AMERICAN

Journal of Epidemiology

Formerly AMERICAN JOURNAL OF HYGIENE

VOL. 96 DECEMBER, 1972 NO. 6

Testa

ORIGINAL CONTRIBUTIONS

STANDARDIZATION OF RISK RATIOS

OLLI S. MIETTINEN[1]

(Received for publication July 6, 1972)

Miettinen, O. S. (Harvard School of Public Health, Boston, Mass. 02115). Standardization of risk ratios. *Am J Epidemiol* 96: 383–388, 1972. The procedure commonly employed for the computation of "standardized" risk ratio estimates (observed-to-"expected" ratios) which characterize different categories of a risk factor do not lead to a set of mutually comparable values. In cohort studies truly standardized risk ratios can be obtained through a simple modification of the prevailing method. The problem is more subtle in case-control studies, but these studies, too, permit the computation of standardized risk ratio estimates with explicit specification of the standard.

biometry; epidemiologic methods

The risks of disease or death for different categories of a risk factor (*RF*) are commonly considered in terms of the ratio of the category-specific risks to that in a selected reference category of the *RF*, and often an attempt is made to standardize these risk ratios ("relative risks" (*RR*'s)) with respect to the distribution of some confounding factor (*CF*). Involved in the computation of such mutually comparable risk ratios for a set of categories of a *RF* are (a) the specification of the reference category of the *RF* (for which *RR* = 1 by definition), (b) the specification of the standard distribution for the *CF* and (c) the calculation of the *standardized risk ratios* (*SRR*'s) according to the two specifications.

Although the use of *SRR*'s is central to epidemiologic research, the methods of computation have remained unsatisfactory. In *cohort* (follow-up) studies it is customary to take the standard to be the *CF* distribution of the group in the reference category and to compute for each of the compared groups a "standardized morbidity (mortality) ratio" (*SMR*) in the spirit of "indirect standardization." However, the actual standard (common distribution of people over the strata) in this procedure is not

Abbreviations: *CF*, confounding factor; *CRR*, crude risk ratio; *RF*, risk factor; *RR*, risk ratio; *SMR*, standardized morbidity (mortality) ratio; *SRR*, standardized risk ratio.

[1] Departments of Epidemiology and Biostatistics, Harvard School of Public Health; and Department of Cardiology, Children's Hospital Medical Center, Boston, Massachusetts.

Supported by grants 5 PO1 CA06373 and HE 10436 from the National Institutes of Health.

provided by the group in the reference category but rather by the group for which the *SMR* is computed; a set of *SMR*'s thus lacks a common standard and, therefore, mutual comparability. Even those who are aware of this problem apply the procedure (1, 2)—partially because this approach is not considered to be very misleading, and presumably also because an attractive alternative has not been available. In *case-control* studies, the practice is to compute some "summary" or "standardized" estimate for the different groups without even considering the selection of the standard in an explicit manner—to say nothing about an actual attainment of standardization.

The purpose of this article is to present a simple modification of the prevailing procedure in cohort studies leading to actual *SRR*'s, and to derive an equivalent procedure for case-control studies.

Cohort studies

A layout and notation system for cohort study data is presented in table 1.

The estimate for the crude risk ratio (*CRR*) for the i^{th} *RF* category is

$$\widehat{CRR}_i = (e_i/F_i)/(g/H) = e_i/(F_i g/H) = (He_i/F_i)/g. \quad (1)$$

It is seen that a sample *CRR* may be regarded, firstly, as the ratio of the observed number of events in the *RF* category at issue (e_i) to an estimate of the "expected" number which would have occurred in it had the crude rate of the reference category prevailed in it ($F_i g/H$). Alternatively, the *CRR* may be seen as the analogous "expected"-to-observed ratio in the reference category.

TABLE 1

Cohort study: layout and notation for the data

Stratum	Components of rate	Category of risk factor	
		i^{th}	Referent
j^{th}	Events* Denominator†	e_{ij} F_{ij}	g_j H_j
Total	Events* Denominator†	e_i F_i	g H

* Number of cases of disease or death.

† Number of individuals studied or person-years of follow-up.

As already noted, in the usual approach to estimating *SRR*'s, the sample *SMR* is computed for each group that is compared to the reference group. Specifically, for the i^{th} group,

$$\widehat{SMR}_i = e_i/(\textstyle\sum_j F_{ij}g_j/H_j), \quad (2)$$

the ratio of the observed number of events in the i^{th} group to an estimate of the corresponding "expected" number which the stratum-specific rates in the reference category would have produced. If in this formula we substitute $\sum_j F_{ij}e_{ij}/F_{ij}$ for e_i, it becomes apparent that SMR_i is the ratio of a ("directly") standardized rate in the category at issue to the correspondingly standardized rate in the reference category. However, the weights of the standardization (F_{ij}'s) derive from the group in the category being characterized rather than from the group in the reference category (which is generally presumed to provide the standard in this procedure). It follows that *SMR* estimates computed in this manner are internally standardized but not mutually comparable.

The *SRR* estimates which truly derive the *standard from the referent* involve as weights the denominators in the referent, H_j's, instead of the F_{ij}'s involved in the *SMR*'s. This modification of the usual procedure gives

$$\widehat{SRR}_i = (\textstyle\sum_j H_j e_{ij}/F_{ij})/(\sum_j H_j g_j/H_j),$$

or

$$\widehat{SRR}_i = (\textstyle\sum_j H_j e_{ij}/F_{ij})/g. \quad (3)$$

Thus the objectives of the *SMR* computations are met by using the "expected"-to-observed ratio in the reference group, with the estimate of the "expected" number based on the observed rates in the group at issue.

With a *general standard* distribution of the *CF* characterized by denominators S_j in the various strata, the i^{th} *SRR* is the corresponding ratio of ("directly") standardized rates, with the S_j's as the weights of the standardization:

$$\widehat{SRR}_i = (\sum_j S_j e_{ij}/F_{ij})/(\sum_j S_j g_j/H_j). \quad (4)$$

Case-control studies

A layout and notation system for case-control study data is presented in table 2.

To compute *SRR* estimates from case-control data, consider first the case of deriving the *standard from the referent.* Here the task is to develop an equivalent of formula 3 based on data from a case-control study. Just as formula 3, its equivalent for case-control studies is to express the ratio of the estimated "expected" number of cases in the reference group to the number of cases observed in this group. Moreover, the estimate for the "expected" number is to be derived on the assumption that the risk characteristics of the i^{th} category prevail in the reference category as well. As has been observed previously (3), the desired "expected" number may be estimated as $\sum_j a_{ij}d_j/c_{ij}$ whereas the observed number at issue is $\sum_j b_j = b$. Therefore

$$\widehat{SRR}_i = (\sum_j a_{ij}d_j/c_{ij})/b. \quad (5)$$

More generally, one might derive the *standard from any one of the compared RF categories, or their combination.* To obtain a formula for the SRR_i in this case, we note first that the general definition of *SRR* in formula 4 may be recast as

$$\widehat{SRR}_i = [\sum_j (S_j g_j/H_j)\widehat{RR}_{ij}]/(\sum_j S_j g_j/H_j). \quad (6)$$

This shows that in general the *SRR* is to be computed as a weighted average of the stratum-specific *RR*'s, and that the weights, $(S_j g_j/H_j)$'s, involve not only the standard distribution but also the stratum-specific risks in the reference category. For the computation of these weights we note first that if in the j^{th} stratum the sampling fraction of noncases is f_j and if from the standard category there are C control subjects in the study, then the weights S_j are proportional to the "expected" values of C_j/f_j. If the sampling fraction of cases is uniform over the strata, then the rates g_j/H_j can be taken to be proportional to $b_j/(d_j/f_j)$. Thus, upon substitutions,

$$\widehat{SRR}_i = [\sum_j (C_j b_j/d_j)\widehat{RR}_{ij}]/(\sum_j C_j b_j/d_j), \quad (7)$$

where $\widehat{RR}_{ij} = a_{ij}d_j/b_j c_{ij}$. It may be noted that formula 5 is a special case of this, obtained by setting C_j equal to the number of control subjects in the reference category within the j^{th} stratum, i.e., by setting $C_j = d_j$. But more generally $C_j = \sum_i c_{ij}$ for some range of i.

For a case of a *general standard*—characterized by stratum-specific denominators S_j—expressions for SRR_i are immediately apparent from the above. If no matching was employed in the selection of the control series, then

$$\widehat{SRR}_i = [\sum_j (S_j b_j/d_j)\widehat{RR}_{ij}]/(\sum_j S_j b_j/d_j). \quad (8)$$

If the control series is a matched one, it is necessary to adjust for the variability of the

TABLE 2

Case-control study: layout and notation for the data

Stratum	Series	Category of risk factor	
		i^{th}	Referent
j^{th}	Cases	a_{ij}	b_j
	Controls	c_{ij}	d_j
Total	Cases	a_i	b
	Controls	c_i	d

sampling fraction of noncases over the strata:

$$\widehat{SRR}_i = [\textstyle\sum_j (S_j b_j f_j / d_j) \widehat{RR}_{ij}] / (\textstyle\sum_j S_j b_j f_j / d_j). \quad (9)$$

(In the use of this formula it will suffice to have numbers proportional to the values of f_j.)

Examples

Example 1. Table 3 presents hypothetical cohort study data such that the stratum-specific rates are identical between *RF* categories 2 and 3. The *CRR*'s for these two categories are different because of confounding by the stratification factor. But even the respective *SMR*'s differ from each other, thus failing to reflect the intra-stratum identities between categories 2 and 3. By contrast, $SRR_2 = SRR_3$.

Example 2. Table 4, based on a cohort study, shows rates of pulmonary or bronchial cancer in relation to cigarette smoking, with age as a stratification factor. For the lighest smokers $\widehat{CRR} = (538/460)/(78/444) = 6.66$, $\widehat{SMR} = 538/[717(0/352) + \cdots + 630(8/849)] = 7.64$, and with the referent as the standard, $\widehat{SRR} = [352(2/717) + \cdots + 849(21/630)]/78 = 7.55$. With the standard derived from all current cigarette smokers, the standard denominators for the successive age categories become 71,700 + 40,600 + 3990 = 116,000, $\cdots$, and 6300 + 1270 + 232 = 7800. The corresponding $\widehat{SRR}$ for the lightest smokers is $[116(2/71.7) + \cdots + 7.80(21/6.30)]/[116(0/35.2) + \cdots + 7.80(4/8.49)] = 7.70$. Analogous calculations are applied to the other categories of smoking. In this example the differences between *SMR*'s and *SRR*'s are rather minor in absolute terms, but considerable in terms relative to the differences between *CRR*'s and *SRR*'s.

TABLE 3

Hypothetical data on risk in categories of a risk factor, by strata of a confounding factor

Stratum	Components of rate	Category of risk factor		
		3	2	1
1	Events	1600	400	250
	Denominator	4000	1000	2500
	Rate	0.40	0.40	0.10
2	Events	500	2000	750
	Denominator	1000	4000	2500
	Rate	0.50	0.50	0.30
Total	Events	2100	2400	1000
	Denominator	5000	5000	5000
	Rate	0.42	0.48	0.20
	*CRR**	2.10	2.40	1.00
	*SMR**	3.00	1.85	1.00
	*SRR**, †	2.25	2.25	1.00

* Referent = category 1.
† Standard = referent.

TABLE 4

Death rate (number of cases per person-years of follow-up) for cancer of the lung or bronchus in relation to current cigarette smoking. Kahn (1)

Age	Current No. of cigarettes per day			
	Occasional-20	21–39	40+	None*
35–44	2/71,700	4/40,600	0/3,990	0/35,200
45–54	2/20,800	10/12,800	2/1,930	0/15,100
55–64	220/212,000	245/103,000	63/19,600	25/214,000
65–74	293/149,000	194/50,000	50/8,940	49/171,000
75+	21/6,300	7/1,270	3/232	4/8,490
Total	538/460,000	460/208,000	118/34,800	78/444,000
$\widehat{CRR}$ †	6.66	12.6	19.3	1.00
$\widehat{SMR}$ †	7.64	17.1	23.8	1.00
$\widehat{SRR}$ †, ‡	7.55	15.8	22.7	1.00
$\widehat{SRR}$ †, §	7.70	16.5	23.2	1.00

* Those who have never been regular smokers.
† Referent = "none" category.
‡ Standard = referent.
§ Standard = current cigarette smokers (occasional to 40+/day).

Example 3. The data in table 5, from a case-control study, relate the risk of breast cancer to parity, with age at first delivery as a confounding factor. With parity 1 as the referent we compute for parity 4–9 $\widehat{CRR} = [100(233)/394]/77 = 0.77$ and, with the referent as the standard, $\widehat{SRR} = [10(24)/50 + \cdots + 22(80)/32]/77 = 1.14$. With parities 1–9 as the standard, the distribution for age at first delivery is related to the set $C_1 = 50 + 47 + 24 = 121$, $C_2 = 312 + 428 + 129 = 869$, $C_3 = 32 + 144 + 80 = 256$, and the rates in the reference category to $b_1/d_1 = 2/24$, $b_2/d_2 = 36/129$, $b_3/d_3 = 39/80$. Thus the weights $C_j b_j/d_j$ for averaging the *RR* estimates over the strata can be taken as $121(2)/24 = 10$, $869(36)/129 = 243$ and $256(39)/80 = 125$. Thus, for parities 4–9, $\widehat{SRR} = [10(2.40) + 242(0.78) + 125\ (1.41)]/(10 + 243 + 125) = 1.03$. Analogous calculations are applied to parities 2–3. In this example, standardization removes rather substantial confounding, and the choice of the standard distribution also makes a considerable difference in the pattern of *SRR*'s.

TABLE 5

Risk of breast cancer in relation to parity. Salber et al. (4)

Age at 1st delivery	Series	Parity 4–9	2–3	1
<20	Cases	10	6	2
	Controls	50	47	24
	$\widehat{RR}$	2.40	1.53	1.00
20–29	Cases	68	144	36
	Controls	312	428	129
	$\widehat{RR}$	0.78	1.21	1.00
30+	Cases	22	47	39
	Controls	32	144	80
	$\widehat{RR}$	1.41	0.67	1.00
Total	Cases	100	197	77
	Controls	394	619	233
	$\widehat{CRR}$*	0.77	0.96	1.00
	$\widehat{SRR}$*, †	1.14	0.94	1.00
	$\widehat{SRR}$*, ‡	1.03	1.03	1.00

* Referent = parity 1.
† Standard = referent.
‡ Standard = parities 1–9.

Discussion

The standardization of *RR*'s involves two rather different problems. Firstly, inasmuch as a *RR* involves *two* risks, the distributions of the respective groups might be standardized to attain *internal* standardization of a given *RR*. Secondly, a set of *RR*'s might be *mutually* standardized by using a common internal standard for all the *RR*'s. Both problems have been dealt with in this article, but the emphasis has been on the attainment of mutual comparability for a set of *RR*'s each of which relates its respective category of a single risk factor to a common reference category of that factor. Often the need for such comparable risk ratios arises in the context of evaluating dose-response relationships, but in the other extreme one may deal with categories on a nominal scale, such as geographic areas.

In discussing the standardization, the concern here has been with the procedures of *computing SRR*'s, without regard to the selection of either the referent or the standard. The choice of the referent is an inconsequential problem, simply a matter of arbitrarily selecting a scale factor, whereas in the selection of the standard one could be guided by the intended application of the results and/or considerations of their stability. These problems are left outside the scope of this presentation.

To test the hypothesis that all the *SRR*'s are identical against the alternative of a monotone trend, the Mantel extension (5) of the Mantel-Haenszel test can be applied in the (usual) large-sample case, with appropriate scoring of the different levels of exposure.

References

1. Kahn HA: The Dorn study of smoking and mortality among U.S. veterans: report on eight and one-half years of observation. *In* Epidemiological approaches to the study of cancer and other chronic diseases. (Edited by W Haenszel.) Nat Cancer Inst Monogr 19: 1–125, 1966
2. Haenszel W, Kurihara M: Studies of Japanese

migrants. I. Mortality from cancer and other diseases among Japanese in the United States. J Nat Cancer Inst 40: 43–68, 1968
3. Miettinen OS: Components of the crude risk ratio. Am J Epidemiol 96: 168–172, 1972
4. Salber EJ, Trichopoulos D, MacMahon B: Lactation and reproductive histories of breast cancer patients in Boston, 1965–66. J Nat Cancer Inst 43: 1013–1024, 1969
5. Mantel N: Chi square tests with one degree of freedom: extensions of the Mantel-Haenszel procedure. J Am Statist Assoc 58: 690–700, 1963