

What Is the Bonferroni Correction?

Matthew A. Napierala, MD

The Bonferroni correction is an adjustment made to P values when several dependent or independent statistical tests are being performed simultaneously on a single data set. To perform a Bonferroni correction, divide the critical P value (α) by the number of comparisons being made. For example, if 10 hypotheses are being tested, the new critical P value would be $\alpha/10$. The statistical power of the study is then calculated based on this modified P value.

The Bonferroni correction is used to reduce the chances of obtaining false-positive results (type I errors) when multiple pair wise tests are performed on a single set of data. Put simply, the probability of identifying at least one significant result due to chance increases as more hypotheses are tested.

For example, a researcher is testing 20 hypotheses simultaneously, with a critical P value of 0.05. In this case, the following would be true:

- P (at least one significant result) = $1 - P$ (no significant results)
- P (at least one significant result) = $1 - (1-0.05)^{20}$
- P (at least one significant result) = 0.64

Thus, performing 20 tests on a data set yields a 64 percent chance of identifying at least one significant result, even if all of the tests are actually not significant. Therefore, while a given α may be appropriate for each individual comparison, it may not be appropriate for the set of all comparisons.

This fact is potentially problematic because in contemporary orthopaedic research studies, numerous simultaneous tests are routinely performed. Thus, to avoid a large number of spurious positives, α must be lowered to account for the number of comparisons being performed.

The Bonferroni correction is based on the idea that if an experimenter is testing n dependent or independent hypotheses on a set of data, the probability of type I error is offset by testing each

hypothesis at a statistical significance level $1/n$ times what it would be if only one hypothesis were tested.

An example of the use of the Bonferroni correction in the orthopaedic literature can be found in a recent article in *The Journal of Bone & Joint Surgery–American*, “The Relationship Between Time to Surgical Débridement and Incidence of Infection After Open High-Energy Lower Extremity Trauma.” The purpose of the study was to evaluate the relationship between the timing of the initial treatment of open fractures and the development of subsequent infection.

The study consisted of 315 patients identified as a subgroup of the Lower Extremity Assessment Project (LEAP). For their analysis, the investigators created two outcome measures (“All Infections” and “Major Infections”) and used four time-to-treatment variables. In their manuscript, the authors stated the following: “Because two outcome measures were tested against four hypothesized predictors, a Bonferroni-adjusted significance level of 0.00625 was calculated to account for the increased possibility of type-I error.” In other words, because 8 hypotheses were being tested on one set of data, the chance of obtaining a false-positive result was 34 percent. Accordingly, the authors used the Bonferroni correction to adjust the P value for each hypothesis to 0.00625 to neutralize this risk.

However, although the Bonferroni correction controls for false positives, it can become very conservative as the number of tests increases. This, in turn, increases the risk of generating false negatives (type II errors).

In sum, the risk of making erroneous false-positive conclusions is increased when testing multiple hypotheses on a single set of data. This fact is often underappreciated by investigators and consumers of orthopaedic literature. The Bonferroni correction is a simple statistical method for mitigating this risk, and its appropriate use can ensure the integrity of studies in which a large number of significance tests are used. Other tests that also control for false positives, without the risk of increasing false negatives, are the Tukey and Dunnett’s tests.

Matthew A. Napierala, MD, is a resident in San Antonio, Texas.

References:

1. Bland JM, Altman DG: Multiple significance tests: The Bonferroni method. *BMJ* 1995;310(6973):170.
2. Bhandari M, Whang W, Kuo JC, Devereaux PJ, Sprague S, Tornetta P III: The risk of false-positive results in orthopaedic surgical trials. *Clin Orthop Relat Res* 2003;(413):63-69.
3. Pollak AN, Jones AL, Castillo RC, Bosse MJ, MacKenzie EJ, LEAP Study Group: The relationship between time to surgical debridement and incidence of infection after open high-energy lower extremity trauma. *J Bone Joint Surg Am* 2010;92(1):7-15.

<http://www.aaos.org/news/aaosnow/apr12/research7.asp>

6300 North River Road Rosemont, Illinois 60018-4262 Phone 847.823.7186 Fax 847.823.8125

© 1995-2014 by the American Academy of Orthopaedic Surgeons. "All Rights Reserved." This website and its contents may not be reproduced in whole or in part without written permission. "American Academy of Orthopaedic Surgeons" and its associated seal and "American Association of Orthopaedic Surgeons" and its logo are all registered U.S. trademarks and may not be used without written permission.