each SNP and specific behavioral phenotypes, and found a significant association and the phenotype of SNPs<sup>[3]</sup>. Because of the GWAS study, the number of SNPs to be tested will usually reach hundreds of thousands to millions, if in each test SNP and phenotype is associated with significant level, is set to 0.05, then the number of false (false positive) there will be tens of thousands of. For this reason, multiple correction must be performed.

Multiple calibration means that, at the same time, a series of tests are carried out to determine the significant level of a single test in accordance with a specific standard, so as to ensure that the overall false control is at a lower level. At present, there are several criteria to control the false statement in the multiple test: family error rate<sup>[4]</sup> and error detection rate<sup>[5]</sup>. The FWER standard and Bonferroni correction method, Holm regression adjustment method and permutation test method and other different methods. There are two methods of BH (Benjamini-Hochberg)<sup>[6]</sup> and BHY (Benjamini-Hochberg-Yekutieli)<sup>[7]</sup> under the FDR standard. So, what are the differences between the two standards and the corresponding correction methods? Which method is more suitable for GWAS? Therefore, this article from the concept and implementation of in-depth discussion on the differences of several correction methods at the same time, through the simulation data to investigate the performance of these methods in analysis of GWAS data, provide empirical support for the real GWAS in multiple correction methods. Specifically, this paper will first introduce the concept of FWER and FDR two error rate control standard, and makes a detailed description of commonly used methods of two standards, including Bonferroni correction of FWER standard, Holm decline adjustment method, test method and arrangement of BH under FDR standard. For quantitative comparison and validation

This paper designs a simulation method of GWAS data, which is suitable for the correction of GWAS. Based on GWAS simulation data, the performance of different methods is compared. Finally, this paper discusses the selection of multiple correction methods in GWAS research.

## Multiple Correction Method

In the GWAS study, the number of SNPs to be tested ( $n$ ), but the number of SNPs associated with the behavior is very small (denoted  $a$ ). If we set the significant level of each test to  $\alpha$ , then the SNPs number  $f$  is  $\alpha * (n - a)$ . Because  $a$  is very small relative to  $N$ , so the number of false about  $*n$ . For example, the number of SNPs to be tested when  $n$  reached 1 million, if the significance of each test is still 0.05, then the number of false SNPs will reach about 50 thousand. That is, the test will get a lot of false positive SNPs, which constitutes a multiple test problem in GWAS research.

**Multiple Calibration Standard.** In order to control the multiple test of the false, the need for multiple correction, that is, according to certain standards and methods, adjust the significant level of a single test, so that the overall false control is still at a low level. At present, there are two kinds of standards for the whole system to control the whole false statement: FWER and FDR.

FWER is the probability that at least one false statement occurs in multiple tests, and FDR is the mathematical expectation of reporting the proportion of false reports in a significant test. The different calibration methods are to adjust the level of a single test to ensure that the overall false FWER or FDR below a certain level. That, as long as the control of FWER is less than or equal to  $\alpha$ , will ensure that FDR is less than or equal to  $\alpha$ , but not vice versa. In other words, the correction method based on FWER standard is stricter than the calibration method based on FDR standard. Correction method commonly used in GWAS Bonferroni correction method based on FWER standard, Holm decline adjustment method, arrangement method and BH method based on FDR standard. The following will discuss the concept and implementation of these methods.

Bonferroni correction method<sup>[8]</sup> using the FWER standard. If the true set of related SNPs  $A$ , based on a collection of events that can be drawn, as long as the single level test for  $a/n$ , you can make the whole test FWER  $(n - a) \alpha / n$ , which is less than. Therefore, this method is very simple to achieve, as long as the significance of a single test is set to  $\alpha/n$ .

The standard of Holm decline adjustment method<sup>[9]</sup> is FWER, which is an improvement to the Bonferroni correction method. The Bonferroni correction significant level of all SNPs are all  $\alpha/n$ , assuming  $C$  has a SNPs through the correction, then the remaining  $N - C$  SNPs significant level, according to the Bonferroni inequality is the only  $\alpha / (n - C)$ , which is no longer  $\alpha/n$ . According to this, the Holm adjustment method is used to improve the Bonferroni correction method: 1) the  $p$  value of  $n$  SNPs from small to large arrangement; 2) if the SNP can be tested by the standard of  $\alpha/n$ ,

then the alpha  $(n - 1)$  as a significant level to examine the SNP, followed by analogy. That is, before the  $I - 1$  tests are corrected, the significant level of the first  $I$  SNP is set to  $\alpha / (n - i + 1)$ ; 3) until a certain SNP cannot pass the test.

The same as the above two methods, the permutation test method[10] is also used FWER standard. The difference is that it can well deal with the correlation between different tests. In the GWAS study, there was a linkage disequilibrium (Linkage Disequilibrium, LD) between adjacent SNPs, that is, the probability of the combination of two SNPs genotypes was higher than that of each genotype. This leads to the existence of correlation between different tests, so that the effective test number is less than the number of SNPs to be tested  $n$ . The correction method of the Bonferroni or the Holm decline adjustment method still assumes that the effective test number is  $n$ , which can lead to the conclusion that is too conservative. Permutation test method based on the maximum statistics, can be a good deal with the problem. In the FWER standard, there is at least one false statement, which is equivalent to the null hypothesis that at least one test statistic exceeds the significant threshold  $u$ , which is equivalent to the maximum statistic in all tests. The test statistic is  $T$ , so that it is significant to test whether a SNP is significant or not, which is equivalent to the test of whether the statistic is significant. The distribution of the maximum statistic under null hypothesis can be obtained by randomly rearranging the corresponding relationship between SNPs and phenotype.

## Experiment

After setting the parameters, we can simulate the SNPs and the corresponding behavior data, and know exactly what SNPs and behavior associated. Then, we use additive model one by one to test the simulation of SNPs. Specifically: whether the genotype is associated with the behavioral phenotype will be tested by a linear model with the dependent variable  $X_i$  values 0, 1, and 2 corresponding to the SNP genotype, respectively. At the end of all tests, 500 thousand SNPs test statistics and P values can be obtained. And then use different correction methods to get the SNPs. It can be concluded that the number of SNPs in which there is a correlation, that is, the number of hits; and the number of SNPs which is not associated with the number of false. The results of the three kinds of methods were shown as below:

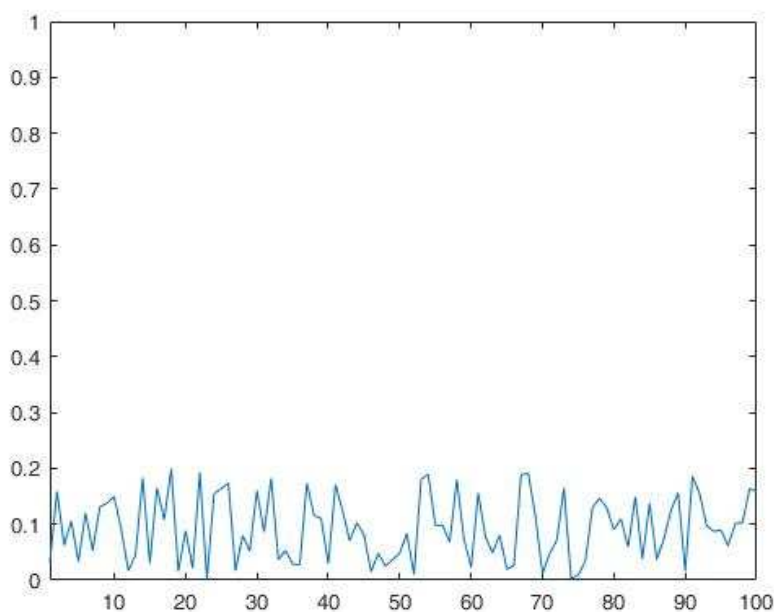


Figure 1. The 100 simulation of different correction methods for average hit and false

## Summary

The results of simulation analysis show that the difference of the multiple correction methods in GWAS research is more intuitive and quantitative. Most of the researchers have focused on the selection of multiple correction methods in GWAS, which have been focused on the selection of different methods in FWER (Wei, 2012). It is true that the Holm method is an improvement to the Bonferroni correction method, and the permutation test method can consider the linkage disequilibrium of SNPs. However, the simulation results tell us that the real difference between the multiple correction methods, not on the standard of FWER, but in FWER and FDR two kinds of false control standard: comprehensive interpretation of standards FDR than FWER standard SNPs on the behavior of the higher. Based on this, we propose to use the FDR based BH method to correct the results in real GWAS studies. In this way, not only can reduce the use of FWER standard false negative, and can still be a good control of false, so as to find those real and behavior related SNPs.

## Acknowledgements

This work was supported by grants from The National Natural Science Foundation of Chi-na (No. 61502343, No. 61373051, and No. 61402423), China Postdoctoral Science Foundation funded (No. 2016M590260), the Guangxi Natural Science Foundation (No. 2015GXNSFBFA139262), the Science Research Funds for the Guangxi Universities (No. KY2015ZD122), Guangxi Colleges and Universities Key Laboratory of Professional Software Technology, Wuzhou University.

## References

- [1] CHARLES J, LESSEY L, ROONEY J, et al. Presentation and care of a family with Huntington disease in a resource-limited community[J]. *J Clin Mov Disord*, 2017, 4: 4.
- [2] SOETERS N, BENNEN E, WISSE R P L. Performing corneal crosslinking under local anaesthesia in patients with Down syndrome[J]. *Int Ophthalmol*, 2017.
- [3] DELAY C, GRENIER-BOLEY B, AMOUYEL P, et al. miRNA-dependent target regulation: functional characterization of single-nucleotide polymorphisms identified in genome-wide association studies of Alzheimer's disease[J]. *Alzheimers Res Ther*, 2016, 8(1): 20.
- [4] DELORME P, DE MICHEAUX P L, LIQUET B, et al. Type-II generalized family-wise error rate formulas with application to sample size determination[J]. 2016, 35(16): 2687-2714.
- [5] IVANOV M V, LOBAS A A, KARPOV D S, et al. Comparison of False Discovery Rate Control Strategies for Variant Peptide Identifications in Shotgun Proteogenomics[J]. 2017.
- [6] GHOSH D. Incorporating the empirical null hypothesis into the Benjamini-Hochberg procedure[J]. *Stat Appl Genet Mol Biol*, 2012, 11(4).
- [7] SHI Y, CUI Y H, WU H G, et al. Effects of mild-warming moxibustion on Bcl-2 and PKC expressions of peripheral blood in elderly people[J]. *J Tradit Chin Med*, 2012, 32(1): 45-51.
- [8] RANSTAM J. Multiple P-values and Bonferroni correction[J]. *Osteoarthritis Cartilage*, 2016, 24(5): 763-764.
- [9] SRIKANTA DANI K G, MARINO G, TAITI C, et al. De novo post-illumination monoterpene burst in *Quercus ilex* (holm oak)[J]. *Planta*, 2017, 245(2): 459-465.
- [10] NGUYEN H D, MCLACHLAN G J, HILL M M. Statistical Evaluation of Labeled Comparative Profiling Proteomics Experiments Using Permutation Test[J]. *Methods Mol Biol*, 2017, 1549: 109-117.