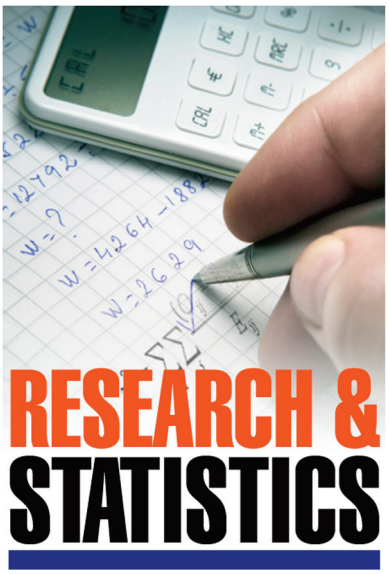
PediatricsinReview®

Research and Statistics :

Sharon R. Ghazarian

Pediatrics in Review 2011;32;73
DOI: 10.1542/pir.32-2-73



Distinguishing Statistical Significance from Clinical Importance: The Value of the *P* Value

Sharon R. Ghazarian, PhD*

Author Disclosure

Dr Ghazarian has disclosed no financial relationships relevant to this article. This commentary does not contain a discussion of an unapproved/investigative use of a commercial product/device.

Case Presentation

You have been providing handouts on infant massage to your families who have newborns for several years. You had read about the benefits of infant massage, including increased bonding between parent and child, improved infant digestion, facilitation of attachment, improved sleep, and better infant weight gain. (1) However, you recently read an article showing the results of a comparative panel design study in which the association between infant massage and weight gain was statistically significant at 2 months but not at 4 months of age for mothers and their infants in Santiago, Chile. (2) You now wonder whether infant massage is as beneficial as you once thought and whether you should continue to provide the information to parents of newborns.

Introduction

Many practitioners wonder whether a lack of statistical significance means a lack of clinical importance for their patients. If a study finds an association to be statistically significant, does it mean that the finding also is clinically important? Alternatively, if a study finds that an association is not statistically significant, does it mean that the finding is not clinically important for the families in a practitioner's practice? Statistical significance demonstrates the likelihood that a difference or relationship really exists and did not just occur by chance. We often are excited by a significant *P* value, but that value does not necessarily mean that the finding is important or useful.

Determining statistical significance involves hypothesis testing and comparison of study results to a predetermined level or *P* value. The level of significance is the probability of reporting an incorrect association or committing a type I error. Conventional *P* values usually are set at either 0.05 or 0.01. A *P* value of 0.05 indicates that the researcher is 95% sure that the finding is a true association and not just due to chance. A *P* value of 0.01 indicates that the researcher is 99% sure that the finding is a true association and that the finding could be due to chance only 1% of the time. Thus, if a study reports that an association between infant massage and infant weight gain is significant at the 0.05 level, the association can be considered as true 95% of the time and possibly due to chance 5% of the time. (3)

Although *P* values measure chance of association, they do not measure strength of association or clinical importance. *P* values tend to be smaller when there is a larger numeric difference between the two groups, a larger sample size, or less variability within treatment groups. (4) Values are also affected by the proportion of those exposed and the proportion who have the targeted outcome. In fact, because of the limitations and misuses of *P* values, the International Committee of Medical Journal Editors suggests that "when possible, quantify findings and present them with appropriate indicators of measurement error or un-

*Director, Bayview Biostatistics, Epidemiology and Data Management Core, Department of Pediatrics, Johns Hopkins University, Baltimore, MD.

certainty (such as confidence intervals). Avoid relying solely on statistical hypothesis testing, such as the use of p-values, which fails to convey important quantitative information." (5)

The authors of the infant massage study (2) found a statistically significant association between infant massage and infant weight gain at 2 months of age (5,537 g for the massage group versus 4,672 g for the control group) with a *P* value of 0.0016. Because the *P* value is less than 0.01, there is less than a 1% likelihood that the findings are simply due to chance. Accordingly, there probably is an association at 2 months. Is this degree of weight change important? The study also shows that the association between infant massage and weight gain at 4 months was not statistically significant (7,044 g for the massage group versus 7,119 g for the control group). Although there was a statistically significant association between massage and weight gain at 2 months, the clinical importance is unclear. What does the statistical difference in weight gain suggest and what does the nonstatistical difference in weight gain at 4 months suggest?

In addition, the study used a sample of 100 newborns in a low-income neighborhood of Santiago, Chile. Unless a practitioner has a practice with the same demographic characteristics of the study sample, study results may not be generalizable to other children. To ascertain whether infant massage is beneficial, numerous studies would need to be examined with samples of families possessing a variety of demographic characteristics, including various levels of income, education, family structure, race/ethnicity, parity, marital status, and more.

Conclusion

Practitioners should consider clinical importance independent from statistical significance, in this case asking whether broader literature, beyond the single study demonstrating limited statistical significance, suggests that infant massage is beneficial for parents and newborns. Final decisions about what to recommend to families should be based on broad literature searches, studies generalizable to a clinician's patient population, and careful consideration of what is statistically and clinically important to the families being served.

References

1. Lorenz L, Moyse K, Surguy H. The benefits of baby massage. *Paediatr Nurs.* 2005; 17:15–18
2. Serrano MSC, Doren FM, Wilson L. Teaching Chilean mothers to massage their full-term infants. *J Perinat Neonat Nurs.* 2010;24:172–179
3. Howell DC. *Statistical Methods for Psychology.* Pacific Grove, CA: Duxbury; 2002
4. Cummings P, Koepsell TD. P values versus estimates of association with confidence intervals. *Arch Pediatr Adolesc Med.* 2010; 164:193–196
5. International Committee of Medical Journal Editors. *Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication.* Accessed November 2010 at: http://www.icmje.org

Research and Statistics :
Sharon R. Ghazarian
Pediatrics in Review 2011;32;73
DOI: 10.1542/pir.32-2-73