

STANDARDIZED MORTALITY RATIO AS AN APPROXIMATION TO RELATIVE RISK

by

M.J. Symons, Ph.D.  
Department of Biostatistics and  
Occupational Health Studies Group  
University of North Carolina, Chapel Hill, N.C. 27514

and

J.D. Taulbee, Ph.D.  
The Procter and Gamble Company  
Miami Valley Laboratories, Cincinnati, Ohio 45247

Institute of Statistics Mimeo Series No. 1294

JULY 1980

STANDARDIZED MORTALITY RATIO AS AN APPROXIMATION TO RELATIVE RISK

M.J. Symons\*, Ph.D.  
Department of Biostatistics and  
Occupational Health Studies Group  
University of North Carolina, Chapel Hill, N.C. 27514

and

J.D. Taulbee\*, Ph.D.  
The Procter and Gamble Company  
Miami Valley Laboratories, Cincinnati, Ohio 45247

ABSTRACT

The standardized mortality ratio is a widely used and often criticized summary statistic for studies of occupational mortality. In this paper we discuss the conditions under which the standardized mortality ratio is a reasonable approximation to relative risk. When the true relative risk is greater than 100%, the standardized mortality ratio over-estimates relative risk no matter how small the mortality rates or how short the age bands utilized in the analysis. However when the excessive mortality is consistent across the age bands, the standardized mortality ratio can usefully approximate relative risk for some applications, such as those involving site-specific cancers, provided the age bands employed are not too large.

---

\*This research was completed while MJS was on leave with an Inter-governmental Personnel Act appointment with the Health Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle, N.C. 27711, and while JDT was on the faculty of the Department of Biostatistics and Occupational Health Studies Group, University of North Carolina, Chapel Hill, N.C. 27514.

# STANDARDIZED MORTALITY RATIO AS AN APPROXIMATION TO RELATIVE RISK

M.J. Symons and J.D. Taulbee

## INTRODUCTION AND SUMMARY

The standardized mortality ratio (SMR) has been generally criticized and its value as a useful summary statistic for studies of occupational mortality has been seriously questioned in two JOM articles.<sup>1,2</sup> The utility of the SMR as an approximation to relative risk (RR) was one of several characteristics considered. In this paper we show that the SMR can be a useful approximation to RR under certain conditions. Specifically we suggest that the SMR reasonably approximates relative risk when

- (i) the mortality rates in the Comparison population for the cause(s) of interest are no larger than about 100 per 10,000 workers per year with ten year age bands. [With twenty (five) year age bands these rates should be no larger than about 50 (200) per 10,000 workers per year.]

and

- (ii) the age-specific mortality rates for the Study and Comparison population are approximately in constant ratio across all age bands.

Several remarks about these conditions and the consequences of their assumption are noteworthy. First, since the SMR is a ratio of

weighted averages, its value will lie between the smallest and largest of the ratios of the Study to Comparison population age-specific mortality rates. Therefore, when these age-specific mortality ratios are all equal, the SMR will be this common value. Second, whenever the true relative risk is greater than unity (100%) the approximation by the SMR is anti-conservative however small the death rates may be. Third, the magnitude of the over-estimation increases with the value of the true relative risk and the length of the age bands. In particular with 10 year age bands and age-specific mortality rates no larger than 100 per 10,000 workers per year, the SMR over-estimates the RR by no more than 6%, 13%, or 20% when the true relative risk is 200%, 300%, or 400%, respectively. Finally, the amount of over-estimation can be controlled by utilizing shorter length age bands, especially when the true relative risk is expected to be quite large or the age-specific rates in the Comparison population become fairly large at the upper age range of the analysis.

#### DEFINITIONS AND DERIVATION

Suppose that the workforce of interest is composed of one sex-race group so that the only factor to be adjusted is age. Table 1 contains the notation for the available data in one age group for the Study (exposed) and Comparison (unexposed) populations. We assume that the data for the Study population are from a cohort of workers and those for the Comparison population are from (cross-sectional) Vital Statistics and Census surveys for the same sex-race group in the U.S. General population, as is often the case<sup>3, 4, 5, 6</sup>.

We note in passing that if the Comparison population is also available as a cohort of individuals, the Mantel-Haenszel<sup>7</sup> weighted

average of the age-specific odds ratios is an appropriate, preferred alternative to the SMR. (Particularly when there is a consistency in the odds ratio estimates across all the age bands, the Mantel-Haenszel statistic has very desirable power properties.<sup>8)</sup> Cornfield<sup>9</sup> showed that the odds ratio is an approximation to relative risk when the probability of disease is small, so that a similar relation between the odds ratio and relative risk exists as will now be shown for the SMR and relative risk.

The ratio of the number of deaths in the exposed population to the number expected without the exposure(s) is the mortality ratio (MR) in the specified age group. Specifically from Table 1, we have

$$MR = d/R_n = r/R. \quad (1)$$

The relative risk for the same age group is the ratio of the probability of an individual in the exposed population, alive at the beginning of the age interval, dying during the age band to the same probability for an individual in the unexposed population. From Table 1 we have

$$RR = q/Q. \quad (2)$$

Notice that relative risk is a concept involving the life chances of two individuals while the mortality ratio in (1) compares the death rates for the two populations. These are two related but different ideas.

The relationship between the mortality ratio in (1) and relative risk in (2) comes from life table theory where it can be shown<sup>10</sup> that under certain reasonable conditions (see Appendix for further comment )

$$1 - q = \exp(-rt), \quad (3)$$

for the exposed population, where  $t$  is the length of one age band and

$\exp(\cdot)$  is the exponential function. When the age-specific death rates are small and time period short, say five or ten years, the Taylor series expansion for (3) can be approximated by the linear portion,  $1 - rt$ . We have then the (approximate) relationship

$$q \doteq rt . \quad (4)$$

As an expression similar to (3) holds for the unexposed population, we have that  $Q \doteq Rt$  and then

$$RR = q/Q \doteq rt/Rt = r/R = MR. \quad (5)$$

Therefore for any single age group, the relative risk is approximately equal to the mortality ratio when the death rates in the two populations are small and the age bands are not too large.

The standardized mortality ratio is given by

$$SMR = \Sigma d / \Sigma Rn = \Sigma rn / \Sigma Rn, \quad (6)$$

the summations being over the age groups. If the mortality rates for the exposed and unexposed populations are in constant ratio, specifically

$$r/R = \theta \text{ or } r = \theta R, \quad (7)$$

for all age groups, then the

$$SMR = \Sigma (\theta R)n / \Sigma Rn = \theta . \quad (8)$$

Using (4) and (7) for the exposed population, we have

$$q \doteq \theta Rt . \quad (9)$$

From (5)  $Q \doteq Rt$  in the unexposed population and then from (2) the relative risk is approximately equal to the same constant  $\theta$ .

To summarize, the standardized mortality ratio will be numerically close to the relative risk when the age-specific deaths rates of the Study and Comparison populations and the age bands are small enough for the linear approximations used in (5) to hold and these rates are in constant ratio across all the age bands.

### NUMERICAL ILLUSTRATIONS

The linear approximation of the exponential relationship (3) can be seen from equation (9) to depend upon three factors. These are as follows:

- (i) the magnitude of the mortality rate,  $R$ , in the Comparison population,
- (ii) the size of the true relative risk,  $\theta$ , and
- (iii) the length of age bands,  $t$ , utilized in the analysis.

The range of occupational mortality applications where the SMR reasonably approximates relative risk is quickly grasped with a couple of numerical illustrations.

First consider the effect of the magnitude of the mortality rates in the Comparison population. This is determined largely by the cause(s) of death and the age range under study. In Table 2 the mortality rates for all malignant neoplasms are given by 10 year age bands for the 1971 U.S. white male population<sup>11</sup>, aged 35-74 years, along with the experience of a fictitious workforce, devised via (3) so that the age-specific relative risks are 200%. This means the probability of death from any malignancy in each 10 year age band, for an individual in the fictitious workforce who is alive at the beginning of the age bands, is twice that of a U.S. white male of the same age in 1971. Depending upon the age distribution of the person-years of experience in the fictitious workforce, we see from (6) that the SMR for Table 2 will be between 200.4% and 211.3%. Therefore the SMR would over-estimate the RR by no more than 5.65%.

With smaller mortality rates the approximation is much better. In Table 3 the mortality rates for lung cancers are presented in parallel fashion to Table 2, except the relative risk is presumed to

be 300%. The resultant mortality ratios range from 300.7% to 311.5%. An over-estimation of RR by the SMR in this situation would be no larger than 3.85%.

The effect of the true relative risk and the length of the age bands is considered in Table 4. The mortality rate of 100 per 10,000 workers per year was chosen for illustration and as a practical upper limit for many occupational mortality applications of the SMR since it is approximately the death rate for all malignant neoplasms for U.S. white males, aged 65-74 years in 1971. The over-estimation of RR by the SMR, as a percentage excess, is given in the right most column of Table 4 for each of three age bands: 5-years, 10-years, and 20-years. Within each age band the true relative risk is presumed to be 200%, 300%, and 400%. The SMR over-estimates the RR by varying amounts that can be seen to steadily increase with the true relative risk and the length of the age bands.

These numerical illustrations can be summarized by noting several things. First, for large age-specific death rates the SMR approximation to relative risk may be a gross over-estimate. Gaffey<sup>1</sup> provides an example in his Table 3 for all causes of mortality. However for some causes of death of interest in occupational mortality the SMR can provide a very satisfactory approximation to RR, as with the lung cancer example in our Table 3. Second, when the true relative risk is very large, greater than 400%, say, the SMR will be even larger, as can be seen from Table 4. Finally, with the use of large age bands, the resulting SMR will over-estimate the true relative risk. However as can also be seen in Table 4, by using short enough age bands all of the above-mentioned over-estimations can be controlled to a satisfactory degree



for many practical occupational mortality applications, such as site-specific cancers.

#### DISCUSSION AND CONCLUSION

The standardized mortality ratio has long been known<sup>12, 13, 14, 15</sup> to have inherent weaknesses as a summary statistic for occupational mortality studies. Since it depends upon the Comparison population rates and the age-distribution of the Study population person-years of experience, comparison of SMR's from different studies can be misleading. Various unrealistic examples are easily constructed<sup>1</sup> to illustrate this point, but differences also occur in practice. See Wong for further discussion and specific references.<sup>2</sup> The weighted average form, displayed in equation (6), is the reason that these difficulties are encountered. Chiazzo's comment<sup>15</sup> that there is no single substitute for a schedule of age-specific rates is very relevant, as such a schedule provides the only way to check for the validity of the fundamental assumption needed for the SMR to approximate RR.

The consistency across all age bands of the relative risk for the exposed individuals is the foundation of this approximation. Consequently, the choice of the age intervals employed in the analysis becomes an important study design consideration for each application. For example, with many site-specific malignancies, a sufficient latent period may need to be allowed for in order to detect an elevated relative risk. Enterline covers a number of these points and provides an excellent discussion of this basic study design issue.<sup>16</sup> Inclusion of age bands where the relative risk is not elevated will only decrease the sensitivity of the analysis.

The consideration of the age range and the length of the age bands for the analysis is important at the design stage of an occupational

mortality study. The appropriateness of the SMR approximation to RR is determined by examining the age-specific mortality ratios. This is an often repeated recommendation in the occupational mortality study methods literature.<sup>15,17</sup> Its relevance for checking a basic condition required for the reasonable approximation of relative risk by the SMR should be added to the list of all the previous reasons.

APPENDIX: Comment of Equation (3)

The exact relationship between the conditional probability of death and the corresponding mortality rate, or more precisely the force of mortality, involves an integral equation. It requires the integral of the force of mortality as a smooth function of age over the appropriate age band. The representation in equation (3) approximates this area by the product of the age-specific mortality rate,  $r$ , and the length of the age band,  $t$ . This is an appeal to the mean-value theorem for integrals, which will provide more accurate results with shorter age bands.

The general conclusions of this paper are not affected by this approximation. However the numerical estimates of over-estimation in the tables may differ slightly from those obtained by modeling the force of mortality by a smooth function of age and integrating this fitted function.

REFERENCES

1. Gaffey W.R.: A critique of the standardized mortality ratio. JOM 18: 157-160, 1976.
2. Wong, O.: Further criticisms on epidemiological methodology in occupational studies. JOM 19: 220-222, 1977.
3. Enterline, P.E.: Mortality among asbestos products workers in the United States. Ann. N.Y. Acad. Sci. 132: 156-165, 1975.
4. Selikoff, I.J., Churg, J., Hammond, E.C.: Asbestos exposure and neoplasia. JAMA 188: 22-26, 1964.
5. Enterline, P.E., DeCoufle, P., Henderson, V.: Respiratory cancer in relation to occupational exposures among retired asbestos workers. Br. J. Ind. Med. 30: 162-166, 1973.
6. Andjelkovic, D., Taulbee, J.D., Symons, M.J.: Mortality experience of a cohort of rubber workers, 1964-1973. JOM 18: 387-394, 1976.
7. Mantel, N., Haenszel, W.: Statistical aspects of the analysis of data from retrospective studies of disease. J. Natl. Cancer Inst. 22: 719-748, 1959.
8. Peto, R. and Peto, J.: Asymptotically efficient rank invariant test procedures (with discussion). J. Royal Stat. Soc. A. 135: 185-207, 1972.
9. Cornfield, J.: A method of estimating comparative rates from clinical data applications to cancer of the lung, breast and cervix. J. Natl. Cancer Inst. 11: 1269-1275, 1951.
10. Chiang, C.L.: A stochastic study of the life table and its applications: I. Probability distributions of the biometric functions. Biometrics, 16: 618-635, 1960.
11. Vital Statistics of the United States. Volume II - Mortality. Part A. Table 1-9: 1-26 to 1-43, 1971.
12. Yule, G.U.: On some points relating to vital statistics, more especially statistics of occupational mortality (with discussion). J. Royal Stat. Soc. 94: 1-84, 1934.
13. Hill, A.B.: Principles of Medical Statistics. Oxford University Press, New York, N.Y. (Sixth Edition), 1959.
14. Miettinen, O.S.: Standardization of risk ratios. Am. J. Epid. 96: 383-388, 1972.
15. Chiazzè, L.: Problems of study design and interpretation of industrial mortality experience. JOM 18: 169-170, 1976.
16. Enterline, P.E.: Pitfalls in epidemiological research, an examination of the asbestos literature. JOM 18: 150-156, 1976.

17. Milham, Jr., S: Methods in occupational mortality studies. JOM  
17: 581-585, 1975.

TABLE 1. Notation for the Mortality Experience for One Age Band in Two Populations

Population	Number of deaths due to cause of interest	Person-years or Mid-year population	Mortality rate	Probability* of dying
Study (exposed)	d	n	$r = d/n$	q
Comparison (unexposed)	D	N	$R = D/N$	Q

\* Full definition is in the text.

TABLE 2: Mortality Ratios by 10 Year Age Bands for All Malignant Neoplasms (ICDA Codes: 140-209) for 1971 U.S. White Males and a Fictitious Workforce with a Relative Risk of 200%

Age Group	Probability of Dying		Relative Risk (%)	Mortality Rate		Mortality Ratio (%)
	1971 U.S. White Males*	Fictitious Workforce		1971 U.S. White Males <sup>††</sup>	Fictitious Workforce*	
35-44	0.004898	0.009796	200	0.000491	0.000984	200.4
45-54	0.016778	0.033556	200	0.001692	0.003413	201.7
55-64	0.047905	0.095810	200	0.004909	0.010072	205.2
65-74	0.096392	0.192784	200	0.010136	0.021416	211.3

\* Derived by using equation (3) in the text; t = 10 years.

TABLE 3: Mortality Ratios by 10 Year Age Bands for Lung Cancer (ICDA Codes: 160-163)  
for 1971 U.S. White Males and a Fictitious Workforce with a Relative Risk of 300%

Age Group	Probability of Dying		Relative Risk (%)	Mortality Rate		Mortality Ratio (%)
	1971 U.S. White Males*	Fictitious Workforce		1971 U.S. White Males <sup>11</sup>	Fictitious Workforce*	
35-44	0.001469	0.004407	300	0.000147	0.000442	300.7
45-54	0.006668	0.020004	300	0.000669	0.002021	302.1
55-64	0.019448	0.058344	300	0.001964	0.006012	306.1
65-74	0.035398	0.106194	300	0.003604	0.011227	311.5

\* Derived by using equation (3) in the text; t = 10 years



TABLE 4: Mortality Ratio as a Function of Relative Risk and Length of Age Interval with a Mortality Rate of 100 per 10,000 Workers Per Year

Length of Age Band	Probability of Dying			Relative Risk RR%	Mortality Rate			Mortality Ratio MR%	Percentage Excess $\frac{(MR - RR)}{RR} \times 100$
	Comparison Population*	Study Population	RR%		Comparison Population	Study Population*	MR%		
5 years	0.048770	0.097540	200	0.010	0.020526	205.3	2.6		
5 years	0.048770	0.146310	300	0.010	0.031637	316.4	5.5		
5 years	0.048770	0.195080	400	0.010	0.043402	434.0	8.5		
10 years	0.095163	0.190326	200	0.010	0.021112	211.1	5.6		
10 years	0.095163	0.285489	300	0.010	0.033616	336.2	12.1		
10 years	0.095163	0.380652	400	0.010	0.047909	479.1	19.8		
20 years	0.181269	0.362538	200	0.010	0.022513	225.1	12.6		
20 years	0.181269	0.543807	300	0.010	0.039242	392.4	30.8		
20 years	0.181269	0.725076	400	0.010	0.064563	645.6	61.4		

\* Derived by using equation (3) in the text,  $t = 5, 10, \text{ or } 20$  years as indicated.