

## VER - Chapter 13 Sample Problems (Not in Text)

You will use five data sets in these exercises. The following three will be provided as Stata data sets

- pig\_adg.dta (real data)
- nocardia.dta (real data)
- calf.dta (real data)

You can find descriptions of these three datasets in Chapter 27 of the Veterinary Epidemiologic Research.

The fictional data for two other problems (one dealing with hoof angle and lameness in horses and one with Potomac Horse fever) will be provided in paper form and you will have to create the Stata data set.

If you would rather work through the datasets one at a time, do the questions in the following order:

- pig\_adg data - questions # 1, 4, 6, 8
- nocardia data - questions # 3, 5
- calf data - questions # 9, 10
- lameness data - question # 2
- Potomac data - question # 7

### Evaluating Confounding

#### 1. Does atrophic rhinitis increase the risk of pneumonia in pigs?

Use the data set pig\_adg.dta to determine if atrophic rhinitis increases the risk of pneumonia. Since both disease variables were originally recorded on multi-level scales, use the following criteria to dichotomize the variables:

- atrophic rhinitis positive if:  $ar \geq 1$  (ie. scores 1 through 5)
- pneumonia positive if:  $lu \geq 1$  (ie. scores 1 through 3)

Is sex a confounding variable for this relationship?

#### 2. Does hoof angle affect the risk of lameness in horses ?

Within the equine community, it is generally accepted that a low hoof trim angle increases the risk of lameness. You decide that you want to study this so you identify 1000 thoroughbreds (TB) and 1000 standardbreds (STB) that you can study over 1 racing season. You measure the hoof angle on all horses at the start of the season and follow them to determine if they become lame. (Of course all 2000 horses are still in your study at the end of the season and no trainer has changed the hoof trimming angle on any horses - Yeah right !). The data you collect are as follows:

Hoof Angle	Lameness	Breed	Frequency
Low	Lame	STB	194
Low	Lame	TB	6

Low	Sound	STB	706
Low	Sound	TB	94
Natural	Lame	STB	21
Natural	Lame	TB	79
Natural	Sound	STB	29
Natural	Sound	TB	871

What do you conclude about the association between hoof angle and the risk of lameness? Note: You will have to build a small (8 rows, 4 variables) dataset in Stata and then remember to use “frequency weights” (ie. [fweights = ???] ) as part of your command.

If you call the variables: hoof, lame, breed, and count .....  
 cs lame hoof [fweights=count], by(breed)

### 3. Is neomycin use a risk factor for Nocardia mastitis ?

Use the data set nocardia.dta to determine if neomycin is a risk factor for Nocardia mastitis. Is herd size a confounding variable for this relationship? Since herd size was originally recorded as a continuous variable, create a dichotomous variable according to the following definition:

herd is “large” if: numcow >= 50

so that we can use the Mantel-Haenszel procedure to evaluate and control confounding.

### 4. Does pneumonia affect average daily gain in pigs ?

Use the data set pig.dta to determine if pneumonia affects average daily gain (adg). Is sex a confounding variable for this relationship? Note: This problem has a continuous variable as its outcome so you will have to use a regression model to evaluate these relationships (refer to Chapter 14 for the basics of linear regression).

## Evaluating Interaction and Confounding

### 5. More on Nocardia mastitis

Is there evidence of interaction between cloxacillin use and neomycin use (in terms of the effect of neomycin on the risk of Nocardia mastitis)?

### 6. Risk factors for pneumonia in pigs

Go back to the pig.dta data set and determine if there is evidence of interaction between sex and atrophic rhinitis (dichotomized as before) on the risk of pneumonia in pigs.

## 7. Are pastured horses at higher risk of Potomac Horse Fever?

Potomac Horse Fever is an infectious disease caused by a rickettsial parasite. In 1993 it was reported in Ontario but back in the 1980's it was becoming a problem on the eastern seaboard of the United States. At the time, the agent of the disease had just been discovered but its mode of spread was not known. Assume that you were working in that region at the time and wanted to determine if there was any sex predilection and if exposure to pasture increased the risk of the disease. (Specifically, exposure to pasture was defined as > 6hrs per day with substantial forage growth on the pasture). Since many horses that become infected show no clinical signs, you chose to use sero-conversion as an outcome measure. You followed 180 horses on several large breeding farms over the summer of 1987 and recorded the following data.

25 horses seroconvert	- 5 male	- 3 exposed to pasture - 2 not exposed
	- 20 female	- 17 exposed to pasture - 3 not exposed
155 horses remain neg.	- 55 male	- 7 exposed to pasture - 48 not exposed
	- 100 female	- 43 exposed to pasture - 57 not exposed

- Compute the: risk ratio, odds ratio, risk difference, attributable fraction exposed, population attributable risk and attributable fraction population for exposure to pasture (ignoring any effects of sex and variation due to chance). Interpret each of the results this is a review of material from Chapter 4).
- Could pasture use be a confounder for the association between sex and seroconversion?
- Use the Mantel-Haenszel procedure to determine if “exposure to pasture” is a risk factor. Is sex either an interacting or a confounding variable for this relationship?

## 8. Factors that affect ADG in pigs

In the data set pig.dta, is there evidence of any interaction between sex and pneumonia on the average daily gain of pigs? As with question 4, you will have to use a regression model to answer this question.

## 9. Factors that are related to the probability of septicemia

Use the dataset calf.dta and sort out the relationships amongst three potential predictors of septicemia (age, attitude and swollen umbilicus). This will involve evaluating the possibility of confounding and/or interaction among the variables. In working out this question we chose to dichotomise -age- at 0-6 and 7+ days and -attd- between “depressed” and “comatose”

## 10. Interaction in the calf dataset

Consider the factors -age- and -umb- in an evaluation of interaction on both the additive and multiplicative scales. To do this, once again dichotomise -age- at 7 days (ie 0-6 and 7+ days).

- (a) Is there evidence of interaction on a multiplicative scale? What direction is it in (towards antagonism or towards synergism) and is it significant.
- (b) Is there evidence of interaction on an additive scale. Do not worry about whether it is statistically significant or not, just determine if it appears to be present, or not, and what direction it is in.

## Matching

### 11. Using matching to control confounding

Base the following on an observational study that is of interest to you (perhaps even relevant for your research program). Identify the outcome, an important predictor and a potential confounding variable. Use the spreadsheet matching.sxc (or .xls) to answer the following questions.

- (a) Assume that the potential confounder is a true confounder and set up a population structure to reflect that. (Put reasonable values in for your population structure).
  - i. If you do a case-control study without doing a stratified analysis, how serious is the bias in the estimate of the OR?
  - ii. Does stratification over the confounder solve the problem?
- (b) Next, assume that the potential confounder is NOT really a confounder because it turns out that it is not related to the predictor of interest. Adjust your population structure to reflect this fact.
  - i. Is there any bias in the OR from a case-control study if you do not do a stratified analysis?
  - ii. Does stratification over the confounder solve the problem in the case-control study?
  - iii. Is there any bias in the OR from a cohort study if you do not do a stratified analysis?