***Case‑control Studies in Epidemiology***

The *Case‑control Studies* lab uses the well‑established association between lung cancer and past smoking to illustrate characteristics of case‑control investigations. It uses descriptions and data from three case‑control investigations to elaborate their primary strengths (cost‑effectiveness) and weakness (effects of bias). The lab illustrates the evolution of case‑control investigations, and demonstrates the persistence of underlying strengths and limitations of this investigative method.

Epidemiologic terms and concepts emphasized in the *Case‑control Studies* laboratory include: *sources of cases and controls, representativeness versus comparability, various sources of systematic error (bias), interview methods, information measured in case‑control studies, distinction between incident and prevalent cases, the requirement for confirmation of disease diagnosis, misclassification, statistical significance, chi‑squared, p‑value, measurement of the magnitude of an association using the odds ratio, and use of multiple comparison (control) groups*.

LABORATORY EXERCISE: CASE‑CONTROL STUDIES

*Cancer of the Lung and Cigarette Smoking*

Edited by John Morgan and Raymond Knutsen

Suggestions that tobacco might be associated with lung cancer had been proposed by clinicians as far back as 1912. However, most medical practitioners first became aware of the relationship when the *Journal of the American Medical Association* published two papers on this subject in May of 1950; one by *Wynder* and *Graham* (1) and the other by *Levin*, *Goldstein* and *Gerhardt* (2). The *British Journal of Medicine* published a similar report by *Doll* and *Hill* (3) on September 30 of the same year.

Both studies evaluated differences in past tobacco use among lung cancer cases as compared to persons not having lung cancer (controls). Although each identified convincing associations between lung cancer and past smoking, all of the results were susceptible to the effects of bias (systematic error). A number of refinements have been made in case‑control methodology and interpretation, yet systematic error remains the greatest obstacle encountered in these studies.

The following laboratory assignment includes a brief description of the study design and a summary of results for each of the early case‑control studies which first linked lung cancer and smoking. These landmark investigations represented the first widespread use of case‑control studies and illustrate the evolution of understanding regarding this research method.

**I. *Wynder and Graham Study***

Wynder and Graham (1) studied 684 cases (644 males and 40 females) of microscopically confirmed bronchogenic carcinoma (ICD‑9, code 162). The patients came from hospitals and private practices of physicians in California, Colorado, the District of Columbia, Illinois, Maryland, Massachusetts, Michigan, Missouri, New Jersey, New York, Ohio, Pennsylvania, and Utah. A standard questionnaire was administered to all cases and included questions on history of other lung disease, smoking, occupation, exposure to dust and fumes, alcohol intake, residence, education, and cause of death of deceased parents and siblings. Personal interviews, apparently conducted by the physician authors of the report, were obtained in 634 cases, with a mailed questionnaire used for 33 cases. Exposure information was obtained from a person who had known the lung cancer patient intimately throughout his/her adult life for 17 cases.

Two non‑medical investigators interviewed 780 male and 552 female non‑lung cancer patients on the general surgical and medical services at three hospitals in St. Louis: Barnes Hospital, Jefferson Barracks Veterans Hospital, and St. Louis City Hospital. The methods used to interview the controls are uncertain.

The age‑adjusted data on smoking habits for the 605 male lung cancer patients and the 780 male controls are presented in Table 1.

**II. *An Attempt to Minimize Bias***

The non‑medical colleagues of Wynder and Graham interviewed an additional series of 100 lung cancer patients and 186 controls having other chest diseases. The interviews were conducted with no previous knowledge of the patients' diagnoses. Both groups of patients were interviewed on the chest services of 4 hospitals in St. Louis. The age‑adjusted results are given in Table 2.

**III. *Doll and Hill ‑ Urban Hospitals***

Doll and Hill (3) studied 709 patients with carcinoma of the lung, 637 patients with carcinoma of the stomach, colon or rectum, and 709 non‑cancer control patients. The study was conducted in 20 London hospitals which notified the investigators of all patients admitted with the relevant carcinomas. Four social workers interviewed the patients using a questionnaire. For each lung cancer patient, the social worker was instructed to interview, as a control, the first patient on the ward lists of the same sex, within the same five‑year age‑group, having a diagnosis other than cancer. For all patients, the final diagnosis on discharge was confirmed by a review of medical records.

The authors reported that the diagnosis in 489 of the lung cancer patients was confirmed by microscopic examination of tissue obtained by autopsy, biopsy, or exploratory surgery. Diagnosis was confirmed using clinical, laboratory and radiologic criteria for the remaining 220 patients.

Doll and Hill found a statistically significant association between lung cancer and tobacco use. More than 99.6% of the 649 male lung cancer patients reported a history of smoking as compared to 95.8% of the controls. These data are presented in Table 3. A non‑significant association between lung cancer and social classes IV and V (semiskilled and unskilled laborers) was also reported. Limited to the 551 case‑control pairs seen at the hospitals having special chest surgery or radiotherapy facilities, Doll and Hill observed a significant association between lung cancer and rural residence.

 Of the male lung cancer patients, 26% were smoking 25 cigarettes or more per day prior to their illness, as compared with 14% of the matched controls. Similar differences were found for the average amount of tobacco ever consumed (total amount/years of use) and for the estimate of life‑time total amount of tobacco consumed. These differences were reported to be "statistically significant." The lung cancer group included a lower proportion of pipe smokers than the control group, although the difference was not statistically significant. The smoking histories of the patients with cancer of the stomach, colon or rectum were similar to those observed for the non‑cancer controls.

Similar to the observation for males, only 31.7% of the 60 female lung cancer patients were non‑smokers as compared with 53.3% of the matched controls. This result was significant at the p=0.02 level.

*Systematic Error Which May Be Acceptable*

Upon checking discharge diagnoses, Doll and Hill discovered that 209 of the patients who had been interviewed as cases were incorrectly identified on admission records as having cancer of the lung. The major reason for this misclassification occurred when 147 persons at one hospital who had undergone bronchioscopy, were mistakenly reported as having bronchogenic carcinoma. The smoking histories of these patients were found to differ significantly from the histories of the 709 lung cancer cases, but not from the controls. This systematic error was detected and corrected after publication of the results presented in Tables 4‑6.

**IV. *An Attempt To Minimize Interview Bias***

Levin, Goldstein and Gerhardt (2) reviewed the records of the Roswell Park Memorial Institute in Buffalo, a state cancer hospital in which a history of tobacco usage has been taken from all patients admitted since 1938. These histories are part of the regular clinical record and are taken before the patient is seen by a physician; before a diagnosis is established.

The age‑standardized proportions of cigarette smokers were as follows: *lung cancer* 236 cases (66%); *other cancer* 666 cases (48%); *lung non‑tumors* 124 cases (53%); and *other non‑cancer* 481 cases (44%). The differences between the lung cancer patients and each of the other groups were statistically significant. This relationship was not found for pipe or cigar smoking.

One of the primary reasons for conducting case‑control studies is an attempt to isolate and quantify the degree of association between a particular exposure factor and disease. This is usually accomplished by arranging the data collected in an investigation into a 2x2 contingency table (Figure 1). This table facilitates comparison of the proportion of cases having past exposure to a defined factor with a similar measure made for controls. The ratio computed is termed the *relative odds*, or *odds ratio* (OR). Although the retrospective nature of case‑control studies precludes the direct computation of incidence, the odds ratio provides an unbiased estimation of the relative risk (risk ratio) for rare diseases.

Figure 1. Contingency table frequently used for case‑control study results.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Disease status | |  |
|  | Cases (Disease +) | Controls (Disease -) | TOTAL |
| Exposed | a | b | a+b |
| Not exposed | c | d | c+d |
| Total | a+c (total # of cases) | b+d (total # of controls) |  |

where: a & c equal either the absolute number or proportion (percent) of cases in each exposure category and b & d equal the number or proportion of controls in each exposure category.

The disease odds ratio is defined as the ratio of the disease odds for persons having exposure (a/b) to the disease odds for persons not having exposure (c/d).

a/b

OR = = ad/bc

c/d

In order to estimate relative risk, an important requirement is that the disease studied be rare. The rate of disease must be low within the population from which cases and controls were derived. This assumption is met for lung cancer and is also true for most other chronic diseases.

Tables 4, 5 and 6 present the major findings from the Doll and Hill case‑control study.

*References*

1. Wynder E, Graham E. Tobacco smoking as a possible etiologic factor in bronchogenic carcinoma. A study of six‑hundred and eighty‑four proved cases. *Journal of the American Medical Association* 143(4):329‑36, May 27, 1950.

2. Levin M, Goldstein H, Gerhardt P. Cancer and tobacco smoking: A preliminary report. *Journal of the American Medical Association* 143(4):336‑38, May 1950.

3. Doll R. Hill A.B. Case‑control study of lung cancer and smoking. *British Journal of Medicine*, September 30, 1950.

Students may access articles through the Loma Linda University Library:

[http://www.llu.edu/library/index.page?rsource=library.llu.edu/](http://www.llu.edu/library/index.page?rsource=library.llu.edu/ )