## Statistical Methods in Epidemiology Lab 9. Case – Control Studies

## I. Matched Case-Control Studies: The Salmonella Typhimurium dataset

In the fall of 1996 an unusually large number of Salmonella Typhimurium cases were recorded in Fyn County in Denmark. The Danish Zoonosis Centre set up a matched case-control study to find the sources. Cases and two age-, sex- and residency-matched controls were telephone interviewed about their food intake during the last two weeks. The data from this study are in the dataframe typhi.dta.

1. Examine the effect of pork using conditional logistic regression.

2. Look through the other variables and find out if any of them have a significant effect on the outcome.

3. You should find that plant7 is a risk factor, while fruit is protective. A question which this finding raises is whether there is any interaction between these two factors. Fit a model with interaction, fill in the table below and interpret the parameters.

log(OR)	fruit	
plant7	0	1
0		
1		

4. One possibility, for example, is that fruit is only protective among subjects who have eaten meat from Plant7. Is this an interaction model? How would you parameterize a model that addresses this question?

5. How does the results from this analysis compare to the ones that you found in question 3? How would you report the results? Fill in the table below.

log(OR)	fruit	
plant7	0	1
0		
1		

## II. Choice of Controls in Case-Control Studies

## Objectives

By the end of the practical students will be able to:

- Identify, for a given case-control study which sampling scheme has been used, and the measure (risk/rate/odds ratio) that is being estimated.
- Identify, for a given case-control study and choice of control group, the advantages and disadvantages of that choice.

1. Read the attached extracts from the paper by Mahmood et al. and discuss the following issues. From what population were the controls recruited? What are the potential advantages and disadvantages of this approach? How was the sampling of controls performed? Which measure of relative incidence will the study estimate? Why do you think this measure was chosen? Does it matter? Why do you think infants "3 months of age and older, with no history of being taken to an MCHC for immunization" were excluded from the cases (criterion(e))?

Read the attached extracts from the paper by Mueller et al. (J Natl Cancer Inst, 1987, 1-5). This study used siblings (brothers and sisters) of cases and controls. What are the potential advantages and disadvantages of this approach?

3. Read the attached extracts from the paper by Collins et al. (Lancet, 1999, 353, 693-697). Why do you think the investigators chose to use random telephone controls? What biases might this method introduce?