Centers for Disease Control and Prevention
Epidemiology Program Office
Case Studies in Applied Epidemiology
No. 811-703

Oral Contraceptive Use and Ovarian Cancer

Student's Guide

Learning Objectives
After completing this case study, the participant should be able to:

☐ Outline the sequence of an epidemiologic analysis;

☐ Discuss the biases of particular concern in case-control studies and ways to minimize their influence;

☐ Describe why and when to use crude and adjusted odds ratios and 95% confidence intervals, and how to interpret them; and

☐ Define and recognize effect modification and confounding.

This case study was developed by Richard Dicker and Peter Layde in 1981. Current version updated by Richard Dicker with input from the EIS Summer Course instructors.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service

PART I

In 1980, ovarian cancer ranked as the fourth leading cause of cancer mortality among women in the United States. An estimated 18,000 new cases and more than 11,000 attributable deaths occurred among American women that year.

Several studies had noted an increased risk of ovarian cancer among women of low parity, suggesting that pregnancy exerts a protective effect. By preventing pregnancy, oral contraceptives (OCs) might be expected to increase the risk of ovarian cancer. On the other hand, by simulating pregnancy through suppression of pituitary gonadotropin release and inhibition of ovulation, OCs might be expected to protect against the subsequent development of ovarian cancer. Because by 1980 OCs had been used by more than 40 million women in the United States, the public health impact of an association in either direction could be substantial.

To study the relationship between oral contraceptive use and ovarian cancer (as well as breast and endometrial cancer), CDC initiated a case-control study – the Cancer and Steroid Hormone (CASH) Study in 1980. Case-patients were enrolled through eight regional cancer registries participating in the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute.

Question 1: Which investigations need to be reviewed by an institutional review board? Does this investigation need to be reviewed?

As the investigators planned this study, they discussed a variety of methods to minimize potential biases.

Question 2: What types of bias are of particular concern in this case-control study? What steps might you take to minimize these potential biases?

As the investigators began to consider what data to collect with their questionnaire, they began to lay out the analyses they wanted to conduct. They did so by sketching out "table shells" -- frequency distributions and two-way tables that contain no data but otherwise include appropriate titles, labels, measures and statistics to be calculated. The tables followed a logical sequence from the simple (descriptive epidemiology) to the more complex (analytic epidemiology) that is often used when results are presented in a manuscript or oral presentation.

Question 3: List, in logical sequential order, the table shells you might use to analyze or present the CASH study data.

PART II

The study design included several features to minimize selection and information bias. Ascertainment bias of disease status – a type of selection bias – was minimized by attempting to enroll as cases all women ages 20-54 years with newly diagnosed, histologically confirmed, primary ovarian cancer who resided in one of the eight geographic areas covered by the cancer registries. Controls were women ages 20-54 years selected randomly using telephone numbers from the same geographic areas. Because 93% of U.S. households had telephones, virtually all women residing in the same areas as the cases were eligible to be controls. (Interestingly, all the women enrolled with ovarian cancer had telephones.)

To minimize interviewer bias, CDC investigators conducted group sessions to train interviewers in the administration of the pretested standard questionnaire. The same interviewers and questionnaires were used for both cases and controls. Neither cases nor controls were told of the specific a priori hypotheses to be tested by the study. Recall bias of oral contraceptive exposure was minimized by showing participants a book with photographs of all OC preparations ever marketed in the United States and by using a calendar to relate contraceptive and reproductive histories to other life events.

The primary purpose of the CASH study was to measure and test the association between OC use and three types of reproductive cancer – breast cancer, endometrial cancer, and ovarian cancer. Enrollment of subjects into the study began in December 1980. During the first 10 months of the study, 179 women with ovarian cancer were enrolled, as well as larger numbers of women with endometrial or breast cancer. During the same period, 1,872 controls were enrolled to equal the number of subjects with breast cancer. The same control group was used for the ovarian cancer analysis; however, the investigators excluded 226 women with no ovaries at the time of interview and four controls whose OC use was unknown, leaving 1,642 women to serve as controls. The distribution of exposure to OCs among cases and controls is shown in Table 1.

Table 1. Ever-use of oral contraceptives among ovarian cancer cases and controls, Cancer and Steroid Hormone Study, 1980-1981

		CASE-CONTROL STATUS		
		Case	Control	Total
USE OF OCs	Ever	a = 93	b = 959	H_1 = 1052
	Never	c = 86	d = 683	H_0 = 769
	Total	V_1 = 179	V_0 = 1642	T = 1821

Question 4: From these data, can you calculate the risk of ovarian cancer among oral contraceptive users? Why or why not?

Question 5: Describe the rationale behind using the odds ratio as an estimate of the risk ratio. When is the odds ratio not an appropriate estimate of the risk ratio?

The investigators used the data in Table 1 and the formulas shown below to calculate an odds ratio, a Mantel-Haenszel Chi, and 95% test-based confidence limits.

Measure	Formula	Calculation from Table 1
Odds Ratio (OR)	$OR = \frac{ad}{bc}$	$\frac{(93)(683)}{(959)(86)} = 0.77$
Expected Value of Cell 'a' (E(a))	$E(a) = \frac{H_1 V_1}{T}$	$\frac{(1052)(179)}{1821} = 103.41$
Mantel-Haenszel (MH) Variance	$MH\ variance = \frac{H_1 H_0 V_1 V_0}{T^2(T-1)}$	$\frac{(1052)(769)(179)(1642)}{(1821)^2(1821-1)} = 39.40$
Mantel-Haenszel Chi*	$X_{MH} = \frac{a - E(a)}{\sqrt{MH\ variance}}$	$\frac{93 - 103.41}{\sqrt{39.40}} = -1.66$
Test-Based Confidence Limits (Note: Z values for 2-sided confidence limits are: 90% = 1.645, 95% = 1.96, 99% = 2.58.)	$LOWER\ CL = OR^{(1-(Z/X_{MH}))}$ $UPPER\ CL = OR^{(1+(Z/X_{MH}))}$	$0.77^{(1-(1.96/1.66))} = 0.57$ $0.77^{(1+(1.96/1.66))} = 1.05$

* The Mantel-Haenszel Chi with one degree of freedom (X_{MH}) is equivalent to a "Z score" and may be used to find the 2-tailed p value from a table of areas in two tails of the standard normal curve. In this case, p = 0.097.

Question 6: What special information does the odds ratio give that you do not get from Chi square and p value? What additional information do you get from the p value and Chi square? From a confidence interval?

Question 7: How might you describe and interpret these results?

In many epidemiologic studies, age is a confounding factor.

Question 8: What is confounding? Under what circumstances would age be a confounder in this study?

PART III

In the analysis of use of oral contraceptives and ovarian cancer, age was related both to OC use and to case-control status. (OC users were younger than never-users; case-patients were younger than controls.) Therefore, the investigators decided to stratify the data by age and calculate stratum-specific and, if appropriate, summary statistics of the stratified data. The Mantel-Haenszel (MH) procedure is a popular method for calculating a summary odds ratio and test of significance for stratified data.

Question 9: What is stratification? Why stratify data? How do you decide on which variables to stratify?

Question 10: What is effect modification? How do you look for it?

Question 11a: Using the data in Table 2, calculate the odds ratio for the 40- to 49-year age stratum.

Question 11b: Using the data in Table 2, calculate the expected value of cell A for the 40- to 49-year age stratum.

Table 2. Ever-use of oral contraceptives and risk of ovarian cancer, stratified by age, Cancer and Steroid Hormone Study, 1980-1981

Ages 20-39 years

	Case	Control	Total	
Ever user	46	285	H_1 = 331	OR = 0.69
Never user	12	51	H_0 = 63	Expected(a) = 48.73
Total	V_1 = 58	V_0 = 336	T = 394	MH variance = 6.66
				MH Chi = -1.06
				95% CLs = 0.34, 1.38

Ages 40-49 years

	Case	Control	Total	
Ever user	30	463	H_1 = 493	OR = ____
Never user	30	301	H_0 = 331	Expected(a) = ____
Total	V_1 = 60	V_0 = 764	T = 824	MH variance = 13.39
				MH Chi = ____
				95% CLs = 0.38, 1.10

Ages 50-54 years

	Case	Control	Total	
Ever user	17	211	H_1 = 228	OR = 0.61
Never user	44	331	H_0 = 375	Expected(a) = 23.06
Total	V_1 = 61	V_0 = 542	T = 603	MH variance = 12.91
				MH Chi = -1.69
				95% CLs = 0.34, 1.08

Question 11c: Using the data in Table 2, calculate the Mantel-Haenszel chi for the 40- to 49-year age stratum.

The investigators had been taught to look for effect modification before looking for confounding.

Question 12: Do you think age is an effect modifier of the oral contraceptive and ovarian cancer association?

The investigators concluded that age was not an effect modifier. They therefore decided to control for confounding by calculating an odds ratio adjusted for age, also called a summary odds ratio or Mantel-Haenszel odds ratio, using the following formula:

$$OR_{MH} = \frac{\Sigma\ (ad/T)}{\Sigma\ (bc/T)}$$

They also calculated a Mantel-Haenszel chi, from which they found a p-value. Finally, they calculated a 95% confidence interval of 0.45 to 0.92.

Question 13a: Using the stratified data in Table 2, calculate the summary odds ratio adjusted for age.

Question 13b: Based on the Mantel-Haenszel chi of -2.55 and the attached table of the standard normal curve, determine the 2-tailed p-value.

Question 14: In terms of the null hypothesis and statistical significance, what do you infer from the p-value? What do you infer from the confidence interval of 0.45–0.92?

Question 15: Do you think age is a confounding variable in this analysis of the association between OC use and ovarian cancer?

Question 16: What are the other ways of eliminating confounding in a study?

In the introduction to this case study, pregnancy was described as apparently protective against ovarian cancer. The investigators were interested in seeing whether the association between OC use and ovarian cancer differed for women of different parity. Table 3 shows parity-specific data.

Table 3. Ever-use of oral contraceptives and risk of ovarian cancer, by parity*, CASH Study, 1980-1981

Parity	Use of OCs	# Case-patients	# Controls	Age-adjusted odds ratios (95% confidence intervals)
0	Ever user	20	67	0.3 (0.1-0.8)
	Never user	25	80	
1-2	Ever user	42	369	0.8 (0.4-1.5)
	Never user	26	199	
≥3	Ever user	30	520	0.7 (0.4-1.2)
	Never user	35	400	

* Excludes seven controls (four never-users and three ever-users) and one case (ever-user) with unknown parity.

Question 17: Is there any evidence for effect modification in the data presented in Table 3?

AREAS IN TWO TAILS OF THE STANDARD NORMAL CURVE

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	1.000	0.992	0.984	0.976	0.968	0.060	0.952	0.944	0.936	0.928
0.1	0.920	0.912	0.904	0.897	0.889	0.881	0.873	0.865	0.857	0.849
0.2	0.841	0.834	0.826	0.818	0.810	0.803	0.795	0.787	0.779	0.772
0.3	0.764	0.757	0.749	0.741	0.734	0.726	0.719	0.711	0.704	0.697
0.4	0.689	0.682	0.674	0.667	0.660	0.653	0.646	0.638	0.631	0.624
0.5	0.617	0.610	0.603	0.596	0.589	0.582	0.575	0.569	0.562	0.555
0.6	0.549	0.542	0.535	0.529	0.522	0.516	0.509	0.503	0.497	0.490
0.7	0.484	0.478	0.472	0.465	0.459	0.453	0.447	0.441	0.435	0.430
0.8	0.424	0.418	0.412	0.407	0.401	0.395	0.390	0.384	0.379	0.373
0.9	0.368	0.363	0.358	0.352	0.347	0.342	0.337	0.332	0.327	0.322
1.0	0.317	0.312	0.308	0.303	0.298	0.294	0.289	0.285	0.280	0.276
1.1	0.271	0.267	0.263	0.258	0.254	0.250	0.246	0.242	0.238	0.234
1.2	0.230	0.226	0.222	0.219	0.215	0.211	0.208	0.204	0.201	0.197
1.3	0.194	0.190	0.187	0.184	0.180	0.177	0.174	0.171	0.168	0.165
1.4	0.162	0.159	0.156	0.153	0.150	0.147	0.144	0.142	0.139	0.136
1.5	0.134	0.131	0.129	0.126	0.124	0.121	0.119	0.116	0.114	0.112
1.6	0.110	0.107	0.105	0.103	0.101	0.099	0.097	0.095	0.093	0.091
1.7	0.089	0.087	0.085	0.084	0.082	0.080	0.078	0.077	0.075	0.073
1.8	0.072	0.070	0.069	0.067	0.066	0.064	0.063	0.061	0.060	0.059
1.9	0.057	0.056	0.055	0.054	0.052	0.051	0.050	0.049	0.048*	0.047
2.0	0.046	0.044	0.043	0.042	0.041	0.040	0.039	0.038	0.038	0.037
2.1	0.036	0.035	0.034	0.033	0.032	0.032	0.031	0.030	0.029	0.029
2.2	0.028	0.027	0.026	0.026	0.025	0.024	0.024	0.023	0.023	0.022
2.3	0.021	0.021	0.020	0.020	0.019	0.019	0.018	0.018	0.017	0.017
2.4	0.016	0.016	0.016	0.015	0.015	0.014	0.014	0.014	0.013	0.013
2.5	0.012	0.012	0.012	0.011	0.011	0.011	0.010	0.010	0.010	0.010
2.6	0.009	0.009	0.009	0.009	0.008	0.008	0.008	0.008	0.007	0.007
2.7	0.007	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.005	0.005
2.8	0.005	0.005	0.005	0.005	0.005	0.004	0.004	0.004	0.004	0.004
2.9	0.004	0.004	0.004	0.003	0.003	0.003	0.003	0.003	0.003	0.003
3.0	0.003	0.003	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002
3.1	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001
3.2	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.3	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.4	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.0005
3.5	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003	0.0003
3.6	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	0.0002	0.0002
3.7	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
≥3.9	<0.0001									

* Use this table to find the 2-tailed p value which corresponds to a Z score or Chi (square root of chi-square) with 1 degree of freedom. For a given value of Z or chi (say, 1.98), find that value to 1 decimal place in the left-most column (1.9). The p value will be in the 1.9 row. Now find the second decimal of your Z or chi across the top row (0.08). The p value is in that column. The 2-tailed p value is at the intersection of the row and column you've identified (for 1.9 and 0.08, $p_2 = 0.048$).

To find the p value for a chi-square with 1 degree of freedom (including any chi-square from a simple 2-by-2 table, the McNemar chi-square from a matched 2-by-2 table, and the Mantel-Haenszel chi-square from stratified 2-by-2 tables), simply take the square root of the chi-square, then proceed as above.

PART IV - CONCLUSION

In their published report, the investigators wrote the following about the possible effect modification by parity:

"Parity appeared to be an effect modifier of the association between oral contraceptive use and the risk of ovarian cancer...[Table 3]. Among nulliparous women, the age-standardized odds ratio was 0.3 (95% confidence interval: 0.1-0.8). Among parous women, however, the odds ratios were closer to, but still less than, 1.0....It is possible, therefore, that oral contraceptives are most protective for women not already protected by pregnancy."

Although this case study deals with the data collected over the first 10 months (phase 1) of the study, an additional 19 months of data (phase 2) were collected and analyzed subsequently. The following table summarizes the apparent role of parity as an effect modifier in the two phases of the study.

Table 4. Age-adjusted odds ratios (aOR) and 95% confidence intervals for the association of oral contraceptive use and ovarian cancer, by parity and phase of study, CASH Study, 1980-1982

Parity	Phase 1 (months 1-10) aOR	(95% CI)	Phase 2 (months 11-29) aOR	(95% CI)	Total (months 1-29) aOR	(95% CI)
0	0.3	(0.1-0.8)	0.7	(0.5-1.2)	0.7	(0.4-1.0)
1-2	0.8	(0.4-1.5)	0.5	(0.3-0.7)	0.5	(0.4-0.8)
≥3	0.7	(0.4-1.2)	0.5	(0.4-0.8)	0.6	(0.4-0.8)
Total	0.6	(0.4-0.9)	0.5	(0.4-0.7)	0.6	(0.5-0.7)

On the basis of the full study results, it appeared that the effect of oral contraceptives on ovarian cancer was not substantially different for nulliparous women and for parous women.

Although there were no published studies of oral contraceptives and ovarian cancer when this study was launched, there were several by the time this study was published. Almost all showed an apparently protective effect of oral contraceptives on ovarian cancer.

Question 18: What are the public health and/or policy implications of the apparently protective effect of oral contraceptives on ovarian cancer?

References – CASH Study

1. Centers for Disease Control. Oral contraceptive use and the risk of ovarian cancer: the Centers for Disease Control Cancer and Steroid Hormone Study. *JAMA* 1983;249:1596-9.

2. Centers for Disease Control. The reduction in risk of ovarian cancer associated with oral contraceptive use: the Cancer and Steroid Hormone Study of the Centers for Disease Control and the National Institute of Child Health and Human Development. *N Engl J Med* 1987;316:650-5.

References - Data Analysis, Confounding, Effect Modification

3. Dicker RC. Analyzing and Interpreting Data. In: Gregg MB. *Field Epidemiology, 2nd edition*. New York: Oxford University Press, 2002.

4. Rothman KJ. *Epidemiology: an introduction.* New York: Oxford University Press, 2002.