

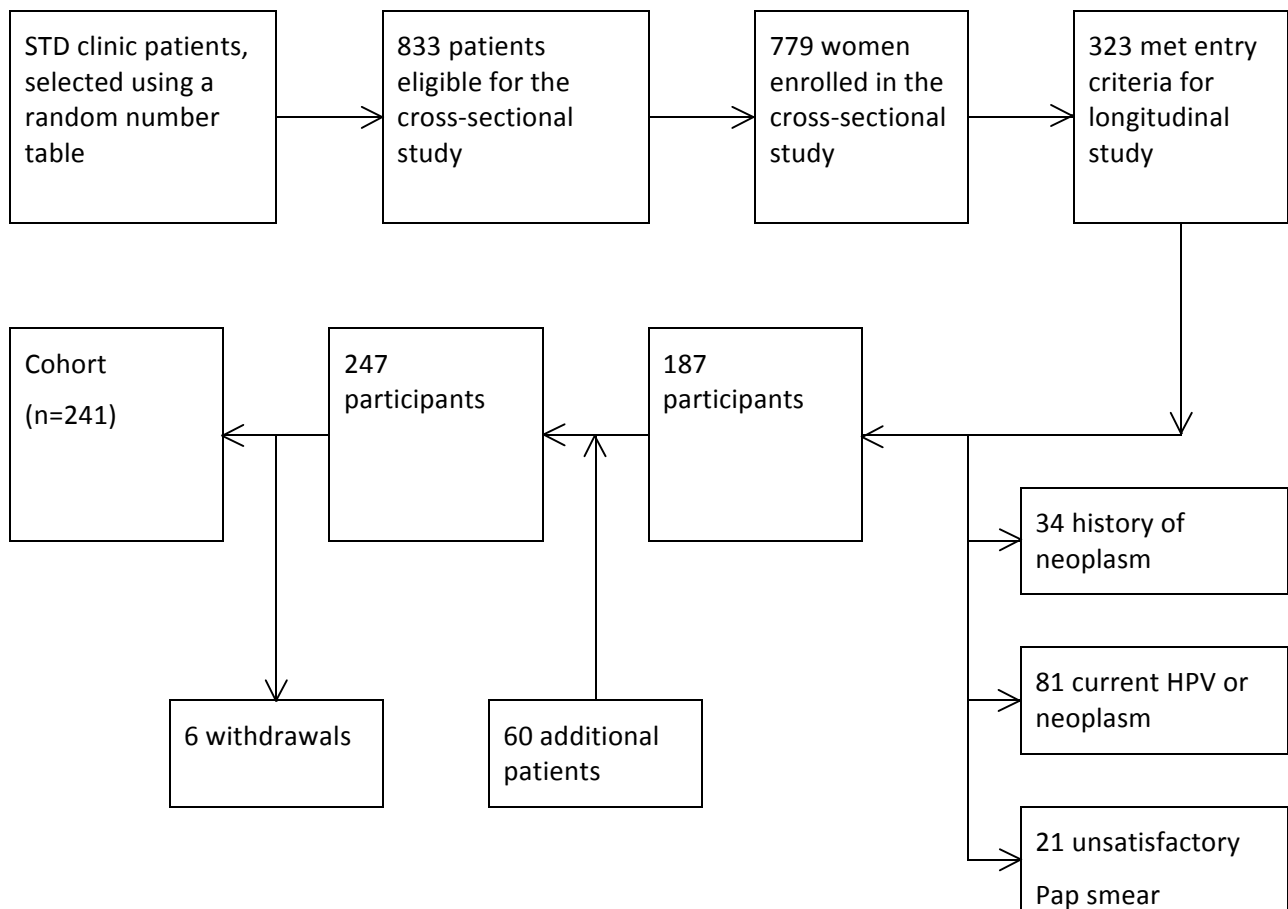
## Measuring Disease and Exposure HPV and CIN

**Objectives:** To identify the exposures and outcomes in a study; calculate incidences in subgroups, and compare the incidences in the form of common measures of association and impact.

**Acknowledgement:** This exercise is based on an exercise created by Vic Schoenbach of the UNC School of Public Health (see footer). An earlier version was developed by Jo M. Harter, M.D., Ph.D.

**Read:** Read Koutsky LA, Holmes KK, Critchlow CW, et al. A cohort study of the risk of cervical intraepithelial neoplasia grade 2 or 3 in relation to papillomavirus infection. *New England Journal of Medicine* 1992;327:1272-8. Full text: [www.sjsu.edu/faculty/gerstman/eks/Koutsky1992.pdf](http://www.sjsu.edu/faculty/gerstman/eks/Koutsky1992.pdf).

Focus initially on the Abstract, Introduction, and Methods sections on your initial reading. Here is a flow diagram showing the complex construction of the cohorts for this study:



### Questions

1. What is the primary explanatory variable (“exposure”) in this study?
2. What is the primary response variable (“disease”)? What is its significance for health?
3. What *specific* issues are addressed by this study? Why are these important?

4. How is the exposure variable defined and measured in this study?
5. How is the disease variable defined and measured in this study?
6. The authors use the term *cumulative incidence*. This term is synonymous with *incidence proportion* and *average risk*. On what basis are data that are reported characterized as *incidences*?
7. Complex survival analysis methods were used to calculate measures of occurrence and effect in the original article. However, we can get reasonably good estimates by addressing these simplified data:

	Biopsy CIN+	Biopsy CIN-	Total
HPV +	24	86	110
HPV -	4	127	131
Total	28	213	241

- a. What is the incidence proportion (“cumulative incidence”) of CIN in HPV+ group?
  - b. What is the incidence proportion in HPV- group?
  - c. Using these data, what is the *RR* associated with HPV exposure? What does this statistic measure, in plain terms?
  - d. Based on these data, what is the attributable fraction in exposed cases (i.e.,  $AF_e$ )? Interpret this statistic.
  - e. The attributable fraction in the population ( $AF_p$ ) can be estimated with these data if and only if we assume the proportion of cases that are exposed to the risk factor ( $p_c$ ) in the sample is representative of the prevalence of exposure in the general population. Let us make this assumption. Then, using formula 8.16, determine the  $AF_p$ . Interpret this statistic.
8. The incidence *rate* is difficult to determine under the circumstances described in the article. Normally, we would consider different lengths of follow-up for individuals in the cohorts. When individual follow-up time is not available, we must make certain assumptions about follow-up time. For learning purposes, let us consider two different assumptions that allow us to estimate person-time and hence.
- a. **Method 1:** Each case contributed maximal person-time before succumbing to illness. This method has each subject contributed the full follow-up period for the cohort. Since the average follow-up time for this study was 2 year, the 110 individuals in the HPV+ sub-cohort contributed  $110 \times 2$  years = 220 person-years of follow-up. Use this method to calculate the incidence rate of CIN in the HPV+ sub-cohort. In addition, calculate the rate in the HPV- sub-cohort.
  - b. **Method 2:** Each case contributes half available person-time before succumbing to the condition. This “modified-life table” approach assumes the average time to illness is half the follow-up interval. By this assumption, each non-case in our cohort contributes the full 2 person-years of observation-time and each case contributes half that, or 1 person-year. For example, in the exposed (HPV+) sub-cohort, the 24 cases contribute 1 person-year each (24 person-years of follow-up), while the 86 non-cases contribute  $86 \times 2$  years = 172 person-years of follow-up. Therefore, the exposed subcohort contributed  $24 + 172 = 196$  person-years at risk. Using this method, calculate the rates of CIN in the HPV+ sub-cohort, and in the HPV- sub-cohort.