

Data analysis and interpretation - Assignment

Part I

The following questions are based on Rosenberg et al., "Oral contraceptive use in relation to nonfatal myocardial infarction". *Am J Epidemiol* 1980; 111:59-66.

1. Control for menopausal status appears to have been accomplished through: (Choose one)
 - A. Restriction
 - B. Matching plus stratified analysis
 - C. Stratified analysis without matching
 - D. Mathematical modeling (logistic regression)
2. Using the data in Table 1, label and complete a 2x2 table for the crude (i.e., not stratified by age of hospitalization) odds ratio for MI and current (versus "never") OC use.
3. Compute the odds ratio for the above table (show all work).
4. Is age a confounder of the relationship between MI risk and current OC use, based on the data in Table 1? Justify your answer (1-3 sentences), referring to specific measures or estimates in the data.
5. Is age an effect modifier of the relationship between MI risk and current OC use, based on the data in Table 1? Justify your answer (1-3 sentences).
6. Based on the data in Table 4:
 - a. Is MI associated with cigarette smoking? Give a sentence to support your answer (yes, no, cannot determine from the data in Table 4).
 - b. Is the association between MI and the combination of smoking and hypertension greater than or less than the association between MI and hypertension alone? Give a sentence to support your answer (greater than, less than, cannot determine from the data in Table 4).
7. Using the data in table 4, create and label a 2x2 table relating current OC use and hospitalization for MI among nurses who have no history of hypertension, regardless of their smoking status.
8. What is the relative risk estimate for women who have all three characteristics (OC, CIG, HYP) compared to women who have none?
9. Assuming that the relative risk estimates in table 4 are precise (i.e., ignoring the variability from small cell sizes), what would be the relative risk estimate for women who have all three characteristics (OC, CIG, HYP) compared to women with none under a multiplicative model? Show your work.

Part II

1. In table 2 from the Rosenberg et al. study (p. 62), which variables are not found to be risk factors for MI?
2. Show the relationship between the logistic regression coefficient for Current OC use and the relative risk estimate for that factor.
3. Has age been controlled in this logistic model?
4. Comparing the information in Table 2 with that in Table 1, do you see evidence of confounding in Table 1 with respect to the relationship between MI risk and OC use? Briefly discuss, citing the most relevant measures or statistics from the two tables.
5. What is peculiar about this logistic regression model, in terms of the form of the variables in it?
6. What provision has been made in this model for possible statistical interaction, i.e., deviation from a multiplicative model?
7. Based on the logistic model shown, what would be the odds ratio for the combination of both cigarette smoking and Current OC use (versus neither)? Compare this result to the corresponding figure(s) in the stratified analysis in Table 4 and suggest possible explanations for the difference, if any.