VOLUME 4, ISSUE 1

FOCUS on Field Epidemiology

DISCUSSION QUESTIONS: Advanced Data Analysis: Methods to Control for Confounding (Matching and Logistic Regression)

1. **This issue of FOCUS discussed restriction and stratification as methods to control for confounding. What is the difference between these two methods? Think of situation(s) in which each method should be used.**

Restriction is when only a specified part of the population is studied or used in an analysis. This population segment has uniform characteristics with regard to the confounder. For example, age is a confounder in the relationship between vaccination status and blood antibody titers against the vaccine, because children have a routine vaccination schedule required for school entry. In a population of interest, a survey to determine whether booster shots are needed to keep the population protected might restrict the analysis to only adults, because children and adults will have different levels of immunity due to recent vaccinations. Another example in an outbreak setting might be if students at several different high schools are reported to have a respiratory syndrome, but the majority of cases are at one school. The initial analysis might be restricted to the high school with most of the cases; environmental/ classroom, neighborhood, and socio-economic exposures differ between high schools and could confound the analysis. If the confounder is not allowed to vary in the population, it can have no effect on the results of the analysis.

Stratification is when all segments of the population are examined, but the analysis is performed separately for each level of the confounder. These results can be compared, and can be averaged into one estimate using Mantel-Haenszel methods. Confounders commonly controlled for using stratification methods include gender and race/ethnicity. For example, in a study of dietary habits and a health outcome, you might need to stratify by race/ethnicity, as people from different races or ethnic groups may have different cultural practices in the foods they eat, and could potentially have different rates of health outcomes.

2. Have you been involved in an investigation in which you had to control for several confounders? If so, describe the situation, the investigation, and how you did the analysis.

Pinpointing the problem in many outbreaks can be done without the need to control for many confounders. However, sometimes the source of a disease or commonalities among case-patients can be difficult to discern. Common confounders are age, gender, and other personal characteristics that may be appropriate to the disease, such as occupation, work or leisure activities, etc.

When conducting epidemiologic investigations in non-outbreak settings (i.e., for research or policy analysis purposes), confounding can be more of an issue. If you expect to have confounding in any type of study, it is definitely useful to gather information on the confounder from study participants, and to enroll enough study subjects to have the power to control for the confounder(s) during the analysis.

3. Are you familiar with any software programs that you can use to conduct matched analyses or logistic regression? If so, what are they? If you are familiar with more than one, which do you prefer to use?

Statistical programs you or members of your group may have used include Epi Info, SPSS, Stata, EpiSheet, and SAS, among others. Some of these programs are menu-driven (Epi Info, SPSS), so that analyzing data is a matter of "point and click." Other programs operate in programming language (SAS, Stata), so the user types in the commands for analyzing the data. Most of these programs have many point-and-click operations and the option of using programming language.

For basic analyses, Epi Info is often used since it is free from the CDC and has many useful features such as the ability to create data entry forms and maps.

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.