Department of Epidemiology
School of Public Health
University of North Carolina

EPID 160 | Principles of Epidemiology | Fall 1967

CASE HISTORY AND COHORT STUDIES

Faculty Guide to Laboratory Exercise IV

Page 1, Table 1

The subjects with lung cancer have a smaller proportion of non-smokers and lighter smokers and a larger proportion of heavy smokers than the subjects without lung cancer.

Possible bias:

The lung cancer patients tended, because of their disease, to exaggerate their smoking habits.

Page 1, Table 2

The results cannot be attributed to the exaggeration of the smoking habits of the lung cancer patients since subjects incorrectly classified as lung cancer patients reported smoking habits similar to subjects without lung cancer.

Page 2, Tables 2 and 3

<u>Case History Study</u>

Ages 39-49 — Coronary Heart Disease Present

- Type A: 41/60 = .68 (68% of CHD patients were Type A)
- Type B: 19/60 = .32 (32% of CHD patients were Type B)

Coronary Heart Disease Absent

- Type A: 1196/2416 = .495 (49.5% of non-CHD patients were Type A)
- Type B: 1220/2416 = .505 (50.5% of non-CHD patients were Type B)

Ages 50-59 — Coronary Heart Disease Present

- Type A: 39/53 = .74
- Type B: 14/53 = .26

Coronary Heart Disease Absent

- Type A: 577/995 = .58
- Type B: 418/995 = .42

Total Subjects	Coronary Heart Disease Present
Type A	80/113 = .71
Type B	33/113 = .29
	Coronary Heart Disease Absent
Type A	1773/3411 = .52
Type B	1636/3411 = .48

Cohort Study

Ages 39-49	Type A	
CHD Present	45/1117 = .0403	(403 CHD per 10,000 Type A or 4% of Type A had CHD)
CHD Absent	1072/1117 = .9597	
	Type B	
CHD Present	18/1204 = .0149	(149 CHD per 10,000 Type B)
CHD Absent	1186/1204 = .9850	
Ages 50-59	Type A	
CHD Present	49/579 = .0846	
CHD Absent	530/579 = .9154	
	Type B	
CHD Present	21/415 = .0506	
CHD Absent	394/415 = .9494	
Total Subjects	Type A	
CHD Present	94/1696 = .0554	
CHD Absent	1602/1696 = .9446	
	Type B	
CHD Present	39/1619 = .0241	
CHD Absent	1580/1619 = .9759	

Case History Study: Behavior pattern Type A was found to be associated with CHD. Among the total subjects, 80 or 70.9 percent of the 113 subjects with CHD were classified as Type A behavior pattern. Whereas, 1773 or 52.0 percent of the 3411 subjects without CHD were classified as Type A and the remaining 48.0 percent as Type B.

Cohort Study: In subjects already classified as Type A behavior pattern, the CHD rate is 94/1696 or 5.5 percent. Conversely, in subjects already classified as Type B behavior pattern, the CHD rate is 39/1619 or 2.4 percent.

Page 3, Table 4

Source: Epidemiological Study of Cancer and Other Chronic Diseases. National Cancer Institute, Monograph 19, U. S. Department of Health, Education, and Welfare, Public Health Service, National Cancer Institute, Bethesda, Maryland, p. 145.

Relative risk of coronary heart disease for cigarette smokers only as compared to persons who have never smoked:

For Males

Ages 45-54	422/150 = 2.81
Ages 75-84	3871/3132 = 1.24

The risk of coronary heart disease is 2.8 times as great for male cigarette smokers between 45-54 years of age as for non-smokers the same age.

For Females

Ages 45-54	66/33 = 2.00
Ages 75-84	2349/1973 = 1.19

The attributable risk of coronary heart disease for cigarette smokers only as compared to persons who have never smoked:

For Males

Ages 45-54	422 - 150 = 272
Ages 75-84	3871 - 3132 = 739

Assuming that causes of coronary heart disease other than smoking had equal effect on the exposed and unexposed groups, stopping smoking would save 272 cases per 100,000 men.

For Females

Ages 45-54	66 - 33 = 33
Ages 75-84	2349 - 1973 = 376

Page 4, Table 5

The 10-year incidence study gives different results than the prevalence studies. The incidence study shows a rise of risk with increasing cholesterol levels. However, the prevalence studies do not show such a trend.

Selective forces may be responsible for the failure of the prevalence results to agree with the incidence study. Cholesterol may be related to the risk of rapidly fatal CHD and therefore such cases would not be included in a prevalence study.

Friedman, et. al., describe the test of this hypothesis as follows:

> To test this hypothesis, the incidence study was repeated using only the cases that survived through Exam VI. When this was done cholesterol was still associated with CHD incidence. Thus the failure of the Exam VI prevalence study to show a risk gradient for cholesterol could not be attributed to a selective removal of fatal cases from the study population. Even when such cases were removed from the incidence data, the association persisted.
>
> There was further evidence that the factors selecting subjects out of the study population for inclusion in the Exam VI prevalence study were not responsible for the lack of the cholesterol rask gradient. The Exam VI prevalence study of cholesterol was repeated, this time classifying subjects by their initial cholesterol level. In this case there was an association between serum cholesterol and CHD. If the Exam VI prevalence study had been lacking those individuals in the population in whom cholesterol was associated with CHD, it was not expected that classification of the Exam VI group by an earlier cholesterol value would have resulted in an apparent association between cholesterol and CHD.
>
> This difference in results using initial and later cholesterol levels, suggested that the change in serum cholesterol with time was different in persons with and without CHD. The mean levels of serum cholesterol and Exams I and VI were compared in men according to whether or not they had CHD at Exam VI. This comparison was made on the 1,768 men in the Exam VI prevalence study who had cholesterol determinations at both Exam I and VI. The mean cholesterol was substantially higher at Exam I in the cases (243.6 milligrams per 100cc) than in the controls (222.0 milligrams per 100cc). However, in the ten years the mean cholesterol level had risen 19.5 milligrams per 100cc in the controls, but only 2.5 milligrams per 100 cc in the cases so that by Exam VI they were about equal (241.5 milligrams per 100 cc respectively). Thus no risk gradient would be demonstrable.

However, the main point to be taught by this data is the possibility of selections.

Page 5, Table 6

In analyzing this prospective study, there is the problem of persons being eliminated from the study for many reasons. In the calculation of rates, the correct denominator must be computed. Possible methods are the person-years and the life table techniques.

The following tables show the calculation of both person-years and the life table.

Calculation of Person-Years

Year X to X+1	Alive and Normal at End of Year l_x	New Entries n_x	Person-Years of Observation L_x	CHD Rate (Percent) m_x
0- 1	0	1013	474.0	0.0
1- 2	948	900	1379.5	0.217
2- 3	1811	0	1797.5	0.445
3- 4	1784	0	1767.5	0.679
4- 5	1751	0	1734.0	0.865
5- 6	1717	0	1697.0	1.237
6- 7	1677	0	1667.0	0.600
7- 8	1657	0	1642.5	1.218
8- 9	1628	0	1619.0	0.679
9-10	1610	0	1601.0	0.812
10	1592			
Total				0.758

Life Table for Calculation of Probability of CHD By the End of Any Specified Year

Year X to X+1	CHD Rate m_x	Probability of CHD (Cases & Deaths) q_x	Number at Risk of CHD at Beginning of Year l_x	Number of Cases and Deaths of CHD During Year $(l_x \times q_x)$	Number of Cases and Deaths of CHD at Beginning of Year $1000 - l_x$
0- 1	0.0000	0.0000	1000*	0.0	0.0
1- 2	0.0022	0.0020	1000.0	2.0	0.0
2- 3	0.0045	0.0043	998.0	4.3	2.0
3- 4	0.0068	0.0066	993.7	6.6	6.3
4- 5	0.0087	0.0085	887.1	7.5	112.9
5- 6	0.0124	0.0122	779.6	9.5	220.4
6- 7	0.0060	0.0058	770.1	4.5	229.9
7- 8	0.0122	0.0120	465.6	5.6	534.4
8- 9	0.0068	0.0066	460.0	3.0	540.0
9-10	0.0081	0.0080	457.0	3.7	543.0

* 1000 is an arbitrary number with which the table started.

Formulae used for Person-Year Method:

w_{x1} = CHD at intake

w_{x2} = Deaths from other causes

w_{x3} = Lost to observation

c_x = Cases of CHD

d_x = Deaths from CHD

n_x = New entries

L_x = Person-years of observation

l_x = Number under observation at beginning of period

m_x = CHD rate

q_x = The probability of CHD (cases and deaths)

(1) $l_{x+1} = l_x + n_x - (w_{x1} + w_{x2} + w_{x3}) - c_x - d_x$

(2) $L_x = l_x + \frac{1}{2}n_x - \frac{1}{2}(w_{x1} + w_{x2} + w_{x3}) - \frac{1}{2}c_x - \frac{1}{2}d_x$

(3) $m_x = c_x + d_x/L_x \times 100$

Formula used for Life Table Method:

(1) $q_x = 2m_x/2 + m_x$

Examples for Person-Years Method:

(1) $l_{(1-2)} = 0 + 1013 - (41 + 2 + 22) - 0 - 0$

$= 1013 - 65$

$= 948$

(2) $L_{(1-2)} = 948 + \frac{1}{2}(900) - \frac{1}{2}(23 + 1 + 10) - \frac{1}{2}(3) + \frac{1}{2}(0)$

$= 948 + 450 - 17 - 1.5$

$= 1379.5$

(3) $m_{(1-2)} = 3/1379.5 \text{ X } 100$

$= 0.217$

Examples for Life Table Method:

(1) $q_{(1-2)} = 2(0.0022)/2 + 0.0022$

$= 0.0020$

Number during year = $(q_x\ l_x)$ for year (1-2)

0.002 (1000) = 2.0

Number at beginning of year = 1000 - l_x for year (1-2)

1000 - 1000 = 0.0

Pages 6 and 7, Tables 7, 8, 9, and 10

Number of Deaths from Lung Cancer Among Smokers and Non-Smokers That Would be Expected to Occur in the Sample in 10 Years if There was no Difference in the Mortality Rates for Smokers and Non-Smokers
(Figures are Rounded)

Ages	Males		Females	
	Smokers	Non-Smokers	Smokers	Non-Smokers
35-44	3	1	0	1
45-54	16	4	1	6
55-64	37	12	1	3
65-74	25	13	1	9

Examples:

Males 35-44

Number of Smokers = (percent of white male smokers for age 35-44) X (number of males in sample age 35-44)

Number of White Male Smokers = (80.5) (4000) = 3220

Expected Deaths from Lung Cancer Among Smokers = (lung cancer mortality rate for white males 35-44) X (number of white male smokers age 35-44)

Expected Number of Deaths = (82.0) (3220)

Expected Number of Deaths = 2.6

Note the fact that although a sample of 35,800 person were selected, there are cells in which zero or one death are expected to occur.