Chapter 9 Exercises

Part I

1. An intrepid epidemiology student was examining statistics from the National Highway Traffic Safety Administration when she made a startling discovery: smoking among drivers appeared to be associated with fatal car accidents. She gathered 2009 data from Florida and arranged it in the 2x2 table below:

	Fatal Accident	Nonfatal Accident	Total
Smoking	1289	164723	166012
No Smoking	1269	688421	689690
Total	2558	853144	855702

Calculate and interpret the risk ratio and risk difference based on the above table.

2. With such compelling results, the student immediately began thinking of reasons for the observed association. Could it be that people who smoked were more distracted? That they tended to drive with only one hand on the wheel and were at greater risk of losing control? That they had reduced visibility because of the smoky car interior? Her mentor had another idea, however, and suggested that she consider whether a third variable, such as alcohol consumption, might be driving the results. The student first considered whether the driver's testing positive for blood alcohol at the scene of the accident was associated with smoking in her data:

	Smoking	No Smoking	Total
Alcohol	124337	335475	459812
No Alcohol	41675	354215	395890
Total	166012	689690	855702

Calculate and interpret the risk ratio and risk difference based on the above table.

3. She then considered whether testing positive for alcohol was associated with being in a fatal car crash in her data:

	Fatal Accident	Nonfatal Accident	Total
Alcohol	1639	458173	459812
No Alcohol	919	394971	395890
Total	2558	853144	855702

Calculate and interpret the risk ratio and risk difference based on the above table.

4. Do you think that the association between smoking and fatal car accidents originally observed by the student may have been confounded by alcohol? Explain.

Part II

You undertake an investigation to assess the effects of hormone replacement therapy (HRT) on coronary heart disease (CHD). You conduct a cohort study whereby you follow women with no history of CHD for ten years. Assume complete follow-up on all women.

We will assume that the data table below is the TRUTH (i.e., not what you measured in your study, but what you would have measured if you had conducted the study perfectly and there were no non-comparability). In the absence of any misclassification of exposure status, this is what you would have observed.

	CHD+	CHD-	Total
HRT+	200	5110	5310
HRT-	170	5259	5429
Total	370	10369	10739

1. Calculate and interpret the risk ratio and risk difference based on these data.

Misclassification Scenario 1:

Women were asked to recall their HRT exposure experience; recall is almost never perfect. Twenty percent of women who actually had taken HRT said that they hadn't, whereas ten percent of women who actually had not taken HRT said that they had. This occurred regardless of disease status.

2. Use the table below to calculate the 2x2 table with misclassification.

	CHD+	CHD-	Total
HRT+			
HRT-			
Total			

Final table with misclassification:

	CHD+	CHD-	Total
HRT+			
HRT-			
Total			

3. What are the risk ratio and risk difference in the misclassified table? Calculate and interpret.

4. Compare the risk ratio in question 3 to the true risk ratio in question 1. How did misclassification of the exposure independent of disease status (non-differential misclassification of exposure) affect the estimates in the study?

Misclassification Scenario 2:
Twenty percent of the nonusers with CHD were categorized as users. Everyone else was classified accurately according to the TRUTH table above.

5. Is this an example of misclassification of exposure or disease?

6. Fill in the table below with the misclassified 2x2 table.

	CHD+	CHD-	Total
HRT+			
HRT-			
Total			

Table with misclassification:

	CHD+	CHD-	Total
HRT+			
HRT-			
Total			

7. What are the risk ratio and risk difference in the misclassified table? Calculate and interpret.

8. Compare the risk ratio in question 7 to the true risk ratio in question 1. How did this misclassification affect the estimates in the study?

Part III

1. In a prospective study of depression and dementia, you recruit participants who are between ages 65 and 75 and not suffering dementia. At intake, you give them a screening test for depression. Five years later, you give them the Mini Mental State Examination, which assesses cognitive function. There is no loss to follow-up. Your observed risk ratio is 2.40: elderly people who are depressed have 2.40 times the risk of dementia compared with elderly people who are not depressed over 5 years.

   a. If, instead of perfect follow-up, you ended up with 20% fewer participants after 5 years. It turns out that depressed people were more likely to drop out than non-depressed. What would the effect be on your risk ratio?

   b. Alternatively, what would be the effect on the risk ratio if the 20% of participants lost to follow-up were all people who were depressed and demented?

Part IV

What are some strategies that may be deployed in the design and data collection phases of epidemiologic studies to minimize misclassification?