VOLUME 3, ISSUE 6

FOCUS on Field Epidemiology

DISCUSSION QUESTIONS: Data Analysis: Simple Statistical Tests

1. In this issue of FOCUS, analytic techniques using chi-square (χ^2) were presented. What information do you get from the chi-square result? How does this differ from the information you get from a measure of association (such as odds ratios or risk ratios)?

When conducting χ^2 analyses, the chi-square statistic is used primarily to identify the corresponding p-value and determine statistical significance. Often, as in the examples given in the issue, the chi-square value itself is not even presented – the p-value is the important result. Thus, the chi-square test is a test of statistical significance. When interpreting results, we usually give statements like "The difference in exposure between those who are ill and those who are well is statistically significant, indicating that the exposure is associated with disease." Alternatively, if there was not a statistically significant result, we would conclude that the exposure was NOT associated with disease.

When calculating odds ratios (OR's) or risk ratios (RR's), however, we are also able to focus on the magnitude of the association. If the RR is 4.5, with a 95% confidence interval of 3.8 – 5.6, we can say that the risk of disease in those who were exposed was 4.5 times the risk in those who were not exposed. The confidence interval gives us an idea of how precise the estimate of the RR is, and can also indicate statistical significance. An interval that includes the value of 1.0 is not statistically significant. In the example confidence interval given here (3.8 – 5.6), 1.0 is not included, so we know that not only is our estimate of the RR fairly precise, it is also statistically significant.

2. How comfortable do you feel interpreting confidence intervals? Do you use them to evaluate the precision of your odds ratio or risk ratio, or as a test of statistical significance?

In field epidemiology, confidence intervals are often used more as a test of statistical significance, since the studies conducted in field epidemiology tend to be small, meaning that the odds or risk ratios have extremely wide confidence intervals. When looking for a way to control an outbreak, you look for a confidence interval that does not include the value of 1.0 to indicate that the exposure in question may be the cause of the outbreak.

However, the confidence interval as a measure of precision is worth keeping in mind. If you have conducted a case-control study to determine the cause of the outbreak, it could be that NONE of the odds ratios you calculate are statistically significant. What would you do then? Just choose the exposure with the largest odds ratio as the most likely cause of the outbreak?

Say that the odds ratio for exposure to a beverage is 59, while exposure to a salad bar item is 22. This would be a good time to look at the confidence interval. What if the beverage confidence interval was (0.1 – 346), while the salad bar item's confidence interval was (0.8 – 34)? Which estimate do you think is more likely to be the cause of the outbreak, and how would you proceed?

Sometimes these decisions must be made considering the original data (what proportion of ill people were actually exposed to each item). And a little gut instinct never hurts.

UNC SCHOOL OF PUBLIC HEALTH

NORTH CAROLINA CENTER FOR PUBLIC HEALTH PREPAREDNESS

The North Carolina Center for Public Health Preparedness is funded by Grant/Cooperative Agreement Number U90/CCU424255 from the Centers for Disease Control and Prevention. The contents of this publication are solely the responsibility o the authors and do not necessarily represent the official views of CDC.