International Journal of Epidemiology

Vol 7, No 2
Printed in Great Britain

Calculation of Attributable Risks from Epidemiological Data

STEPHEN D WALTER[1]

Walter S D (Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, CT 06510, USA). Calculation of attributable risks from epidemiological data. *International Journal of Epidemiology* 1978, 7: 175–182.

Attributable risk is an important epidemiological index, and may be more relevant to many health planning situations than other indices such as the relative risk. This paper outlines the requirements of typical epidemiological data in order to estimate attributable risks, and illustrates the calculations for some common types of research study. The examples chosen have no special significance *per se*, but represent the main findings of several typical investigations reported in the epidemiological literature. The combination of estimates from several studies and extrapolation to other populations are discussed. Finally the concept of attributable risk as the causal effect of a risk factor is presented.

INTRODUCTION

The use of relative risk as a measure of association between the incidence of a disease and exposure to a postulated risk factor is well established in epidemiology. Briefly defined, it is the ratio of the incidences of disease in persons exposed and not exposed to risk. It is of course quite possible to adopt other measures of association such as the difference in incidence rates between the two groups, but perhaps the reason that relative risk has enjoyed such widespread use is that it can be estimated directly from longitudinal and cross-sectional data, and also approximately (by using the odds ratio) in retrospective case-control studies. The latter is done under the assumption that the absolute value of the disease incidence rate is small. These three kinds of research design (cross-sectional, prospective and retrospective) are used in the vast majority of epidemiologic work; the simple large sample formulae for the estimate of the odds ratio and its associated standard error are the same in all three, an appealing characteristic.

Relative risk may be thought of as indicating the strength of the physiologic or biologic effects of exposure to the hazard under study. On the other hand, this measure does not take into account the number of individuals exposed to risk in the population, and so it would be quite possible to have a risk factor with a high relative risk, but which is not an important health problem because very few individuals were exposed to it. Attributable risk is an alternative measure which takes into account not only the strength of the physiologic effect of exposure, but also the number exposed to the risk factor in question; it is loosely defined as the fraction (or percentage) of all cases of disease which are associated with the risk factor. Under certain assumptions (discussed below), the attributable risk is also the proportional reduction in the disease load which would occur if exposure were prevented; it is therefore a very useful measurement of the public health importance of a risk factor, and may be used to assess the relative contribution of each of several risk factors in the generation of disease(s). Such calculations may be used as a rational foundation for the choice between alternative preventive strategies.

In symbols, if we let I_e be the incidence of disease in persons exposed to a risk factor, and I_o the incidence in persons not exposed, then $I_e - I_o$ is the excess risk of disease among exposed persons which may be ascribed to the exposure *per se*. Hence the fraction of cases among the exposed which are due to the exposure is $(I_e - I_o)/I_e$, or

1 Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, CT 06510, USA.
Selected for presentation at 8th International Scientific Meeting of the International Epidemiological Association, Las Croabas, Puerto Rico, September 1977.

$(r-1)/r$ where r is the relative risk I_e/I_o. If the proportion of the population exposed to risk is denoted by p, then it may be shown (1, 2, 3) that the proportion of all cases (in both exposed and non-exposed groups) which are associated with exposure is

$$AR = \frac{p(r-1)}{p(r-1)+1} \quad (1)$$

This quantity is known as the attributable risk, but it has also been termed the population attributable risk (4), the aetiologic fraction (5) and attributable fraction (3); the terms attributable risk percent and population attributable risk percent have also been suggested for use in the obvious ways, but the definitions without the word 'percent' are commonly applied to the percentage equivalents of these parameters. The associated mortality parameter (6, 7) is also related to attributable risk.

From equation (1) we may see that in order to estimate the attributable risk, in general we must be able to derive from the data estimates of the relative risk (or its odds ratio approximation) and the fraction of the population exposed to risk. The former is available from most epidemiologic studies as just discussed, but the latter is usually only directly available from cross-sectional studies, or from prospective studies when the sampling of study individuals is not stratified by exposure status. In retrospective work, an approximate estimate of the fraction exposed may sometimes be available from the control group, as shown below in the numerical examples. In contrast to the odds ratio, the formula for the standard error of the attributable risk is different in the various types of study design. The examples below consider the estimation procedures in typical prospective and retrospective situations, and the combination of information from several studies of the same risk factor-disease association is also discussed.

EXAMPLES

Throughout the numerical examples following, the sample frequencies of cases and controls by exposure status will be denoted as in the general contingency table (Table 1). The simplest situation, with one dichotomous (present or absent) risk factor and a dichotomous disease will be considered, although more general formulations of the problem are possible.

Example 1. Prospective study

The illustrative data for the calculation of attributable risk from prospectively collected information comes from the Framingham study of heart disease, using a small subset of data reported by Cornfield (8). Table 2 shows the number of cases of coronary heart disease (CHD) which had developed after six years of follow-up of 1329 men in the age range 40–59 years, the subjects being grouped by their initial level of serum cholesterol which will be used as the risk factor. For simplicity, this variable is divided into only two ranges, and other factors which may influence the probability of contracting CHD have been ignored.

In this study the method for sampling the 1329 men observed did not involve any stratification on the risk factor (initial level of serum cholesterol); rather, the individuals were obtained from an unstratified sampling scheme and subsequently grouped into the two risk groups shown. For this reason we may take the observed fraction of men exposed to risk (756/1329 = 0.569) as a valid estimate of the fraction exposed in the general population. There is, however, sampling variation associated with this estimate which must be taken into account when estimating the attributable risk. The relative risk may be estimated in the usual way as $72/756 \div 20/573 = 2.73$; the corresponding odds ratio estimate is $ad/bc = 2.91$, and a 95% confidence interval is (1.75, 4.84). By

TABLE 1

General 2 x 2 sample contingency table arising in epidemiological studies

		Disease		
		Present	Absent	Total
Risk factor	Present	a	b	$m_1 = a + b$
	Absent	c	d	$m_2 = c + d$
	Total	$n_1 = a + c$	$n_2 = b + d$	$N = a + b + c + d$

TABLE 2

Results from a prospective study of coronary heart disease

		CHD after 6 years	No CHD after 6 years	Total
Initial serum cholesterol(mg%)	220+	72	684	756
	<220	20	553	573
	Total	92	1237	1329

Measure	Estimate	95% confidence limits
Odds ratio:	2.91	1.75, 4.84
Attributable risk	49.6%	30.5%, 68.7%

using these estimates of the fraction exposed and the relative risk in equation (1) one may derive the estimate of attributable risk as 0.496 or 49.6%. The appropriate formula ((3) − equation A7) to estimate the corresponding variance is

$$V(AR) = \frac{cN[ad(N - c) + bc^2]}{(a + c)^3 (c + d)^3} \qquad (2)$$

which in this example gives V(AR) = 0.00951, and a standard error of $[V(AR)]^{1/2} = 0.0975$, yielding in turn an approximate 95% confidence interval 0.496 ± 1.96 x 0.0975, or (30.5%, 68.7%), making the assumption that the estimate is approximately normally distributed. Thus we may state that the data indicate that with 95% probability, and ignoring the effects of other risk factors, the percentage of CHD cases associated with, or attributable to serum cholesterol levels above 220 mg% is between 30% and 69% for men of this age group. Note that the use of 220 mg% as the dividing point between normal and abnormal values may be questioned. If a higher level were chosen, however, (say 250 or 260 mg%) the change in the attributable risk estimate may not be very great; the reason for this will be discussed below in the context of combining data from several studies where quite different definitions of exposure to risk may have been used.

Example 2. Retrospective study

Table 3 shows the outcome from a typical retrospective case-control study of the association of smoking and oral cancer; these data are originally from a study by Wynder *et al.* (9) and have also been discussed by Rothman and Keller (10). In the usual way, the odds ratio will be adopted as an approximation to the relative risk; the estimated odds ratio from Table 3 is 2.57, and its 95% confidence interval is (1.61, 4.11). The additional parameter required to calculate the attributable risk is the fraction of the population exposed to risk, which in this study is defined as smoking 16 or more cigarettes per day; in most retrospective studies the prevalence of exposure in the population

TABLE 3

Results from a retrospective study of oral cancer

		Oral cancer cases	Controls	Total
Smoking status (cigarettes/day)	16+	255	93	348
	<16	49	46	95
	Total	304	139	443

Measure	Estimate	95% confidence limits
Odds Ratio:	2.57	1.61, 4.12
Attributable risk	51.3%	34.3%, 68.3%

is not directly available, and so it must be estimated from some ancillary data or alternatively from the control group, the latter being adopted if it can be reasonably assumed that the class of persons from which the controls were selected have approximately the same pattern of exposure as does the entire population. This will often be the case if the absolute incidence of the disease is low (so that only a small fraction of the population contracts the disease), and if there is no control selection bias related to exposure. A specific case is worthy of mention; if the cases and controls are matched in some way (either individually or in groups), then it is very likely that the matching process will yield a different exposure distribution in the controls as compared to what would have arisen in a comparable unmatched control series. Therefore in a matched study, data in addition to that of the outcome contingency Table 1 (or its matched format equivalent) would be required in order to estimate the level of exposure in the general population.

For simplicity, in the example of Table 3 we will suppose that the controls are in fact fairly representative of the general population, at least with respect to smoking as categorised there. Thus we can estimate that the fraction of individuals who smoke more than 15 cigarettes per day is 93/139 = 0.669 or 66.9%. Now combining this with the previous estimate 2.57 of the odds ratio, equation (1) yields the value 51.3% as the estimated attributable risk. Using the variance formula appropriate for a retrospective study ((3) — equation A2) we have

$$V(AR) = \left(\frac{cn_2}{dn_1}\right)^2 \left[\frac{a}{cn_1} + \frac{b}{dn_2}\right] \qquad (3)$$

which in this example may be enumerated as 0.00751 to give the standard error as $[V(AR)]^{1/2} = 0.0867$; as in example 1, if we assume this estimate to have approximately a normal distribution, the 95% confidence interval is (34.3%, 68.3%). [Note: The alternative assumption that log (1 − AR) is normally distributed (11) yields a very similar interval.] Notice that although the width of this confidence interval is roughly comparable to that in example 1, the prospective study has three times as many study subjects as the retrospective. It is usually the case that retrospective studies are much more efficient in this regard than are prospective studies of comparable size (12).

Example 3. Combination of data from several studies

The illustrative data for this topic come from ten retrospective studies of the association of smoking and lung cancer (see Table 4). These ten are in fact a subset of 14 studies assembled by Dorn (13), and constitute a group which have been shown to be statistically homogeneous in their odds ratios (14). These data have also been examined by Armitage and Gart (15, 16).

Procedures to combine studies to give an estimate of their overall odds ratio are well-established, the best known being the Mantel-Haenszel method (17). If it can be assumed that all the variation between studies in the estimates of odds ratio is due to sampling error (i.e. the same underlying true relative risk applies to all the studies individually), then by using an appropriate weighting system for each of the individual odds ratio estimates, it is possible to create a weighted average for the estimate of the overall odds ratio and to use the weights to calculate a confidence interval (15). An obvious choice for the weight is the reciprocal of the variance of the odds ratio estimate, so that studies which yield very precise values for the odds ratio will receive more weight and vice versa. The previously demonstrated homogeneity of odds ratios in the ten studies make this approach reasonable here; column 5 of Table 4 indicates the odds ratio estimate arising from each study (being the usual estimate ad/bc). Column 6 shows the associated weight w_1 for the logarithm of each of these estimates, being the reciprocal of the estimated variance, i.e. $[1/a + 1/b + 1/c + 1/d]^{-1}$, and scaled to indicate the percentage contribution of each study relative to the total weight; thus study 1 provides an individual estimate 5.38 for the odds ratio, the logarithm of which receives 2.2% of the total weight in the combined estimate of the log odds ratio. The combined estimate for the overall odds ratio is 4.62, and the 95% confidence interval is (3.82, 5.60).

Column 7 of Table 4 indicates the estimated percentage of the population who smoke (or are "exposed") for each of the ten studies; the estimates all come from the control groups, e.g. the estimate for study 1 being 72/86 = 84%. Column 8 shows the estimated attributable risk, derived from equation (1), and column 9 shows the associated weights w_2 which are correspondingly taken as the reciprocals of the variances of these estimates (the variance being calculated from equation 3), and again scaled to indicate the contribution of each study relative to the total weight. A combined

TABLE 4

Combination of relative and attributable risks from 10 retrospective studies of smoking and lung cancer

Study	Lung Cancer Cases (1) Smokers	Lung Cancer Cases (2) Non Smokers	Controls (3) Smokers	Controls (4) Non Smokers	(5) $\hat{OR}$	(6) w_1	(7) θ(%)	(8) $\hat{AR}$(%)	(9) w_2
1	83	3	72	14	5.38	2.2	84	78.6	2.6
2	90	3	227	43	5.68	2.5	84	79.7	3.3
3	129	7	81	19	4.32	4.4	81	76.9	3.5
5	412	32	299	131	5.64	21.2	70	76.3	24.1
8	1350	7	1296	61	9.08	5.9	95	88.5	22.3
9	60	3	106	27	5.09	2.4	80	76.5	2.4
10	459	18	534	81	3.87	13.2	87	71.9	8.8
12	499	19	462	56	3.18	12.7	89	66.1	6.0
13	451	39	1729	636	4.25	31.6	73	70.4	21.4
14	260	5	259	28	5.62	3.9	90	80.7	5.4
Total	3793	136	5065	1096		100.0			100.0

OR = Odds ratio; w_1 = Weight for log OR; θ = Percentage of smokers among controls;
AR = Attributable risk; w_2 = Weight for AR.
Weighted mean for OR = 4.62; Weighted mean for AR = 77.2%.

estimate of the attributable risk may be derived by taking a weighted average of the individual attributable risk estimates, and this yields the value 77.2%. Also, using the reciprocal of the total weight (before scaling) as the variance of the combined estimate we may obtain a 95% confidence interval for the overall attributable risk as (73.0%, 81.4%).

It should be noted that estimating the attributable risk from the 'pooled' 2 x 2 table (i.e. the table with the total number 3793 of lung cancer cases who smoked, the total number 136 of cases who did not smoke, etc.) will be unsatisfactory in general, as is true for odds ratio estimation. The odds ratio from the pooled table is 6.03, a value which exceeds nine of the ten estimates for the individual studies. The overall percentage exposed to risk, which may also be obtained from the pooled table, is 82%, and in combination with the value 6.03 for the relative risk, one derives 80.5% as the attributable risk estimate from the pooled table. This is larger than eight of the ten corresponding individual study estimates of attributable risk, but seems somewhat less extremely placed in the distribution of individual attributable risk estimates than occurs in the relative risk; nevertheless this procedure is not recommended as reliable in general.

As an aside, it may be pointed out that in three of the 14 original studies of Dorn, individual matching of cases and controls was employed (13). These were studies 7, 8 and 12 (using Cornfield's numbering (14)), and all three used hospital patients as the source of controls. In study 7 (not included in Table 4) two controls per case were matched on age and hospital; in study 8 pairwise matching of controls on age and hospital was used; in study 12 controls were individually matched on age and race. The data from all 14 studies collectively constitute a familiar example used in methods to analyse several studies simultaneously, particularly in the context of obtaining an overall estimate of the odds ratio. As far as this writer is aware, none of these analyses takes account of the matching but this is not surprising since the studies themselves all report the data as if they had been unmatched; strictly speaking the relative risk confidence intervals derived without regard to the matching will probably be too liberal, i.e. the analysis could have been strengthened if it had been possible to analyse the matched studies allowing for the matching. For attributable risk, ignoring matched variables may have serious consequences because the matching process will usually influence the distribution of the matched variable in the controls, and so will distort the estimate of the factor prevalence in the general population if such an estimate is made from the control group. Even where group matching has been used the same considerations will apply, but again because the data are not available in matched format the matching in all of the studies will have to be ignored. It is, however, salutory that study 8, one of the two pair-matched studies in Table 4, has the highest estimate of the proportion smoking among the controls.

The ten studies vary greatly in their estimates of

the percentage of the population exposed to risk, despite their homogeneity with respect to relative risk. Apart from the effects of matching, such variation may be due either to true differences in the factor prevalence and the type of smoking encountered in the various study populations, or to differences in the operational definitions of smoking as applied by the various researchers involved. Fortunately the latter discrepancies tend to be annulled by a useful 'cancelling' feature of attributable risk which arises because the relative risk (minus one) is multiplied by the proportion exposed (see equation 1); thus if one study adopted a very stringent definition of exposure (e.g. 40+ cigarettes per day), this would lead to a high relative risk but quite a small proportion of individuals exposed. On the other hand, a more lenient definition of exposure (e.g. ever smoked at all) would produce a lower relative risk but a much higher proportion of individuals exposed to risk. These two opposite effects tend to cancel in the estimation of attributable risks for the various studies, and so a set of studies with quite different definitions of exposure and resulting relative risks may still produce compatible attributable risks.

Extrapolation to other populations

Despite the above comforting feature of attributable risk, the variation in the percentage exposed may still be troublesome as it may indeed represent (at least partially) real differences in smoking patterns between the various populations; if this is so, then even if the relative risks are constant one would still expect the attributable risks to be different, and the interpretation of the combined estimate of attributable risk from such a group of studies would be difficult. In order to use such a combined estimate to assess the attributable risk of a factor in a particular population, ideally one would require that the factor prevalence there and in the populations where the research studies were carried out are fairly similar. It might be, however, that all the available literature on the association in question related to populations for which the exposure levels were all quite different from that in the population for which an attributable risk estimate is required. With the exception of study 8 which was performed in England, all the studies of Table 4 which have moderate or heavy weight in the calculation of relative and attributable risks were done in the United States. Studies 1, 2 and 3 were carried out in Germany and the Netherlands and all receive rather low weight in the calculations because of their smaller size. All ten studies were done in Western industrialised countries, and so an assessment of the contribution of smoking in relation to other factors to the development of lung cancer for countries or societies where the prevalence of smoking was not comparable would require a strategy in the attributable risk estimation different from that of simply taking a weighted average of the individual attributable risks; this is quite apart from allowances which may be needed for other differences between the countries involved.

One possible approach to this problem is to think of the relative risk as a kind of biological constant which would apply in a wide variety of circumstances; if this constant relative risk can be estimated it may then be used in combination with a previously determined percentage exposed to risk appropriate for the population of interest; this then yields an attributable risk for that population. Thus in the case of smoking and lung cancer, we might suppose that after allowing for any other relevant factors a smoker had approximately a four to fivefold chance of contracting lung cancer as compared to a non-smoker, this being so regardless of the population in which the smoker is found. If one knew with some precision that, say, 50% of the population under study are exposed to risk (defined in a manner consistent with that which defines the relative risk), then this prevalence may be used together with the assumed value of relative risk to yield an attributable risk for that population.

This is done in Table 5 for the ten studies of smoking and lung cancer, using 4.62 as the estimate

TABLE 5

Attributable risk for smoking in lung cancer; relative risk from Table IV; various assumed values for the prevalence of smoking.

Percent of population who smoke	Attributable risk (%)
5	15.3
10	26.6
20	42.0
30	52.1
40	59.2
50	64.4
60	68.5
70	71.1
80	74.4
90	76.5
95	77.5

Relative risk taken as 4.62.

of relative risk and with a variety of assumed values for the percentage of the population exposed. As it turns out, the smoking prevalence ranging from 60% to 95% changes the attributable risk by less than 10%, and the decline of the attributable risk with decreasing factor prevalence is quite slow. This is reassuring because it indicates that when the factor prevalence is not known with great precision, a crude estimate may sometimes be used with no serious effect on the attributable risk. The effect of varying the factor prevalence is, however, more marked in situations where the relative risk is closer to the value one. Lilienfeld gives a table of attributable risk for various values of the factor prevalence and the relative risk (18).

It is interesting at this point to compare how the weights of the ten studies differ between the combined estimates of the relative and attributable risks. Studies 5, 8, 10, 12 and 13 provide the five highest weights for both parameters, with the other studies making rather small contributions in both cases (Table 4). The ordering within these first five is different, however, as is the distribution of weight. Study 13 has the highest weight for the relative risk calculation, followed by study 5 and study 10 with considerably less weight in turn. On the other hand, the three highest weights for attributable risk, those of studies 5, 8 and 13, are approximately equal. The most striking individual difference is for study 8 which is weighted in excess of three times more heavily in the attributable risk than in the relative risk. In general terms this is because the variance of the odds ratio is high for a study of its size due to the small number of non-smoking lung cancer cases; study 8 does, however, possess a large control group which leads to a quite precise estimate of the percentages exposed to risk. The effect of a poorly dichotomised case group is of somewhat lesser importance in the estimation of attributable risk than in the odds ratio; in the latter the balance between smokers and non-smokers is equally important in both groups in order to reduce the standard error, whereas for attributable risk the control group is more important because it contributes to both components of this measure, i.e. the proportion of individuals exposed and the relative risk.

Although the set of studies discussed in this section were all retrospective, it would of course be quite possible to use the same strategies to evaluate a combined attributable risk estimate from a set of prospective studies, or indeed to a mixture of the two types of research design.

DISCUSSION

The numerical examples discussed in this paper have deliberately been somewhat simplified. The risk factors were assumed to have only two levels, and the disease was not allowed to vary in severity; only one risk factor for each disease was analysed, and the period between exposure and the onset of disease (the latent period) has been ignored. The possibility of competing risks of mortality/morbidity have been omitted. All of these simplifying manoeuvres are not essential, and all of the arguments presented can be generalised in principle. More important to the interpretation of the attributable risk measure is its relationship to the actual reduction in the disease rate which would occur if exposure to the risk factor were prevented. It can be shown that if the factor is indeed a causal agent of the disease and the risk in question is distributed in the population independently of all other causal risk factors, then the reduction and the attributable risk are equal. On the other hand if risk factors are confounded, then the attributable risk will be either greater or lesser than the actual reduction in disease, depending on whether the confounding is positive or negative.

If a risk factor is not in fact causal, some impact on the disease may still occur by secondary changes in risk exposure induced by action directed against the non-causal factor. For example, if a particular dietary constituent is postulated as a risk factor for heart disease, then diet modification might have an effect on the probability of this disease even though the postulated factor was not causal; this might occur if the consumption of other dietary components were altered by the primary modification, or even by such indirect mechanisms as an improved general awareness of health hazards in the patient after being instructed to modify his diet. It may be shown by an algebraic argument (not reproduced here) that the size of the impact which occurs by acting on non-causal or spurious factors depends primarily on the degree of confounding between the causal and non-causal factors, and also on the probability that the level of exposure to the causal factor is altered by the primary preventive measure; it turns out that the effect of modifying a non-causal agent which is heavily confounded with a causal factor is often well estimated by the apparent attributable risk of the non-causal factor, and may even exceed the attributable risk in certain situations. This is a useful finding because many philosophers take the position that causality is impossible to demonstrate by observational data of the kind often used by epidemiologists.

To conclude, it should be said that the calculation of attributable risks should be only one facet in a multidisciplinary approach to the choice between alternative preventive programmes. Issues such as the acceptability and cost of preventive programmes must also be taken into account, together with an evaluation of the expected duration of the programme benefits and the period of time required for the benefits to be realised. The calculations and numerical methodology expounded here represent but one small tool in the extensive kit of techniques which the epidemiologist working in this area will require.

REFERENCES

(1) Levin M L. The occurrence of lung cancer in man. *Acta Unio Inter Contra Cancrum* **19**: 531, 1953.

(2) Rouquette C and Schwartz D. Méthodes en épidémiologie. Éditions Médicales. Flammarion, 1970.

(3) Walter S D. The estimation and interpretation of attributable risk in health research. *Biometrics* **32**: 830, 1976.

(4) Cole P and MacMahon B. Attributable risk percent in case-control studies. *British Journal of Preventive and Social Medicine* **25**: 242, 1971

(5) Miettinen O S. Proportion of disease caused or prevented by a given exposure, trait or intervention. *American Journal of Epidemiology* **99**: 325, 1974.

(6) Seigel D G and Markush R E. Prevalence at death. II. Methodological considerations for use in mortality studies. *American Journal of Public Health* **58**: 772, 1968.

(7) Markush R E. Levin's attributable risk statistic for analytic studies and vital statistics. *American Journal of Epidemiology* **105**: 401, 1977.

(8) Cornfield J. Joint dependence of risk of coronary heart disease on serum cholesterol and systolic blood pressure; a discriminant function analysis. *Federation Proceedings* **21**: 58, 1962.

(9) Wynder E L, Bross I J and Feldman R M. A study of the etiological factors in cancer of the mouth. *Cancer* **10**: 1300, 1957.

(10) Rothman K and Keller A. The effect of joint exposure to alcohol and tobacco on risk of cancer of the mouth and pharynx. *Journal of Chronic Diseases* **25**: 711, 1972.

(11) Walter S D. The distribution of Levin's measure of attributable risk. *Biometrika* **62**: 371, 1975.

(12) Walter S D. Determination of significant relative risks and optimal sampling procedures in prospective and retrospective studies of various sizes. *American Journal of Epidemiology* **105**: 387, 1977.

(13) Dorn H F. The relationship of cancer of the lung and the use of tobacco. *American Statistician* **8**: 7, 1954.

(14) Cornfield J. A statistical problem arising from retrospective studies. *Proceedings of the 3rd Berkeley Symposium* **4**: 135, 1956.

(15) Armitage P. Statistical methods in medical research. Blackwell, Oxford and Edinburgh. 1971.

(16) Gart J J. On the combination of relative risks. *Biometrics* **18**: 601, 1962.

(17) Mantel N and Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* **22**: 719, 1959.

(18) Lilienfeld A M. Foundation of epidemiology. Oxford University Press, New York. 1976.

(received 12 October 1977)

Book News

The following book has been received and may be of interest to readers.

The Heart Patient Recovers
by *Sydney H Croog and Sol Levine.*
An authoritative account of social and psychological factors involved in recovery after a first heart attack. pp 432. Human Sciences Press, New York, 1977.