

STATISTICAL ANALYSIS OF PATIENT-CONTROL STUDIES IN EPIDEMIOLOGY

FACTOR UNDER INVESTIGATION AN ALL-OR-NONE VARIABLE

M. C. PIKE, Ph.D.*

AND

R. H. MORROW, M.D.

*Department of Preventive Medicine, Makerere Medical School,
P.O. Box 7072, Kampala, Uganda*

THERE is considerable confusion in the medical literature as to the correct statistical method to apply to results obtained in retrospective studies in which patients with a particular disease are *individually* matched to control subjects.

For example, in a recent paper Vessey and Doll (1968) reported on their investigations of the relation between the use of oral contraceptives and thromboembolic disease. Fifty-eight married patients admitted to hospital with deep vein thrombosis were individually matched for age, parity, and date of admission with two married control patients admitted to the same hospital for an acute surgical or medical condition. They arranged their results as follows:

Diagnostic Group	Oral Contraceptives		All Women
	Used (no. of patients)	Not Used	
Thromboembolism	26	32	58
Control	10	106	116
Both groups	36	138	174

and tested them by chi-squared in the usual 2×2 table manner ($\chi^2 = 28.7$, $n = 1$, $P < 0.001$).

But, as Vessey and Doll noted in their paper, such amalgamation of results ignores the individual matching between affected and control patients; and this is only legitimate if the probability of taking oral contraceptives is independent of the matching variables age, parity, and date of admission, that is, if matching is irrelevant.

*On secondment from the Medical Research Council's Statistical Research Unit, London
Present address: Department of the Regius Professor of Medicine, Radcliffe Infirmary, Oxford

A correct approach to the analysis of investigations in which there is a single control for each patient is known (McNemar, 1947; Billewicz, 1964) and is developed systematically by Stuart (1957), Cox (1958), and Mantel and Haenszel (1959). In the next section we present, and justify, from a slightly different viewpoint this test for the single control situation.

In the following sections we extend the test to the situation with an arbitrary number of controls. This extension is a special case of the general test given by Mantel and Haenszel (1959) and may also be derived using Cox's (1958) arguments.

SINGLE CONTROL

Suppose there are n_1 patient plus control pairs, then, if the factor under study is an all-or-none variable (K, not-K), the raw results fall into three sets as follows:

Set (no. of Ks)	Patient plus Matched Control	No. of Sets	No. of Patients with K
0	2 not-K	$n_{1,0}$	$m_{1,0} = 0$
1	1 K, 1 not-K	$n_{1,1}$	$m_{1,1}$
2	2 K	$n_{1,2}$	$m_{1,2} = n_{1,2}$
Total		n_1	

These figures are often arranged in the following tabular form:

Factor	Patients	Matched Controls	Total
K	$W = m_{1,1} + n_{1,2}$	$X = (n_{1,1} - m_{1,1}) + n_{1,2}$	$R_1 = n_{1,1} + 2n_{1,2}$
not-K	$Y = n_{1,0} + (n_{1,1} - m_{1,1})$	$Z = n_{1,0} + m_{1,1}$	$R_2 = 2n_{1,0} + n_{1,1}$
Total	n_1	n_1	$2n_1$

And to test for statistical significance of association between the factor and the disease under study a standard 2×2 table chi-squared test on 1 degree of freedom (corrected for continuity),

$$\chi^2 = \frac{2 (|WZ-XY| - n_1)^2}{n_1 R_1 R_2} \dots\dots\dots (1)$$

is then applied. This is incorrect, as we stated above, because it ignores the individual matching of patients and controls and may lead to considerable loss of power (see below).

The common correct test (McNemar's test) of association concentrates solely on Set (1) pairs and ignores Set (0) and Set (2) pairs.

Set (1) pairs have had only one person exposed to K. Under the null hypothesis of no association, there should be an equal number of these pairs with the patient having K and with the control having K; *i.e.*, on the null hypothesis the expected value, $E(m_{1,1})$ of $m_{1,1}$ is $\frac{1}{2}n_{1,1}$, and its variance, $V(m_{1,1})$, is $\frac{1}{4}n_{1,1}$. The test statistic is then

$$\chi^2 = \frac{(|m_{1,1} - E(m_{1,1})| - \frac{1}{2})^2}{V(m_{1,1})} = \frac{(|2m_{1,1} - n_{1,1}| - 1)^2}{n_{1,1}} \dots\dots\dots (2)$$

which is distributed approximately as chi-squared on 1 degree of freedom.

Mantel and Haenszel (1959) develop this test by considering each of the n_1 pairs as a separate 2×2 table:

Set (0) pair

	K	not-K	Total
Patient	0	1	1
Control	0	1	1
Total	0	2	2

Set (1) pair

	K	not-K	Total
Patient	1 (0)	0 (1)	1
Control	0 (1)	1 (0)	1
Total	1	1	2

Set (2) pair

	K	not-K	Total
Patient	1	0	1
Control	1	0	1
Total	2	0	2

Formula (2) is a summary chi-squared for these n_1 tables. The tables formed by Set (0) and Set (2) pairs have zero variability when they are considered as having fixed marginal totals as is usually done,

and the summary chi-squared is therefore based solely on Set (1) pairs.

Ignoring Set (0) and Set (2) pairs can also be justified intuitively by noting that a retrospective matched pair study can only provide evidence of association between a disease and a given factor if it is carried out in a population where the factor is present in some people but not all. Now Set (0) pairs provide no evidence that they are not from a population of 100% not-K; and *vice versa* for Set (2) pairs. They should therefore be omitted.

For example, consider Vessey and Doll's (1968) study of thromboembolic disease and oral contraceptives. Suppose that it is carried out in three areas, area A where no women take the pill, area B where some do and some do not, and area C where all do, and the results are combined. There is 'overmatching' in areas A and C; area B is the *only* area that can contribute to our knowledge of whether taking oral contraceptives increases the risk of thromboembolic disease. In this example, the nature of the problem, *viz.*, results being highly correlated with area, would have been noticed by an observant investigator. In general, however, correlations of this type may easily be overlooked (see, for example, Pike, Morrow, Kisuule, and Mafigiri, p. 39 in this issue).

The unpaired 2×2 table analysis χ^2 , formula (1), will be greater than the paired χ^2 , formula (2), if $4n_{1,0}n_{1,2} < n_{1,1}^2$, and *vice versa*.

For example, if $n_{1,1} = 20$, $m_{1,1} = 16$, the paired chi-squared $\chi^2 = 6.05$ ($P = 0.01$); and the following table shows the value of the unpaired chi-squared for a range of values of $n_{1,0}$ and $n_{1,2}$: ($n_{1,0}$ assumed equal to $n_{1,2}$ for convenience):

$n_{1,0} = n_{1,2}$	Unpaired χ^2
0	14.40
10	6.05
20	4.03
30	3.03
100	1.10

We see, therefore, that when there is considerable overmatching (*i.e.*, $n_{1,0} \times n_{1,2}$ 'large') the unpaired χ^2 may be very much less than the paired χ^2 .

TWO CONTROLS

When each patient is individually matched with two controls, the natural extension of the above approach may be adopted.

Suppose that there are n_2 triples of patient plus two controls, then the raw results fall into four Sets as follows:

Set	Patient plus Matched Controls	No. of Sets	No. of Patients with K
0	3 not-K	$n_{s,0}$	$m_{s,0}=0$
1	1 K, 2 not-K	$n_{s,1}$	$m_{s,1}$
2	2 K, 1 not-K	$n_{s,2}$	$m_{s,2}$
3	3 K	$n_{s,3}$	$m_{s,3}=n_{s,3}$
Total		n_s	

Set (0) and Set (3) triples do not contribute to any test of association.

Set (1) triples each have one person being K and two being not-K. On the null hypothesis of no association $m_{s,1}$ should be, subject only to random (binomial) variation, $\frac{1}{3}$ rd of $n_{s,1}$, i.e., the expected value, $E(m_{s,1})$, of $m_{s,1}$ is $n_{s,1}/3$, and its variance, $V(m_{s,1})$, is $2n_{s,1}/9$. Similarly, $E(m_{s,2}) = 2n_{s,2}/3$ and $V(m_{s,2}) = 2n_{s,2}/9$. The test statistic (corrected for continuity) is then

$$\chi^2 = (| \frac{\{m_{s,1} - E(m_{s,1})\} + \{m_{s,2} - E(m_{s,2})\}}{\{V(m_{s,1}) + V(m_{s,2})\}} | - \frac{1}{2})^2 \quad (3)$$

which is again distributed approximately as chi-squared on 1 degree of freedom. This summary chi-squared can be shown to be 'optimal' in a certain sense (Cox, 1958).

c CONTROLS

The above generalizes immediately to the situation in which each patient is individually matched with c controls.

The raw results fall into $(c + 2)$ Sets as follows:

Set	Patient plus Matched Controls	No. of Sets	No. of Patients with K
0	$(c+1)$ not-K	$n_{c,0}$	$m_{c,0}=0$
1	1 K, c not-K	$n_{c,1}$	$m_{c,1}$
\vdots		\vdots	\vdots
i	i K, $(c+1-i)$ not-K	$n_{c,i}$	$m_{c,i}$
\vdots		\vdots	\vdots
$c+1$	$(c+1)$ K	$n_{c,c+1}$	$m_{c,c+1}=n_{c,c+1}$
Total		n_c	

On the null hypothesis

$$E(m_{c,i}) = n_{c,i} \times i / (c + 1)$$

$$V(m_{c,i}) = n_{c,i} \times i \times (c + 1 - i) / (c + 1)^2$$

And the summary test statistic is

$$\chi^2 = (|\sum m_{c,i} - \sum E(m_{c,i})| - \frac{1}{2})^2 / \sum V(m_{c,i}) \quad \dots (4)$$

which is again distributed approximately as chi-squared on 1 degree of freedom.

If patients do not all have the same number of controls then form raw result tables as above for each subset of patient plus controls with $c = 1, 2, \dots, T$ ($T =$ maximum number of controls for any patient). The summary chi-squared test on 1 degree of freedom is then

$$\chi^2 = (|\sum_{j=1}^T \sum m_{jt} - \sum \sum E(m_{jt})| - \frac{1}{2})^2 / \sum \sum V(m_{jt}) \quad (5)$$

DISCUSSION

This test given by formula (5) above is easy to apply and involves very little increase in computational effort *vis-à-vis* the usual incorrect 2×2 table approach. There is thus no possible reason for not testing for statistical significance by a correct method.

This test is certainly not new. The purpose of this note is to draw epidemiologists' attention to it.

REFERENCES

BILLEWICZ, W. Z. (1964). Matched samples in medical investigations. *Brit. J. prev. soc. Med.*, **18**, 167.

COX, D. R. (1958). Two further applications of a model for binary regression. *Biometrika*, **45**, 562.

MCNEMAR, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, **12**, 153.

MANTEL, N., and HAENSZEL, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *J. Nat. Cancer Inst.*, **22**, 719.

STUART, A. (1957). The comparison of frequencies in matched samples. *Brit. J. statist. Psychol.*, **10**, 29.

VESSEY, M. P., and DOLL, R. (1968). Investigation of relation between use of oral contraceptives and thromboembolic disease. *Brit. med. J.*, **2**, 199.