

## ATTRIBUTABLE RISK PERCENT IN CASE-CONTROL STUDIES

PHILIP COLE AND BRIAN MacMAHON

*Department of Epidemiology, Harvard School of Public Health, 665 Huntington Avenue, Boston, Massachusetts 02115*

The difference between the disease rates of two groups, one exposed and one not exposed to a suspected aetiological factor, is a useful measure of the risk associated with the factor. This measure is of most interest when the association is considered causal, and has been termed the attributable risk (MacMahon and Pugh, 1970). Expressed as a percentage of the total disease rate in a group (the attributable risk percent), the measure describes the percentage of the group's total risk which is in excess of the risk among persons not exposed to the suspect factor. Both measures (attributable risk and attributable risk percent) may be computed either for persons exposed to the suspect factor or for the population as a whole.

Both measures are readily obtained when estimates of absolute disease rates are available for the groups to be compared. However, in many case-control studies subjects are not referable to an enumerated population, and, while relative risks can be obtained, absolute disease rates can not. If an estimate of the disease rate in the population is available from some other source, disease rates in exposed and non-exposed persons, and hence the attributable risks, may still be estimated (MacMahon and Pugh, 1970). It appears not to have been recognized, however, that even in the absence of knowledge of the disease rate in the population, estimates of attributable risk per cent can be obtained—both for exposed persons and for the population as a whole.

When estimates of absolute disease frequency are available the attributable risk percent among the exposed is determined as:

$$(1) \quad A_e\% = \frac{R_e - R_o}{R_e} \times 100\%$$

where  $A_e\%$  = attributable risk percent among exposed

$R_e$  = absolute risk among exposed

$R_o$  = absolute risk among unexposed.

The estimate of the attributable risk percent is unaltered if the data are reduced from absolute to relative frequencies of disease. Assigning the value 1, as is conventional, to the risk among the unexposed, expression (1) may be rewritten as:

$$(2) \quad A_e\% = \frac{R - 1}{R} \times 100\%$$

where  $R$ , the relative risk, is the ratio of the risk among the exposed to that among the unexposed.

The numerator of (2), termed the excess relative risk, is the segment of the relative risk among the exposed which exceeds the risk among the unexposed. Since relative risk reflects total risk, expressing a group's excess relative risk as a percentage of its relative risk yields the attributable risk percent. Expression (2) applies then to relative risk estimates such as are available from the usual case-control study. For example, if a group of exposed persons has a relative risk of 2.0, the attributable risk percent is 50%; if the relative risk is 3.0, the attributable risk percent is 67%, etc.

If, in addition to knowing the relative risk among the exposed, the proportion of the 'population' exposed can be estimated, the population attributable risk percent can also be obtained. ('Population' refers here to those persons who, although not enumerated, form the common source of cases and controls in a study.) Since controls should be representative of unaffected persons in the population, the data available from them can be used to estimate the proportion of the population exposed. The population attributable risk percent is then determined as:

$$(3) \quad A_p\% = \frac{P_e(R - 1)}{1 + P_e(R - 1)} \times 100\%$$

where  $A_p\%$  = population attributable risk percent

$P_e$  = proportion of the population exposed.

The denominator in expression (3) represents the total disease experience of the population and has two components. The first, 1, is the relative amount of disease not associated with exposure. This component represents the total population multiplied by the relative disease frequency in the absence of exposure; both of these values and hence their product are always 1. The second component of the denominator is the relative amount of disease which is associated with exposure. This is expressed as the proportion of the population exposed multiplied by

their excess relative risk; the value is the same as the numerator.

The attributable risk percent among the population is entirely analogous to that among the exposed. A population with no exposed members may be said to have a relative risk of 1. If some of the population become exposed, the population in the aggregate will exhibit higher risk. The amount of increase depends both on the proportion of the population exposed and on their relative risk. The proportions of the population unexposed and exposed can be estimated by data from the controls. The relative risk for each group is also known, being 1 among the unexposed and the estimated value among the exposed. The 'population relative risk' is then determined as the weighted average of the relative risks for the two groups comprising the population:

$$(4) \quad R_p = \frac{P_o(1) + P_e R}{P_o + P_e}$$

where  $R_p$  = population relative risk  
 $P_o$  = proportion of the population unexposed.

Expression (4) is readily simplified to:

$$(5) \quad R_p = P_o + P_e R$$

By subtracting 1 from the population relative risk, it is converted to the population excess relative risk. The population attributable risk percent can then be determined as was the attributable risk percent among the exposed:

$$(6) \quad A_p\% = \frac{R_p - 1}{R_p} \times 100\%$$

Data from a case-control study of bladder cancer and cigarette smoking (Cole, Monson, Haning, and Friedell, 1971) illustrate the method. For male cigarette smokers the observed relative risk of bladder cancer was 1.89 as compared to a risk of 1 for non-smokers. The attributable risk percent among male smokers can be determined as:

$$A_e\% = \frac{R - 1}{R} \times 100\% = \frac{1.89 - 1}{1.89} \times 100\% = 47\%$$

Based on the distribution among the controls, it was determined that 0.73 of the male population were smokers. The attributable risk percent among men can then be determined as:

$$A_p\% = \frac{P_e(R - 1)}{1 + P_e(R - 1)} \times 100\% = \frac{0.73(1.89 - 1)}{1 + 0.73(1.89 - 1)} \times 100\% = 39\%$$

The method provides only estimates of the attributable risk percent. The estimates should be meaningful, however, if three considerations are met. First, the relative risk has been referred to as if it were the usual measure of risk obtained in a case-control study. In fact, the usual measure is the relative odds or odds ratio. However, since most diseases are uncommon the relative odds is generally equivalent to the relative risk. Second, the relative risk, and hence the corresponding excess relative risk, should be controlled for the influence of other known or suspected aetiological agents which overlap in distribution with the exposure of interest. This consideration applies to all measures of risk derived from epidemiological studies.

The third and major consideration is that controls usually will have been stratified to correspond to the cases in one or more characteristics and thus will not be representative of the population. The stratification will introduce error if the proportion of the population exposed, or the associated relative risk, differs between categories of the stratification variable. This error can be minimized and summary estimates of the attributable risk percent obtained. For the population, the attributable risk percent is first determined in each category by the method presented. The summary estimate is

TABLE  
 SEVERAL MEASURES OF BLADDER CANCER RISK, ACCORDING TO SINGLE AND MULTIPLE EXPOSURE CATEGORIES, MEN AGED 20-89

High Risk Occupation*	Cigarette Smoking†	No. of Cases	No. of Controls	Relative Risk (R)	Excess Relative Risk (R-1)	Annual Rate per 100,000	Attributable Annual Rate per 100,000
No	No	43	94	1.00	0.00	20.3	0.0
No	Yes	173	189	2.00	1.00	40.6	20.3
Yes	No	26	20	2.84	1.84	57.7	37.4
Yes	Yes	111	72	3.37	2.37	68.4	48.1
Total		353	375	—	—	41.8	—

\*'Yes' applies to men ever employed in an industry shown in this study to be associated with increased risk.

†'Yes' applies to men who have smoked at least 100 cigarettes.

obtained by taking a weighted average of these category-specific values, the weighting factor being the total number of cases in the respective categories. For the exposed, the summary estimate is obtained as a weighted average of the category-specific attributable risks percent, the weighting factor being the number of exposed cases in the respective categories.

The concept of excess relative risk, relative risk minus 1, is essential to the method presented. The measure has another application especially useful in the study of diseases for which multiple causal factors are suspected. For such diseases it is often asked whether the multiple factors act independently or together ('synergistically') in persons exposed to more than one such factor. Unlike relative risks, but like attributable risks, the excess relative risks are additive. They can thus be used to estimate the excess of disease to be expected among persons exposed to multiple factors under the assumption that each factor acts independently. The data in the Table illustrate this application. Whether the excess relative risks (column 6) or the attributable risks (column 8) are used, men experiencing both exposures have about the excess of disease that would be expected from simple addition of the independent risks. For excess relative risks this is 2.37

observed, compared to 2.84 (1.00 plus 1.84) expected, and for attributable risks this is 48.1 observed compared to 57.7 expected. Both approaches suggest that the two exposures do not act synergistically in the production of disease. (Data in the Table are from a case-control study of a total incidence series of cases from an enumerated population and of controls selected from the same population. This allows computation of the absolute annual rates shown in column 7.)

#### SUMMARY

A method is presented whereby, using data from a case-control study, the attributable risk percent can be computed for persons exposed to a suspect aetiological factor or for the population as a whole.

This work was supported by a research grant (5 PO1 CA 06373) from the National Cancer Institute, National Institutes of Health, U.S. Public Health Service.

#### REFERENCES

- COLE, P., MONSON, R. R., HANING, H., and FRIEDEL, G. H. (1971). Smoking and cancer of the lower urinary tract. *New Engl. J. Med.*, **284**, 129.
- MACMAHON, B., and PUGH, T. F. (1970). *Epidemiology Principles and Methods*. Little, Brown, Boston.