

LETTERS TO THE EDITOR

Confidence Intervals, “P”s and Lights

I García-Doval

Servicio de Dermatología, Complejo Hospitalario de Pontevedra, Spain

To the Editor:

The General Universitario y San Juan de Alicante hospital group has been publishing in *Actas-Dermosifiliográficas* a series of very useful, interesting, and painstaking bibliometric studies.^{1,2} For the sole purpose of improving them, I would like to mention a repeated error which, although not serious, could indicate a methodological weakness in the editing process of the journal.

In research, the usual approach consists of analyzing a sample in order to obtain results applicable to the population from which the sample was taken. Since a sample is used, the measurement is associated with a random error, which may be quantitated by statistical estimation. To do this, confidence intervals are most commonly used, as they provide an actual value for the sample and a range in which the actual value of the study population is likely to be found.

However, on rare occasions, results are obtained for the entire population rather than a sample and estimation makes no sense because the actual result is available for the population. Such is the case of the articles

mentioned. If, for instance, all *Actas* papers from 2003 to 2005 are studied and compared to those from 2000 to 2002, estimation is not needed, and the confidence intervals provided in the paper are unnecessary and false (they are based on an analysis of a much larger population sample): the true confidence interval includes only the value found.

The same can be said of hypothesis testing: the *P* value indicates the probability that the results of a sample will be found when the null hypothesis is true (usually, that there is no difference) and considering the existence of a random variation associated with sampling. This is applicable when samples are studied. In this case, entire populations are studied and it makes no sense to use hypothesis testing (all *P* values “would be” 0 when different and 1 when not). This case is also an example of the actual value of *P*: all the differences found are statistically significant ($P=0$); the difficulty lies in knowing if they are significant from the “clinical” point of view, something the statistics cannot reveal.

I feel that these papers would be improved if the *P* values and confidence intervals were eliminated, and measures of dispersion (standard deviations, percentiles, or ranges) were used to describe the data instead.

As a curious anecdote, I am reminded of Castle’s excellent English humor³ in his book on an introduction to statistics where he said that physicians tend to use statistics in the same way that drunks use lampposts: for support, rather than lighting. Warm regards.

References

1. Belinchón I, Miralles J, Ramos JM. Análisis de la producción científica de la revista *Actas Dermosifiliográficas* en el trienio 2003-2005. *Actas Dermosifiliogr.* 2006;97:428-35.
2. Ramos JM, Belinchón I, Miralles J, Ballester R, Sánchez-Yus E. Análisis de la producción científica nacional e internacional de los dermatólogos españoles (1988-2000). *Actas Dermosifiliogr.* 2006;97:436-43.
3. Castle WM. *Statistics in operation*. Edinburgh: Churchill Livingstone; 1979. p. 4.

Reply

JM Ramos^a and I Belinchón^b

^aServicio de Medicina Interna del Hospital General Universitario de Elche and ^bSección de Dermatología del Hospital General Universitario de Alicante, Spain

To the Editor:

We appreciate the interest shown by Dr García-Doval¹ in our bibliometric studies of the *Actas Dermo-Sifiliográficas* journal. In reply to his comments on the methodology, the first article, “Análisis de la producción científica de la revista

Actas Dermo-Sifiliográficas en el trienio 2003-2005” (Analysis of Scientific Production of *Actas Dermo-Sifiliográficas* between 2003 and 2005),² “all” papers in the two periods are compared, as the title indicates. As the author points out, because the entire

sample for the journal is available, it is not entirely correct to use the confidence interval as an estimated measure of the sample since we used all documents (entire population). In this article we used the confidence interval as a measure of dispersion in the quantitative values.