SOME MISUSES OF STATISTICS AND HOW TO AVOID THEM

Ferenc Rárosi & Krisztina Boda
Department of Medical Physics and Informatics, Bolyai Institute, University Szeged, Hungary
<rarosi.ferenc@med.u-szeged.hu>

BACKGROUND:
In medical literature, the misuse or misinterpretation of statistics is still a frequent problem. Several papers deal with this problem. (Phylip Schatz et al., 2005; Matthew S. Thiese et al., 2015; Goodman S., 2011) Among many typical mistakes, this poster will show some of such mistakes found frequently in our practice in collaboration with medical scientists.

FINDINGS:
Correlations. The existence of correlation does not necessarily mean a casual relationship between two variables (http://twentytwowords.com/funny-graphs-show-correlation-between-completely-unrelated-stats-9-pictures).
Misinterpretation of statistical significance. It is well known that hypothesis tests tend to reject H_0 for larger samples and even a practically irrelevant result will be significant. For example Pearson correlation of *R=0.1, n=400* will be statistically significant at *5%* level, but only 1% of the total variance of the dependent variable could be explained by the fitted straight line. (The determination coefficient is just *0.01*.) The statistical significance does not mean necessarily a practically relevant result. There are several other possible misinterpretations of statistical significance (Goodman S, 2011).
Misinterpretation of the lack of the significance. Researchers often interpret the lack of significance as the 'proof' of the H_0, even if the confidence intervals are very wide. Another problem is that non-significant results are often not published (Song F et al, 2010).
Multiplicity problem. 'Fishing for p-values'. After several hypothesis test there is a serious increase of type I error rate. After 13 hypothesis tests it is more likely to find false significant result (at 5% level) than not to find. This problem is mainly connected to the interpretation of the results of multiple comparisons.

CONCLUSIONS:
Authors believe that these examples and similar ones help (medical) students and scientists to avoid these mistakes. It is important to bring these possible misinterpretations and similar ones to students' and scientists' attention and emphasize critical thinking.

References:
Darrell Huff (1991). *How to Lie with Statistics*. Penguin; New Edition.
Phylip Schatz et al. (2005). Misuse of statistical tests in Archives of Clinical Neuropsychology publications. *Archives of Clinical Neuropsychology 20,* 1053–1059. doi: 10.1016/j.acn.2005.06.006.
Matthew S. Thiese et al (2015). The misuse and abuse of statistics in biomedical research. *Biochemia Medica 25,* 5–11. doi: 10.11613/BM.2015.001.
http://twentytwowords.com/funny-graphs-show-correlation-between-completely-unrelated-stats-9-pictures
Goodman S. (2011). A dirty dozen: twelve p-value misconceptions. *Semin Hematol. 48 (4),* 135-140. doi: 10.1053/j.seminhematol.2008.04.003.

