Bias in the Standardized Mortality Ratio when Using General Population Rates to Estimate Expected Number of Deaths

Michael E. Jones and Anthony J. Swerdlow

Cohort studies often compare the observed number of cases arising in a group under investigation with the number expected to occur on the basis of general population rates. The general population is taken to represent unexposed persons, but it is almost inevitably biased in that it comprises all types of people including exposed ones. To identify circumstances when this bias matters, the authors modeled its effect in relation to the size of the observed standardized mortality ratio (SMR) and the prevalence of exposed individuals in the general population. The authors found that bias may be a major problem, causing substantial underestimation of the true relative risk, when either the prevalence of exposure in the general population or the SMR are large. The bias can cause an apparent trend in SMRs with age when none exists. It also places a limit on the maximum size of the observed SMR, no matter how large the true relative risk. A table is provided showing the extent of bias in different circumstances. Cohort studies of people with common diseases or exposures, or that find large SMRs, when using general population expectations, need to consider the extent of bias from this source.


bias (epidemiology); mortality; standardized mortality ratio

Many cohort studies compare the observed number of cases arising in a cohort under investigation with the number expected to occur on the basis of general population rates. The investigator usually assumes that the disease rate in the general population reflects the rate that would occur in unexposed persons. Virtually all general populations, however, contain exposed individuals, if only because the cohort live within the population; it is usually not possible to determine rates for general populations excluding all exposed persons, because the latter are not separately identifiable. There is thus a bias in such studies (1), but its extent and the circumstances where it is of importance or trivial do not seem to have been investigated, and reports of cohort studies using general population expectations have generally not considered the problem.

We were initially concerned about this bias when analyzing renal disease mortality in a cohort study of patients with diabetes mellitus (2), since an appreciable proportion of renal disease in the general population is due to diabetes and there is no simple way to derive renal disease mortality rates for the general population excluding all people with diabetes. The problem, however, is more general, applying to other cohorts of persons with relatively common diseases or congenital conditions, and to cohorts defined, for example, by occupation, social class, or ethnic group, or a particular diet or use of a common medication, or exposure to a prevalent but not ubiquitous pollutant.

In this paper, we model the bias in the SMR and derive a mathematical expression for its size. We use the mathematical expression to investigate the extent to which the bias occurs in different circumstances, and hence when it is of practical importance.

METHODS

The bias described above occurs in cohort studies when the group being followed are considered to be exposed to the risk factor under investigation and general population rates are taken to represent the rate that would occur in an unexposed population. The extent of the bias depends on the prevalence of the exposure in the general population and the size of the true relative risk (rate ratio) in exposed compared with unexposed persons.

To estimate the true relative risk, let us first consider the situation where both exposed and unexposed groups are within the same cohort (i.e., using an internal comparison group). With a small extension to conventional notation (3), let $\lambda_{e,i}$ and $\lambda_{u,i}$ be the death rates in the exposed ($e$) and unexposed ($u$) population in stratum $i$ (where, for example, the exposed popula-
tion may be patients with diabetes and stratification is on age, sex, and period). The relative risk (rate ratio (RR)) in stratum i is given by

$$RR_i = \frac{\lambda_{e,i}}{\lambda_{u,i}}.$$

Let $n_{e,i}$ and $n_{u,i}$ be the person-years of follow-up in the exposed and unexposed parts of the cohort in stratum $i$. The expected number of deaths in stratum $i$, $E_i$, is given by

$$E_i = \lambda_{u,i} \times n_{e,i},$$

the observed number of deaths, $D_i$, is given by

$$D_i = \lambda_{e,i} \times n_{e,i},$$

and, as usual,

$$\text{SMR} = \frac{\sum D_i}{\sum E_i} = \frac{\sum \lambda_{e,i} \times n_{e,i}}{\sum \lambda_{u,i} \times n_{e,i}},$$

where $\sum$ is the summation over all $i$ strata. Assume $RR_i$ is constant over strata (i.e., $RR_i = RR$ for all $i$) and thus,

$$\lambda_{e,i} = RR \times \lambda_{u,i}.$$

Hence, by substitution it can be shown that:

$$\text{SMR} = RR,$$  \hspace{1cm} (1)

which is to say that the SMR correctly estimates the true relative risk when the expected number of deaths is estimated from the rate in the unexposed group.

If the expected number of deaths is estimated from general population rates, a bias is introduced into the SMR. Let $\lambda_{g,i}$ be the rate in the general population in stratum $i$. The expected number of deaths in stratum $i$, $E_{g,i}$, is given by $E_{g,i} = \lambda_{g,i} \times n_{e,i}$, and the expression for the observed number of deaths remains the same,

$$D_i = \lambda_{e,i} \times n_{e,i}.$$

The general population rate, $\lambda_{g,i}$, is, however, an average of the rate among the exposed and unexposed parts of the general population. If $p_{e,i}$ is the prevalence of exposure in the general population in stratum $i$, and so $(1 - p_{e,i})$ is the prevalence of not being exposed in stratum $i$, then

$$\lambda_{g,i} = p_{e,i} \times \lambda_{e,i} + (1 - p_{e,i}) \times \lambda_{u,i}.$$

Assume again that $RR_i$ is constant over strata $i$ (i.e., $RR_i = RR$ for all $i$), and in addition $p_{e,i}$ is constant over strata $i$ (i.e., $p_{e,i} = p$ for all $i$), or that our analysis is limited to those strata where this is so, then the usual combination of observed deaths divided by expected deaths gives the observed SMR:

$$\text{SMR} = \frac{\sum D_i}{\sum E_{g,i}} = \frac{\sum \lambda_{e,i} \times n_{e,i}}{\sum \lambda_{g,i} \times n_{e,i}}.$$

Substituting for $\lambda_{g,i}$ gives:

$$\text{SMR} = \frac{\sum \lambda_{e,i} \times n_{e,i}}{\sum [\lambda_{e,i}(1 - p_{e,i}) + (1 - p_{e,i}) \times \lambda_{u,i}]} \times \frac{1}{n_{e,i}}.$$

Substituting $RR \times \lambda_{u,i}$ for $\lambda_{e,i}$ and $p$ for $p_{e,i}$ gives:

$$\text{SMR} = \frac{\sum RR \times \lambda_{u,i} \times n_{e,i}}{\sum [p \times RR \times \lambda_{u,i} \times n_{e,i} + (1 - p) \times \lambda_{u,i} \times n_{e,i}]} = RR \times \frac{\sum \lambda_{u,i} \times n_{e,i}}{[p \times RR + (1 - p)] \times \lambda_{u,i} \times n_{e,i}}.$$

The $\sum \lambda_{u,i} \times n_{e,i}$ term in the numerator and denominator cancel out, giving:

$$\text{SMR} = RR \times \frac{[p \times RR + (1 - p)]}{[1 - p \times RR]}.$$

which is a biased estimate of the true relative risk (rate ratio) when $p > 0$. The bias is always toward the null. The above expression also implies that there is an upper bound to the observed SMR, which tends to $1/p$ as the true relative risk becomes very large.

Rearrangement of the formula for the observed SMR gives an expression for the true relative risk, as follows:

$$\text{SMR} \times [p \times RR + (1 - p)] = RR,$$

giving

$$RR = \text{SMR} \times (1 - p)[1 - (p \times \text{SMR})].$$

In practice, the SMR is estimated from real data containing random error and the true relative risk given by equation 3 is not strictly "true" because it too is subject to this random error.

The size of the bias, expressed as percent of the SMR, may be estimated as:

$$\text{Bias} = \frac{(RR - \text{SMR})}{\text{SMR}} \times 100.$$  \hspace{1cm} (4)

If the same approach is taken with the comparative mortality factor (CMF) (3), where $w_i$ are the weights used in the direct method of standardization (e.g., the world standard population), we get the same result except that the observed SMR is replaced by the observed CMF:

$$\text{CMF} = \frac{\sum (\lambda_{e,i} \times w_i)}{\sum (\lambda_{g,i} \times w_i)} = RR[(1 - p) + p \times RR].$$

$$\text{CMF} = \frac{\sum (\lambda_{e,i} \times w_i)}{\sum (\lambda_{g,i} \times w_i)} = RR[(1 - p) + p \times RR].$$

$$\text{CMF} = \frac{\sum (\lambda_{e,i} \times w_i)}{\sum (\lambda_{g,i} \times w_i)} = RR[(1 - p) + p \times RR].$$

and

\[ RR = CMF \times (1 - p) / [1 - (p \times CMF)]. \] (6)

We used these derived expressions to produce a table of true relative risks for various values of the observed SMR against a range of values for the prevalence of exposure in the general population. The table is equally applicable to the bias in the CMF. We also estimated the true relative risk and the size of the bias for some published studies where the bias may be of importance.

RESULTS

Table 1 shows the estimated relative risk (rate ratio) for a given observed SMR at different values of \( p \), the prevalence of exposure to the risk factor under study in the general population, if general population rates are used to estimate the expected number of cases. For example, if the observed SMR was 5.0 and the prevalence of exposure in the general population was 5 percent, the true relative risk would be 9.0. The lower right hand section of the table is blank because when the relative risk is greater than unity there is an upper bound to the observed SMR for any given prevalence of exposure in the general population. For instance, if 20 percent of the general population are exposed in the same way as the cohort under study, the SMR cannot exceed 5.0 regardless of the size of the true relative risk. The bias is not so pronounced when the SMR is smaller than unity.

Figure 1 shows the degree of bias for a series of SMRs in relation to the prevalence of exposure in the general population. If the observed SMR is modestly raised (e.g., 1.5), the bias will only be large if a considerable proportion (e.g., \( \geq 20 \) percent) of the general population is exposed. For large SMRs, however, even a low prevalence of exposure may cause a large bias.

As an example, consider the cohort of patients with type 2 diabetes mellitus diagnosed at a hospital in Osaka Prefecture, Japan (4). The SMR for renal disease was 12.4 for males and 16.4 for females when compared with the death rate in the general population of Osaka Prefecture. Assuming the prevalence of diabetes in Osaka Prefecture was at least 4 percent (5), the relative risk of renal disease mortality would be 90 percent larger for men and 180 percent larger for women than the reported SMR.

The bias may be important even for rarer conditions. Insulin-dependent diabetes mellitus (IDDM) had a cumulative incidence to age 14 years ranging from 0.14 percent to 0.28 percent in US white or predominantly white populations during 1965 to 1988 (6). Among the cohort of IDDM patients diagnosed before age 17 years at the Children’s Hospital of Pittsburgh, the SMRs for renal disease as a cause of death were 556

<table>
<thead>
<tr>
<th>Observed SMR</th>
<th>Prevalence of exposure in general population (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>0.5</td>
<td>0.50</td>
</tr>
<tr>
<td>1.0</td>
<td>1.00</td>
</tr>
<tr>
<td>1.5</td>
<td>1.50</td>
</tr>
<tr>
<td>2.0</td>
<td>2.00</td>
</tr>
<tr>
<td>2.5</td>
<td>2.50</td>
</tr>
<tr>
<td>3.0</td>
<td>3.00</td>
</tr>
<tr>
<td>4.0</td>
<td>4.00</td>
</tr>
<tr>
<td>5.0</td>
<td>5.00</td>
</tr>
<tr>
<td>7.5</td>
<td>7.50</td>
</tr>
<tr>
<td>10</td>
<td>10.01</td>
</tr>
<tr>
<td>25</td>
<td>25.06</td>
</tr>
<tr>
<td>50</td>
<td>50.25</td>
</tr>
<tr>
<td>75</td>
<td>75.56</td>
</tr>
<tr>
<td>100</td>
<td>101.0</td>
</tr>
</tbody>
</table>

*Bold italic: true relative risk is at least 10% different from the observed SMR.
† Lower right hand section of the table is blank because the bias causes an upper bound to the observed SMR for relative risks greater than unity.
in white males and 561 in white females, when compared with rates in the US white population (7). If the prevalence of IDDM was 0.14 percent during the period under study, the actual relative risk for renal disease death would be at least 4.5 times larger, around 2,500. Even if the prevalence was as low as 0.10 percent, the actual relative risk would be at least twice the observed SMR.

Studies of morbidity can also be affected by this bias. Cryptorchidism affects about 1.5 percent of boys (8) and is a strong risk factor for testicular cancer (9). It has not been feasible to provide a comparison group by following a large cohort of boys who never suffered from cryptorchidism, and cohort studies of cryptorchid boys have compared their incidence of testicular cancer with incidence rates from population-based cancer registries. In the largest such cohort study, the standardized incidence ratio for testicular cancer was 7.5 (9), which, given the prevalence noted above, must have underestimated the true relative risk by about 10 percent.

The bias can also occur in cohorts defined by occupation, social class, or ethnic group. In England and Wales, 6.9 percent of males aged 15–64 years were classified as social class V (unskilled occupations) in the 1970–1972 national analysis of occupational mortality (10). The published SMR for homicide was 3.4 in this group, compared with the rate in the male population of England and Wales. The true relative risk would have been 4.1 if social class V had been compared with the remainder of the male population, an increase of 22 percent on the observed SMR. The SMR for rheumatic heart disease among Australian aborigines in the Northern Territory of Australia was reported to be 39, when compared with national Australian death rates (11). Northern Territory aborigines make up about 0.24 percent of the total Australian population (12), so the true SMR would be 10 percent larger if they were excluded from expectations. If, as seems likely, the raised risk of rheumatic heart disease is true for Australian aboriginal people overall (1.6 percent of the total Australian population (12)), rather than just those who live in the Northern Territory, the estimated SMR would be 162 percent larger.

The results presented so far assume that the relative risk is constant across strata by which the SMR has been standardized (e.g., by age, sex, and calendar period). This is a usual assumption when calculating SMRs (3), but if the relative risk is heterogeneous across strata, it is arguable whether standardized summary measures such as SMRs should be used at all. The prevalence of exposure in the general population is also assumed to be constant across strata, but for many chronic diseases and several other exposures this
latter assumption is not even approximately true. For numerous chronic diseases, for instance, prevalence is far greater at older than at younger ages. This changing prevalence by age (or sex or other stratification variable) can then lead to an artefactual trend in the SMR by age (or sex, etc.) as the extent of the bias alters from slight, at ages when prevalence is low, to great at ages when the prevalence is much higher. An example is shown in table 2. The prevalence of non-insulin-dependent diabetes mellitus in US non-Hispanic whites (13) varies from under 0.1 percent at ages 20–24 years to over 20 percent at ages 70–74 years. If the true relative risk for a particular cause of death in patients with diabetes was constant by age at 6.0, then, as shown in table 2, the observed SMR would appear to decrease by more than 50 percent with increasing age simply because of the bias. Cohort studies of mortality among diabetic patients do indeed show all-cause SMRs decreasing with increasing age after age 30 years (2, 14–17), and although the decreases are generally larger than can be attributed to this bias alone, the bias must account for a proportion of the downward trend.

**DISCUSSION**

We have shown that there is a potential bias in the standardized mortality or morbidity ratio when general population rates are used to estimate the expected number of deaths in a cohort. The same bias can affect the comparison of directly standardized rates (CMFs), because the only difference between indirect standardization (SMRs) and direct standardization (CMFs) is the set of weights used as the standard (18).

As we have shown, the bias is conservative, i.e., toward a relative risk of unity, and consequently the observed number of deaths will appear less extreme than it should when compared with the calculated expected number of deaths. The bias will therefore make p values less significant, because the p value is usually calculated by estimating the probability that the observed number of deaths, or a more extreme number, could have arisen from a Poisson distribution characterized by the expected number of deaths (3).

The ideal solution to this bias would be to follow up an unexposed group of individuals, in addition to those exposed. This solution is often impracticable because of the cost of ascertaining and monitoring a sufficiently large unexposed group for an adequate length of time, or the difficulty in obtaining an unexposed but otherwise comparable group (for instance, for an occupational cohort). Alternatively, general population rates could be modified by subtracting the deaths and person-years experienced by the exposed cohort, if the study cohort itself constituted all exposed persons in the general population. Another possibility, if the prevalence of exposure was known exactly, would be to estimate the true relative risk from equation 3, then use this to calculate the unbiased expected number of deaths, the p value, and confidence interval.

Generally, the prevalence is not known precisely and no exact correction for the bias can be made; the latter approach suggested above can be regarded as a qualitative assessment of the effect of the bias rather than a definitive adjustment for it. Our results, however, show that where the effect of contamination by exposed individuals is small and the SMR is modest, the bias may be ignored. Where the SMR is large or the exposure is not rare in the general population, and especially where both apply, the bias can be appreciable. Furthermore, we have shown that the bias gives an artefactual upper limit to the observed SMR in studies using general population expectations, the magnitude of this limit depending on the general population prevalence of the exposure. Thus, where prevalence is high, large SMRs will not be observed, no matter how large the true relative risk. Indeed, there may be occasions when the bias is so strong that it is inappropriate to calculate SMRs. The bias will also lead to apparent trends in the SMR by age or other such factors if the prevalence of exposure in the general population varies by these factors. Studies using general population expectations therefore need to consider the impact of the bias on their results and conclusions, especially if the prevalence of exposure changes with age, sex, or other analytic variables.
ACKNOWLEDGMENTS

The Epidemiological Monitoring Unit is funded by Medical Research Council (UK) programme grant no. G8119454. Dr. Michael E. Jones is funded by Medical Research Council (UK) project grant no. G9533916.

REFERENCES