

## Methodology Round



### OPEN ACCESS

**Received:** Feb 7, 2023  
**Revised:** Mar 12, 2023  
**Accepted:** Apr 9, 2023  
**Published online:** May 3, 2023

#### Correspondence to

##### Sanggyu Kwak

Department of Medical Statistics, Daegu Catholic University School of Medicine, 33 Duryongwon-ro 17-gil, Nam-gu, Daegu 42472, Korea.  
Email: sgkwak@cu.ac.kr

**Copyright** © 2023 The Korean Society of Lipid and Atherosclerosis.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

#### ORCID iDs

Sanggyu Kwak <https://orcid.org/0000-0003-0398-5514>

#### Funding

None.

#### Conflict of Interest

Sanggyu Kwak is a statistical editor of Journal of Lipid and Atherosclerosis; however, he was not involved in the peer reviewer selection, evaluation, or decision process of this article. No other potential conflicts of interest relevant to this article were reported.

#### Data Availability Statement

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

# Are Only $p$ -Values Less Than 0.05 Significant? A $p$ -Value Greater Than 0.05 Is Also Significant!

Sanggyu Kwak

Department of Medical Statistics, Daegu Catholic University School of Medicine, Daegu, Korea

## ABSTRACT

Statistical hypothesis testing compares the significance probability value and the significance level value to determine whether or not to reject the null hypothesis. This concludes “significant or not significant.” However, since this process is a process of statistical hypothesis testing, the conclusion of “statistically significant or not statistically significant” is more appropriate than the conclusion of “significant or not significant.” Also, in many studies, the significance level is set to 0.05 to compare with the significance probability value,  $p$ -value. If the  $p$ -value is less than 0.05, it is judged as “significant,” and if the  $p$ -value is greater than 0.05, it is judged as “not significant.” However, since the significance probability is a value set by the researcher according to the circumstances of each study, it does not necessarily have to be 0.05. In a statistical hypothesis test, the conclusion depends on the setting of the significance level value, so the researcher must carefully set the significance level value. In this study, the stages of statistical hypothesis testing were examined in detail, and the exact conclusions accordingly and the contents that should be considered carefully when interpreting them were mentioned with emphasis on statistical hypothesis testing and significance level. In 11 original articles published in the *Journal of Lipid and Atherosclerosis* in 2022, the interpretation of hypothesis testing and the contents of the described conclusions were reviewed from the perspective of statistical hypothesis testing and significance level, and the content that I would like to be supplemented was mentioned.

**Keywords:** Statistics; Statistical data interpretation; Statistical data analysis

## INTRODUCTION

In research, a topic to be identified is selected and hypotheses are established accordingly. In order to calculate the evidence to support this, related data is collected, and the collected data is analyzed using a statistical hypothesis test method suitable for the hypothesis. The method of statistical hypothesis testing is determined according to the type of data corresponding to whether the data is a quantitative variable or a qualitative variable, the research design and hypothesis, etc., but in the end, the hypothesis test is performed using the significance probability value calculated as a result of statistical analysis.

When the significance probability value is less than the significance level value, the conclusion and interpretation described are “significant” or you can easily come across papers that express it as “significantly different.” However, the conclusion through the comparison of the significance probability value and the significance level value is the conclusion through the statistical hypothesis test step. Therefore, the researcher is “significant” or “statistically significant under the value of the significance level” rather than “significantly different.” Alternatively, the expression “is statistically significantly different under the value of the significance level” should be used.

In addition, it is easy to see papers that set the significance level to 0.05. In statistical hypothesis testing, the significance level value is a value that determines whether or not to reject the null hypothesis by comparing it with the calculated significance probability value. The significance level value is a value set by the researcher according to the circumstances of the study. If the significance level was set at 0.05 in other studies, there is no need to set it at 0.05 in your own study. In statistical hypothesis testing, the conclusion depends on the setting of the significance level value, so the researcher must carefully set the significance level value.

In this study, we will examine the steps of statistical hypothesis testing in detail, what the exact conclusions will be, and what should be carefully considered when interpreting them, focusing on the significance level. In 13 original articles published in the *Journal of Lipid and Atherosclerosis* for one year in 2022, I would like to look at the interpretation of hypothesis testing and the contents of the described conclusions from the point of view of the significance level, and mention what I would like to see supplemented.

## **STEPS IN STATISTICAL HYPOTHESIS TEST**

Statistical hypothesis testing consists of 5 steps: 1) hypothesis setting, 2) significance level setting, 3) test statistic calculation, 4) rejection area or significance probability calculation, and 5) conclusion. In the conclusion stage, if the value of the test statistic is not included in the rejection range or if the value of the significance level  $< p$ -value, “the null hypothesis cannot be rejected under the value of the significance level,” so the conclusion is made with the contents of the null hypothesis and interpreted. If the value of the test statistic is included in the rejection region or the value of the significance level  $> p$ -value, “the null hypothesis is rejected under the value of the significance level,” so the conclusion is made with the contents of the alternative hypothesis opposite to the null hypothesis and interpreted. For example, when performing a statistical hypothesis test, if the significance level is set to 0.05 and the calculated significance probability is 0.002, the set null hypothesis is rejected and the conclusion is made with the content of the alternative hypothesis. The detailed steps of each statistical hypothesis test are as follows.

### **1. Hypothesis establishment**

When establishing hypotheses in statistical hypothesis testing, 2 hypotheses must be established: the null hypothesis and the alternative hypothesis. The null hypothesis is a tentative argument proposing that it is true about the characteristics of a population, and the alternative hypothesis is a hypothesis that is alternatively true if the null hypothesis is false. The 2 hypotheses are determined depending on which statistical hypothesis test method is used, and based on the sample data collected by the researcher, it is determined whether or not to reject the null hypothesis. This is called the significance test of the null hypothesis.

## **2. Set significance level**

The significance level means the probability of rejecting the null hypothesis due to incorrect judgment even though the null hypothesis is actually true in hypothesis testing and corresponds to a type 1 error. When performing a statistical hypothesis test, the significance level can be set to 1%, 5%, or 10%. If the significance level is set to 5%, it means that the null hypothesis is rejected 5 times out of 100 even though it is true. In other words, you are 95% sure that you will test the correct hypothesis. This concludes that it was rejected under the significance level of 5% as a result of statistical hypothesis testing.

## **3. Calculation of test statistics**

In order to perform a statistical hypothesis test, a basis for determining whether the null hypothesis is true or not is required, and the value used as the basis is the test statistic. The test statistic is a statistic used in the stage of statistical hypothesis testing, and the test statistic value is calculated based on the collected sample data. The formula for calculating the test statistic value is determined depending on which statistical hypothesis test method is used. Each test statistic value corresponds to a random variable and follows a continuous probability distribution such as normal distribution, t distribution, chi-square distribution, and F distribution.

## **4. Rejection area or significance probability calculation**

Rejection area means the area in which the null hypothesis is rejected under a specific probability distribution of the test statistic according to the significance level set in step 2 in the statistical hypothesis test. Depending on whether the type of statistical hypothesis test is a 2-sided test or a one-sided test, it is the area corresponding to the significance level value at both extremes/2 or the area corresponding to the left or right significance level value on the probability distribution. The non-rejection region of the probability distribution is the acceptance region in which the null hypothesis is not rejected. In addition, the significance probability is the minimum probability to reject the null hypothesis and has a value between 0 and 1 as a random variable. The significance probability value is a value twice or 1 time the area corresponding to the left or right test statistic value of the test statistic on the probability distribution, depending on whether the type of statistical hypothesis test is a 2-sided test or a one-sided test.

## **5. Conclusion**

Depending on whether or not the test statistic value calculated under the probability distribution is included in the rejection range generated by the significance level value set in step 2, whether to reject the null hypothesis or not is determined. In addition, whether to reject the null hypothesis or not is determined according to whether the calculated significance probability value is greater than or less than the significance level value set in step 2. Therefore, the sentences “The value of the test statistic belongs to the rejection region generated by the value of the significance level” and “The value of the significance probability is smaller than the value of the significance level” are sentences with meaning. If the null hypothesis is rejected, the correct conclusion is “It can be said that the null hypothesis is not statistically valid under the value of the significance level,” and if the null hypothesis cannot be rejected, “It cannot be said that the content of the null hypothesis is not statistically rejected under the value of the significance level.” is the correct conclusion. Statistical hypothesis testing does not determine whether to reject or accept the alternative hypothesis. Rejecting the null hypothesis does not mean accepting the alternative hypothesis.

## THE REAL MEANING OF STATISTICAL HYPOTHESIS TESTING

### 1. Under the significance level

As mentioned above, the statistical hypothesis test concludes by comparing the significance probability value and the significance level value. If the significance probability value is smaller than the significance level value, it is concluded that “It can be said that the content of the null hypothesis is not statistically under the significance level value.” And if the value of the significance probability is greater than the value of the significance level, the correct conclusion is “It cannot be said that it is not statistically null hypothesis under the value of the significance level.”

### 2. It is up to the researcher to set the significance level

As mentioned above, the statistical hypothesis test concludes by comparing the significance probability value and the significance level value. For example, when performing a statistical hypothesis test, if the significance level is set to 0.05 and the calculated significance probability is 0.07, the null hypothesis cannot be rejected, and the null hypothesis is concluded. However, agreeingly, if the significance level is set to 0.10 when the significance probability is 0.07, the null hypothesis is rejected, and the conclusion is made with the content of the alternative hypothesis contrary to the null hypothesis. Similarly, if the value of the significance level is set to 0.05 and the calculated significance probability value is 0.03, the set null hypothesis will be rejected, but if the value of the significance level is set to 0.01, the null hypothesis cannot be rejected. That is, depending on the value of the significance level, the null hypothesis may or may not be rejected even with the same significance probability value. Earlier, the level of significance was mentioned as a type I error. In other words, setting the value of the significance level is to determine the tolerance for errors that may occur in statistical hypothesis testing. In statistical hypothesis testing, there are no standards, principles, or formulas for determining the value of the significance level, and it is up to the researcher to set it. If the significance level, which is a type 1 error, is set to a very large value of 1, the null hypothesis is rejected because the significance level value is always greater than or equal to the significance probability. However, if the value of the significance level is 1, the probability of making an incorrect conclusion according to the statistical hypothesis test is 100%, so the statistical hypothesis test is meaningless. Therefore, the researcher should conduct statistical hypothesis testing by setting an appropriate significance level value according to the study.

## EXAMPLE

In this session, in 13 original articles published in the *Journal of Lipid and Atherosclerosis* for one year in 2022, we will review the interpretation of hypothesis testing and the contents of the conclusions described from the point of view of the significance level, and mention what we would like to see supplemented. Among the 13 papers, a total of 11 papers were reviewed, excluding 2 papers that did not use statistical hypothesis testing (**Table 1**). There were 10 papers that described the significance level set in the contents of statistical analysis. In 9 papers, the significance level was set to 0.05, and in one paper, the significance level was set to 0.01. In addition, 3 papers described that they were statistically significant in the sentences written in relation to the results of statistical hypothesis testing, and the remaining 8 papers did not include the word statistical, but only stated that they were significant.

**Table 1.** Assessment of sentences that describe the level of significance and describe the statistical hypothesis test results of 11 original articles published in the *Journal of Lipid and Atherosclerosis* in 2022

No.	Last name of first author	Year	Volume	Issue	Sentences that describe the level of significance	The value of significance level set by the researcher	Examples of sentences describing statistical hypothesis test results	Suggested sentences
1	Kim et al. <sup>1</sup>	2022	11	1	<i>p</i> -value less than 0.05 was considered statistically significant.	0.05	There were no significant differences in statin use, HDL-cholesterol, or LDL-cholesterol between patients with and without an extreme risk for ASCVD.	There was no <b>statistically</b> significant difference in statin use, HDL-cholesterol, or LDL-cholesterol between patients at extreme risk for ASCVD and those without.
2	Jeon et al. <sup>3</sup>	2022	11	1	Statistical significance was considered when <i>p</i> <0.05.	0.05	There was no significant difference in all-cause mortality among the 13 studies comparing SUs and DPP4 inhibitors as add-on therapies to metformin (random-effect relative risk [RR], 1.14; 95% confidence interval [CI], 0.98–1.33; I <sup>2</sup> =0%; <i>p</i> =0.098)	There was no <b>statistically</b> significant difference in all-cause mortality among the 13 studies comparing SUs and DPP4 inhibitors as add-on therapies to metformin (random-effect relative risk [RR], 1.14; 95% confidence interval [CI], 0.98–1.33; I <sup>2</sup> =0%; <i>p</i> =0.098)
3	Lee et al. <sup>4</sup>	2022	11	2	The volcano plots were described through the “EnhancedVolcano” package with a significant cutoff, lower <i>p</i> -value than 0.01.	0.01	These eight genes showed significant associations with AD pathology.	These 8 genes were <b>statistically</b> associated with AD pathology.
4	Cho et al. <sup>5</sup>	2022	11	2	A 2-sided <i>p</i> -value of less than 0.05 was considered statistically significant.	0.05	There were significant differences in mortality rates across the 4 groups.	There were <b>statistically</b> significant difference in mortality between the 4 groups.
5	Ochin et al. <sup>6</sup>	2022	11	2	<i>p</i> -values <0.05 were considered to be statistically significant.	0.05	This distribution was statistically significant ( <i>p</i> <0.001) for palmitic acid, LA, oleic acid and stearic acid across all groups.	No suggestion.
6	Lan et al. <sup>7</sup>	2022	11	2	A 2-sided <i>p</i> -value of <0.05 was used to define statistical significance.	0.05	The proportion of patients attaining lipid targets according to year of CABG surgery was not significantly different, as presented in Supplementary Table 1.	The proportion of patients attaining lipid targets according to year of CABG surgery was not <b>statistically</b> significantly different, as presented in Supplementary Table 1.
7	Cho et al. <sup>2</sup>	2022	11	2	Not presented.	None.	The risks of MI ( <i>p</i> for interaction <0.001) and death ( <i>p</i> for interaction=0.0006) for the 4th quartile of the TyG index in women were significantly higher than those of men.	The risks of MI ( <i>p</i> for interaction <0.001) and death ( <i>p</i> for interaction=0.0006) for the 4th quartile of the TyG index in women were <b>statistically</b> significantly higher than those of men <b>at the 5% significance level.</b>
8	Tarkhnishvili et al. <sup>8</sup>	2022	11	2	Significant difference was accepted when <i>p</i> <0.05.	0.05	Mean values of palmitate oxidation were increased by 2.7-fold in vehicle-treated db/db mice compared to vehicle-treated C57BLKS/J mice; statistical significance was however not achieved (Fig. 2E).	No suggestion.
9	Jeong et al. <sup>9</sup>	2022	11	3	<i>p</i> <0.05 was considered statistically significant.	0.05	Among participants not taking statins, LDL-C levels ≥120 mg/dL were significantly associated with MI risk (1.33 [1.27–1.40]). LDL-C ≥120 mg/dL was significantly associated with increased MI risk in statin users compared with statin users with LDL-C <80 mg/dL.	Among participants not taking statins, an LDL-C level ≥120 mg/dL was <b>statistically</b> significantly associated with MI risk (1.33 [1.27–1.40]). LDL-C ≥120 mg/dL was <b>statistically</b> significantly associated with an increased risk of MI in statin users compared to statin users with LDL-C <80 mg/dL.
10	Hayiroğlu et al. <sup>10</sup>	2022	11	3	As the level of statistical significance, a <i>p</i> -value of ≤0.05 was chosen.	0.05	In regard to laboratory measurements, the patients with long-term MACEs had significantly higher levels of FBG, creatinine, TG, the TyG index.	In regard laboratory measurements, the patients with long-term MACEs had statistically significantly higher levels of FBG, creatinine, TG, and TyG indices
11	Nam et al. <sup>11</sup>	2022	11	3	All variables with a <i>p</i> <0.05 were considered statistically significant.	0.05	CMBs did not have statistical significance with AIP.	No suggestion.

HDL, high-density lipoprotein; LDL, low-density lipoprotein; ASCVD, atherosclerotic cardiovascular disease; SU, sulfonyleurea; DPP4, dipeptidyl peptidase-4; AD, Alzheimer’s disease; LA, left atrial; CABG, coronary artery bypass graft; MI, myocardial infarction; TyG index, triglyceride-glucose; LDL-C, low-density lipoprotein cholesterol; MACE, major adverse cardiovascular event; FBG, fasting blood glucose; TG, triglyceride; CMB, cerebral microbleed; AIP, atherogenic index of plasma.

The test of “significant” or “not significant” based on the significance probability value is a conclusion based on statistical hypothesis testing, so the expression “statistically significant” or “not statistically significant” is a more appropriate expression. It is foreseeable. Therefore, sentences expressed as simple significant or not significant are not appropriate in terms of statistical hypothesis testing and significance level. Therefore, in the “suggested sentences” column in **Table 1**, the proposed sentence for the sentence described in the existing paper was modified and described. For example, in Kim et al.,<sup>1</sup> the sentence “There were no significant differences in statin use, HDL-cholesterol, or LDL-cholesterol between patients with and without an extreme risk for ASCVD.” It is suggested to add only since “*p*-value less than 0.05 was considered statistically significant” was described in the content of statistical analysis, it is not necessary to include the significance level value. However, in Cho et al.,<sup>2</sup> there was no description of the significance level set in the statistical analysis, so “The risks of MI (*p* for interaction <0.001) and death (*p* for interaction=0.0006) for the 4th quartile The sentence expressed as “of the TyG index in women were significantly higher than those of men.” added the words “statistically” and “at the 5% significance level” to obtain “The risks of MI (*p* for interaction <0.001) and death (*p* for interaction=0.0006) for the 4th quartile of the TyG index in women were significantly higher than those of men at the 5% significance level.”

## CONCLUSION

In this study, what is the correct conclusion according to the statistical hypothesis test and what should be considered carefully when interpreting it is described. The meaning of the significance level was looked at precisely and it was emphasized that statistical hypothesis testing is performed under the value of the significance level. In addition, since it is up to the researcher to determine the value of the significance level, the researcher emphasized that the statistical hypothesis test should be performed by setting the significance level suitable for the researcher. By utilizing this, many researchers expect to draw accurate conclusions and make judgments consistent with the hypothesis when conducting statistical hypothesis testing.

## REFERENCES

1. Kim KS, Hong S, Han K, Park CY. Assessing the validity of the criteria for the extreme risk category of atherosclerotic cardiovascular disease: a nationwide population-based study. *J Lipid Atheroscler* 2022;11:73-83.  
[PUBMED](#) | [CROSSREF](#)
2. Cho YK, Han KD, Kim HS, Jung CH, Park JY, Lee WJ. Triglyceride-glucose index is a useful marker for predicting future cardiovascular disease and mortality in young Korean adults: a nationwide population-based cohort study. *J Lipid Atheroscler* 2022;11:178-186.  
[PUBMED](#) | [CROSSREF](#)
3. Jeon WK, Kang J, Kim HS, Park KW. Correction to: “cardiovascular outcomes comparison of dipeptidyl peptidase-4 inhibitors versus sulfonylurea as add-on therapy for type 2 diabetes mellitus: a meta-analysis”. *J Lipid Atheroscler* 2022;11:89-101.  
[PUBMED](#) | [CROSSREF](#)
4. Lee J, Park S, Kim Y, Kim HM, Oh CM. Exploring the genetic associations between the use of statins and Alzheimer’s disease. *J Lipid Atheroscler* 2022;11:133-146.  
[PUBMED](#) | [CROSSREF](#)
5. Cho KH, Shin MH, Kim MC, Sim DS, Hong YJ, Kim JH, et al. Prognostic value of baseline neutrophil-to-lymphocyte ratio combined with anemia in patients with ST-segment elevation myocardial infarction: a nationwide prospective cohort study. *J Lipid Atheroscler* 2022;11:147-160.  
[PUBMED](#) | [CROSSREF](#)

6. Ochin CC, Wilson T, Garelnabi M. Dietary oxidized linoleic acids modulate fatty acids in mice. *J Lipid Atheroscler* 2022;11:197-210.  
[PUBMED](#) | [CROSSREF](#)
7. Lan NS, Ali US, Yeap BB, Fegan PG, Larbalestier R, Bell DA. Attainment of lipid targets following coronary artery bypass graft surgery: can we do better? *J Lipid Atheroscler* 2022;11:187-196.  
[PUBMED](#) | [CROSSREF](#)
8. Tarkhnishvili A, Koentges C, Pfeil K, Gollmer J, Byrne NJ, Vosko I, et al. Effects of short term adiponectin receptor agonism on cardiac function and energetics in diabetic *db/db* mice. *J Lipid Atheroscler* 2022;11:161-177.  
[PUBMED](#) | [CROSSREF](#)
9. Jeong H, Han K, Yoo SJ, Kim MK. Low-density lipoprotein cholesterol level, statin use and myocardial infarction risk in young adults. *J Lipid Atheroscler* 2022;11:288-298.  
[PUBMED](#) | [CROSSREF](#)
10. Hayiroğlu MI, Çınar T, Çiçek V, Palice A, Ayhan G, Tekkeşin AI. The triglyceride-glucose index can predict long-term major adverse cardiovascular events in Turkish patients with high cardiovascular risk. *J Lipid Atheroscler* 2022;11:280-287.  
[PUBMED](#) | [CROSSREF](#)
11. Nam KW, Kwon HM, Park JH, Kwon H. The atherogenic index of plasma is associated with cerebral small vessel disease: a cross-sectional study. *J Lipid Atheroscler* 2022;11:262-271.  
[PUBMED](#) | [CROSSREF](#)