Lab: A Matter of Interpretation

Cervical Intraepithelial Neoplasia: HPV or Something Else?

Read

1. The National Cancer Institute fact sheet on "Human Papillomaviruses and Cancer." http://www.cancer.gov/cancertopics/factsheet/Risk/HPV/
2. Schiffman, M. H., Bauer, H. M., Hoover, R. N., Glass, A. G., Cadell, D. M., Rush, B. B., et al. (1993). Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. *Journal of the National Cancer Institute, 85*(12), 958-964.

Questions

Cervical intraepithelial neoplasia (CIN) is a precursor of invasive cancer of the cervix. By definition, CIN is limited to the epithelial lining of the cervix, the external entrance to the uterus. In the typical Pap smear, cells are obtained from the cervix and are examined under the microscope for evidence of atypical cells or for clearly abnormal cells classified into CIN-1, CIN-2, CIN-3, depending on the degree of abnormality. An additional procedure, namely a cervicovaginal lavage, is required to obtain specimens for HPV testing, as described in the journal article to be discussed.

NOTE: The relative risks presented in this article are actually odds ratios. Recall that odds ratios from case-control studies are estimates of relative risks under certain conditions.

1. Crude odds ratios

a. Use the data in **Table 1** to construct a 2-by-2 table to calculate the crude odds ratio Types 16 or 18 HPV and CIN. Use the "negative" HPV category as the nonexposed group. Keep straight the number of exposed cases, nonexposed cases, exposed controls, and nonexposed controls as you build your 2-by-2 table. Show all work.

b. Use the data in **Table 4** of the article to construct a 2-by-2 table exposure to *Types 16 or 18 of HPV* and of *CIN1*.

2. Study design analysis. Appraise the article by Schiffman *et al.* in relation to these five elements of epidemiologic study design:

a. Research hypotheses: Are the research questions clear? Are they relevant? Do they follow logically from what is known scientifically about the exposure and outcome, based on the existing literature?
b. Study design: Is this study design experimental or observational? Is the sample cohort or case-control? Is this design appropriate in light of past research, the research question and the nature of the disease and exposures?

c. Outcome variable: What is the primary outcome variable in this study? Is it relevant? How is it being defined and measured? How accurate is the outcome/disease measurement?
d. Exposure variable: What are the main exposure variable*s* in this study? Are they relevant? How are they being measured? With what level of accuracy? How is exposure quantified: how valid is the cutoff point for distinguishing exposed from unexposed? Are biological markers used to define exposure or is it self-report, medical records etc.?
e. Analysis: Does the analysis address the research question? Is the analysis appropriate for the study design and type of data collected? Do the analysis and presentation provide information on the precision of estimates?

3. Confirmation of cases. Schiffman and coworkers made considerable efforts to review the original cytological diagnoses that served as the basis for defining cases. What type of bias do these efforts seek to avert?

4. Possible nondifferential misclassification of cases. The majority (*n* = 319) of the 500 cases were defined as having *condylomatous atypia* which the authors considered to be borderline rather than definite cases of CIN (see p959-c2- para5). It is possible that some of these borderline cases were noncases that have been misclassified as cases. HPV status could influence the accuracy of diagnosing these borderline lesions if the clinicians making the diagnoses were aware of exposure status or if exposure status affects the appearance of lesions in a way that makes them more susceptible to being misdiagnosed. Assuming that the diagnoses were made without knowledge of the exposure status of the women and that no direct (biologic) effect of HPV on diagnostic accuracy exists, then the misclassification should be nondifferential. What is the likely effect of this type of nondifferential misclassification on the estimate of relative risks calculated in Table 1?

5. Adjustment for HPV test results (Table 2). In the comparison of the RR #1 (age adjusted) and RR #2 columns (adjusted for all the other risk factors in the table) in Table 2 (p. 961), what potential bias are the authors addressing? What conclusion can you draw when you compare the the effects of smoking when you compare RR #1 and RR #2?

Clues for interpreting Table 2

- The authors present various *RR* for the association of cervical cancer and different risk factors (number of sexual partners, age at first intercourse, etc.). Three columns of *RR*s are presented. Each column has been adjusted for different confounding factors. Read the notes below the table carefully.

- The column for *RR* #1 is adjusted for age only. This column indicates, for example, that women with 10+ lifetime sexual partners have 4.4 times the risk of developing cancer as compared to women who had 1 lifetime sexual partner. (The reference group is women with one lifetime sexual partner; i.e., the *RR* for the baseline (referent) group is always 1.

- The column for *RR* #2 adjusts for additional confounders (e.g., lifetime number of sexual partners, age at first intercourse, etc.). This column indicates that adjusting for these

confounders does not very much alter the *RR* for women with 10+ partners vs. 1 partner. *RR*#2 = *RR*#1 ≈ 4.4.

- The column for *RR* #3 adjusts for age and HPV infection. The *RR* for 10+ sex partners vs. 1 partner in this column is 1.8, indicating a much smaller increase in risk after adjusting for HPV.

6. Increasing lifetime number of sex partners. Compare in Table 2 the analysis for lifetime number of sex partners using RR #1 (adjusted for age) and RR#3 (adjusted for age and HPV status). Write a short paragraph (under 100 words) explaining the differences between these two columns, and how this analysis sheds light on the effect of lifetime number of sex partners.

7. Ancillary Analyses. [Thought question.] At the bottom of page 960, the authors describe two ancillary analyses based on subsets of cases. What is the purpose of sub-setting the data in this way? Do the results for these subsets increase your confidence in the validity of the overall conclusion? Why or why not?

8. Is the number of sex partners a risk factor for CIN *independent* of HPV risk? Schiffman *et al.* argue (on 961-1-2) that multivariate analyses including both lifetime numbers of sex partners and HPV test results pointed to HPV infection as the primary risk factor for each of the three categories of cases. What data in Table 4 support this argument? Explain.

9. Assessing causality. In their discussion section on page 962, the authors argue that the HPV association with CIN satisfies all of the accepted criteria for assessing causality. Which of these criteria are strongly satisfied and which somewhat weakly in the evidence discussed by Schiffman *et al.*? Assess strength, biological gradient (dose-response_, temporality, specificity, consistency, and plausibility.

10. Persistent errors in HPV classification. [Thought question.] Explain Schiffman *et al.*'s statement in 962-2-4 that even though some risk factors persist among HPV-negative women, this finding could result from errors in HPV measurements. Suggest some ways in which HPV measurement errors could be reduced.