Is the overall SMR a useful measure of mortality in occupational cohort studies?

Ånund Hobbesland

Telemark Central Hospital, Department of Occupational and Environmental Medicine, N-3710 Skien, Norway

Telephone + 47 35 00 35 00  Telefax + 47 35 00 31 05  E-mail anund.hobbesland@tss.telemax.no

ABSTRACT

Objectives. This study was performed to examine the importance of the age structure of a cohort and the impact of the healthy worker effect (HWE) on the overall standardised mortality ratio (SMR), and to investigate the components of the HWE.

Methods. A previously defined cohort of 14,730 male workers from 12 Norwegian ferroalloy plants was studied with analyses of the overall SMR stratified by age at risk, employment status, periods of observation and duration of follow-up. The age distribution and the proportions of active and inactive person-time were also examined among workers from each of the 12 plants. An alternative SMR notion designated SMR_{ptw} is presented which is the sum of the stratum specific observed/expected ratios weighted by the person-time distribution.

Results. The proportions of person-time attributed from the highest age stratum (≥ 65 years) and the proportions of all inactive person-time varied considerably between the populations from the 12 plants. Among individuals aged < 65 years the overall SMRs increased by increasing proportions of inactive person-time. The inactive person-time stratum among individuals aged 65 years or more showed an SMR close to 1.00. The active person-time SMRs were fairly stable over increasing strata of duration of follow-up. A traditional SMR of 1.08 and an SMR_{ptw} of 1.18 were estimated for this cohort.

Conclusions. Traditional SMR analyses of occupational cohorts are in general influenced by the distribution of age and active/inactive person-time. To reduce the influence of random variations in these distributions it is proposed to perform person-time weighting of the age, employment status, and calendar time strata. The healthy hire effect and the healthy worker survivor effect (HWSE) are the main components of the HWE which did not decline by increasing duration of follow-up.

INTRODUCTION

The Standardised Mortality Ratio (SMR) has been used for a long time as a measure of mortality in occupational cohort studies. It is the ratio of the sums of the observed and expected cases in the exposed group (1). The expected value is based on the race, sex, age, and calendar time specific mortality rates in a reference population multiplied with the corresponding person-years of observation in the exposed group. The SMR is thus dependent on the age and calendar time distribution of the exposed group. The reference population is commonly a large regional or the whole national population.

One of the most important limitations in the interpretation of the overall SMR is the well-known phenomenon called “the healthy worker effect” (HWE) which contributes to incorrectly low overall SMRs. Already in 1885 Ogle presented that an occupational cohort may be selected by the time of hire and also during employment which both are important parts of the HWE (2). Many authors have discussed the elements of this topic during the last decades (e.g. 3-15). Most have agreed with Ogle in the initial selection of healthy individuals to the work force which recently has been referred to as the “healthy hire effect” (12). The survival of healthy individuals within the work force has also been considered an important part of the HWE (4,7,9,10). This phenomenon is known as the “healthy worker survivor effect” (HWSE), but the nature of this effect is still not quite clear (12). The HWSE is at least partly explained by the selection out from the workforce of individuals who develop disease, but it has also been argued that the employment itself may imply improved health due to better economic status, better access to medical care and positive changes in lifestyle which may contribute to the healthy worker effect (7).

The healthy hire effect and the HWSE are both expressed through low mortality rates during the employment period. The person-time at risk when a person is employed has been called “active” person-time. The “inactive” person-time is the person-time after termination of employment at the plant under
The inactive person-time shows mortality rates which commonly are 2-3 times the rates during active person-time (16,17). The variable used to categorise the person-years in active and inactive person-time is known as the employment status (16). The importance of the employment status was investigated in a large US cohort study which comprised ten occupational cohorts where no occupational risk had been observed (16). In this study 89 376 workers were observed for 1 984 505 person-years (46% active and 54% inactive). Overall SMRs of 0.39 and 1.12 were observed for active and inactive person-time, respectively. The corresponding SMRs for ischaemic heart disease were 0.55 and 1.14 and for cancers 0.28 and 1.14, respectively. The results indicate that the mortality pattern of active and inactive person-time concerns occupational cohorts in general and should be evaluated in such studies (16). The explanations to the increased mortality during inactive person-time is not further discussed in this paper.

In addition to the healthy hire effect and the HWSE many authors also include the “time since hire” or the duration of time the population has been followed up as a modifier of the HWE (4,7,9,10,12). In the recent years it has been claimed, however, that the time since hire should not be treated as a positive confounder as the mortality has been found not to increase with time since hire (or duration of follow-up) in actively employed workers (13).

The other main limitation of the SMR is the influence from the age structure of the exposed group on the summary SMR. In the present paper this phenomenon is considered to be a limitation of the SMR methodology even if some authors consider it a part of the HWE. An argument for excluding it from the HWE-concept is that it arises as an effect of the way the SMR is calculated and is not specifically associated with a low overall SMR as might be expected if it was a HWE. It has been shown that SMRs of 0.88 and 1.32 for all causes of death in two cohorts may be explained only by different age structures as the age-specific mortality ratios were the same (7). These cohorts showed low mortality rates in the lower age strata and high mortality rates in the higher age strata. Though this example is illustrative it is probably not typical for occupational cohorts. In the aforementioned study of the large US cohort it was observed that the overall SMR for inactive person-time decreased from a range of 1.22–1.28 for age categories below 65 years to 1.00 for age at risk of 65 years or more (16). The same pattern was observed for heart disease and for all cancers. Thus, as this may be a common pattern of the mortality rates of an occupational cohort the proportion of person-years attributed from persons above 65 years (or from another retirement age) will influence the overall SMR. Thus, the age structure of a cohort should be evaluated in occupational mortality studies.

In order to investigate the extent of influence on the overall SMR from the age structure and the HWE related variables employment status and time since hire, a re-examination was performed of a Norwegian cohort of ferroalloy workers from which studies of mortality also previously have been reported (18-20). From this cohort a slightly increased SMR of 1.08 (3390 deaths) from all causes of death was reported. During active person-time the overall SMR was significantly increased (SMR 0.71) while the SMR was significantly increased (SMR 1.17) during inactive person-time. The active person-time SMRs did not increase by increasing duration of employment (range of SMRs 0.69–0.72 for duration of employment < 3, 3-9, 10-19, and 20 years or more). For age at risk < 30, 30-49, 50-69, and 70 years or more the overall SMRs were 1.35, 1.18, 1.11, and 1.00, respectively (18).

The aim of this paper is to study how variations in the age structure and the distribution of active and inactive person-time may influence the overall SMR. The potentially modifying effect on the overall SMR from duration of follow-up is also examined. Also an alternative to the traditional SMR analysis is proposed.

**Material and methods**

The cohort consists of 14 730 men who had been employed at least six months and for the first time during 1933–1990 in any of 12 Norwegian ferroalloy plants. The information on individual employment periods and kind of work was taken from the manual employment records where individual employment data had been recorded by personnel office workers. Up to ten employment periods were recorded in a database for each man. The vital data was controlled and the information on individual date of death and cause of death supplied by linking the cohort to Statistics Norway (the national register for causes of death). Only 3% (462 men) of the primary male cohort was lost to follow-up. The cohort was observed for deaths during 1962–1990 and comprised of 288 886 person-years.

The expected numbers of deaths were calculated by use of 5-year age specific rates for each year of observation from 1962 throughout 1990 for the total Norwegian male population. The counting of deaths and the calculation of the expected numbers was performed with the Epicure programme package which also created event time tables which might be stratified by fixed and time dependent variables (21). Active person-time was defined as all person-time from the beginning of follow-up until the end of last employment added by 0.1 year as the dates of termination of employment were inaccurate for some of the workers. Active person-time thus included periods off work before the last date of termination of employment for individuals with more than one employment period. All deaths which occurred while the men were employed in one of these plants or within 0.1 year...
after termination of employment were attributed to the active person-time category of the employment status variable. Inactive person-time was defined as the remaining person-time of observation (for those who terminated work alive) and comprised all deaths during this period. The active person-time category comprised 38% of all person-years of observation. Further details about the cohort and the analyses including calculation of 95% confidence intervals (95% CI) of the SMRs have been published elsewhere (18).

It is shown in the Results section that the overall SMR is influenced by the distribution of age and active/inactive person-time in the study cohort. This is a negative aspect of the SMR methodology as SMRs from various cohorts cannot be compared. Other measures of mortality may thus be wanted. In this paper results are presented from an alternative estimation based on a weighted average of the stratum-specific rate ratios of e.g. age where the stratum-specific person-years of observation are used as the weights (wi). This estimation is derived from the general standardised rate ratio (RRi) which may take the form

$$RR_i = \frac{\sum w_i (a_i / N_{i1})/(b_i / N_{i0})}{\sum w_i}$$

where $a_i$ and $b_i$ are the numbers of events in the exposed and non-exposed groups, respectively. $N_{i1}$ and $N_{i0}$ are the person-years of the exposed and non-exposed groups, respectively (1). The stratum specific rate ratios equal the stratum specific ratios of the observed (O) and expected (E) values from SMR analyses as the person-years for each stratum are identical for the exposed and reference groups. It is thus possible to weight the stratum specific O/E-ratio by the person-years of the stratum. This person-time weighted SMR is here designated SMRptw. Based on equation (1) the SMRptw takes the form

$$SMR_{ptw} = \frac{\sum pyr_i (O_i / E_i)}{\sum pyr_i}$$

where $pyr_i$ is the person-years of the $i^{th}$ stratum (of e.g. age). The stratum specific numbers of observed and expected deaths are the same as for the traditional SMR analyses.

**RESULTS**

Table 1 shows the distribution of deaths, the person-years, and the crude mortality rates in four strata of age at risk. The lowest age-stratum (< 45 years) comprised 48% of all person-years of observation but only 8% of all deaths (SMR 1.26). The highest age stratum (65 years or more) comprised 54% of all deaths but only 12% of all person-time (SMR 1.02). The corresponding crude mortality rates were 2.0 and 51.0 deaths per 1000 person-years in these two strata, respectively.

<table>
<thead>
<tr>
<th>Age at risk</th>
<th>No. of deaths</th>
<th>Person-time years</th>
<th>Crude rate</th>
<th>SMR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 45 years</td>
<td>280</td>
<td>138 797</td>
<td>48</td>
<td>2.0</td>
<td>1.26</td>
</tr>
<tr>
<td>45-54 years</td>
<td>392</td>
<td>62 713</td>
<td>22</td>
<td>6.3</td>
<td>1.10</td>
</tr>
<tr>
<td>55-64 years</td>
<td>877</td>
<td>51 280</td>
<td>18</td>
<td>17.1</td>
<td>1.15</td>
</tr>
<tr>
<td>≥ 65 years</td>
<td>1841</td>
<td>36 097</td>
<td>12</td>
<td>51.0</td>
<td>1.02</td>
</tr>
</tbody>
</table>

1 Deaths per 1000 person-years

Figure 1 shows the proportions of person-time in the same four age-strata in the sub-cohorts from each of the 12 plants. The proportion of person-time attributed from the highest age stratum ranged from 3% (plant 12) to 25% (plant 11). The corresponding proportions of deaths in the highest age stratum compared to all deaths in each sub-cohort ranged from 52% in plant 12 to 77% in plant 11 (not shown).

Table 2 presents the age-specific mortality stratified by active and inactive person-time. The SMRs in active person-time were in the range of 0.84–0.66 and were slightly decreasing by increasing age. The deaths during active person-time comprised 14% of all deaths. The SMRs during inactive person-time ranged from 1.35 to 1.61 in the three age strata below 65 years while an SMR of 1.03 was found for inactive person-
Time of observation, particularly from 65 years or more. This category of inactive person-time among men aged 65 years or more comprised 52% of all deaths but only 11% of all person-time (not shown).

Figure 2 shows the distribution of all active and inactive person-time in the sub-cohorts from each of the 12 plants. The overall proportions of inactive person-time ranged from 22 to 70% (plant 12 and 1, respectively). The proportions of inactive person-time for age at risk below 65 years ranged from 20 to 67% (not shown). An overall SMR of 0.68 was found for age at risk below 65 years in the sub-cohort with the lower proportion (20%) of inactive person-time. A high overall SMR of 1.41 was found for the corresponding sub-cohort with the highest proportion (67%) of inactive person-time.

The proportions of active and inactive person-time may also vary over different parts of the observation period which is shown in Table 3. For age at risk below 65 years the proportions of inactive person-time during 1962-1970, 1971-1980, and 1981-1990 were 49%, 54% and 62%, respectively. The corresponding SMRs for all person-time during these successive periods of observation were 1.04, 1.15 and 1.23. For the age stratum of 65 years or more the proportions of inactive person-time were 74%, 89%, and 97%. However, the overall SMRs for the same periods of observation were 1.06, 1.04, and 1.00 for this age-stratum.

Table 4 shows the SMRs in four strata of duration of follow-up. For these strata only a small increase in the SMRs was found during active person-time for increasing duration of follow-up (range of SMRs 0.61–0.80). For the corresponding inactive person-time categories the SMRs decreased from 1.39 to 1.10.

By use of equation (2) overall person-time weighted SMR$_{ptw}$s for the total cohort were computed. Person-time weighting of the 5-year age strata showed an SMR$_{ptw}$ of 1.17 (95% CI 1.13–1.21). The result means that the mortality for this cohort is at average 17% higher for each person-year of observation than in a comparable national population. An overall SMR$_{ptw}$ was further calculated weighted by the person-

### Table 2. SMRs of active and inactive person-time categorised by age at risk.

<table>
<thead>
<tr>
<th>Age at risk</th>
<th>Active person-time</th>
<th>Inactive person-time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obs.</td>
<td>SMR</td>
</tr>
<tr>
<td>&lt; 45 years</td>
<td>83</td>
<td>0.84</td>
</tr>
<tr>
<td>45-54 years</td>
<td>114</td>
<td>0.76</td>
</tr>
<tr>
<td>55-64 years</td>
<td>208</td>
<td>0.66</td>
</tr>
<tr>
<td>≥ 65 years</td>
<td>62</td>
<td>0.67</td>
</tr>
</tbody>
</table>
time distribution of both age and employment status. Each 5-year age-stratum from 15 to 70 years was then estimated from 43% of the stratum specific O/E-ratio of active person-time and 57% of the O/E-ratio of inactive person-time which were the average proportions of active and inactive person-time below 70 years. Above 70 years of age at risk almost all person-time was inactive. The $SMR_{ptw}$ then became 1.18. This shows that the distribution of active and inactive person-time was fairly equal over the age strata below 70 years. The effect on the overall $SMR_{ptw}$ from random variations of the distribution of active and inactive person-time was examined by using the proportions from plant number 1 (70% inactive person-time) and plant number 12 (22% inactive person-time). The overall $SMR_{ptw}$ for the total cohort adjusted for age and employment status then became 1.27 and 0.94, respectively.

**Table 3.** Proportion of inactive person-time, observed (obs.) numbers of deaths and SMRs in different parts of the observation period categorised by age at risk.

<table>
<thead>
<tr>
<th>Period of observation</th>
<th>% inactive person-time</th>
<th>Obs.</th>
<th>SMR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at risk &lt; 65 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1962–1970</td>
<td>49%</td>
<td>339</td>
<td>1.04</td>
<td>0.94–1.16</td>
</tr>
<tr>
<td>1971–1980</td>
<td>54%</td>
<td>589</td>
<td>1.15</td>
<td>1.06–1.25</td>
</tr>
<tr>
<td>1981–1990</td>
<td>62%</td>
<td>621</td>
<td>1.23</td>
<td>1.14–1.33</td>
</tr>
<tr>
<td>Age at risk ≥ 65 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1962–1970</td>
<td>74%</td>
<td>191</td>
<td>1.06</td>
<td>0.92–1.22</td>
</tr>
<tr>
<td>1971–1980</td>
<td>89%</td>
<td>554</td>
<td>1.04</td>
<td>0.95–1.13</td>
</tr>
<tr>
<td>1981–1990</td>
<td>97%</td>
<td>1096</td>
<td>1.00</td>
<td>0.94–1.06</td>
</tr>
</tbody>
</table>

**Figure 2.** The proportions of active and inactive person-time of observation of workers in 12 ferroalloy plants.

**Table 4.** SMRs of active and inactive person-time categorised by length of follow up.

<table>
<thead>
<tr>
<th>Duration of follow-up</th>
<th>Active person-time</th>
<th>Inactive person-time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obs.</td>
<td>SMR</td>
</tr>
<tr>
<td>&lt; 3 years</td>
<td>63</td>
<td>0.61</td>
</tr>
<tr>
<td>3-9 years</td>
<td>156</td>
<td>0.71</td>
</tr>
<tr>
<td>10-19 years</td>
<td>170</td>
<td>0.72</td>
</tr>
<tr>
<td>≥ 20 years</td>
<td>78</td>
<td>0.80</td>
</tr>
</tbody>
</table>

The various plants

Inactive person-time
Active person-time
DISCUSSION

At least four main results may be summarised from this study. Firstly, the age distribution and the distribution of inactive person-time may be quite different in various cohorts and in different parts of the observation period. Below age of 65 years the overall SMR was found to be closely related to the proportion of inactive person-time (table 3). Secondly, the mortality during active person-time (i.e. the employed period) is significantly lower than during inactive person-time. The SMRs for active person-time are remarkably stable over various categories of age at risk (SMRs 0.66–0.84), duration of follow-up (0.61–0.80), and duration of all employment (SMRs 0.69–0.72) which has previously been presented (18). Thirdly, for the inactive person-time category of age at risk of 65 years or more the mortality rate was close to the expected value (SMR 1.03). This category thus differs from lower age-strata of inactive person-time where the SMRs for inactive person-time ranged from 1.35 to 1.61. A great proportion of person-time attributed from the highest age stratum will thus decrease the overall SMR towards 1.00 while a great proportion of inactive person-time below 65 years will increase the overall SMR in the present cohort. Fourthly, the overall SMRptw for the total cohort adjusted for age and active/inactive person-time was slightly higher (1.18) than the traditional overall SMR (1.08) while application of weights for active/inactive person-time from two of the sub-cohorts gave SMRptw values of 1.27 and 0.94.

The distribution of age and active/inactive person-time may thus influence the overall SMR. Various more or less random conditions may affect these distributions. Firstly, the age of a plant is important as increasing age will increase the average age of the cohort members and increase the proportion of inactive person-time during the follow-up period. In the present study plant number 12 and 8 are the two youngest plants (established in 1968 and 1973, respectively) and are the two plants with the lowest proportion of person-years from the age-stratum of 65 years or more (3 and 5%, respectively). These two plants also showed the lowest proportions of inactive person-time (22 and 37%, respectively). Plant number 11 was the third youngest plant and showed the third lowest proportion of inactive person-time (41%). However, the cohort members from this plant had the highest proportion of person-time attributed from men aged 65 years or more (25%). This plant was established in the beginning of the 1960s at the same place where an old plant recently had been shut down. During the first years of production the employer may have recruited workers from the workforce of the previous plant which implied that they at average were older than other workers in this industry at the time of first employment. Thus, the workers who are recruited during the first years may have an age structure which is different from the later employed workers. Secondly, the choice of inclusion and exclusion criteria in relation to the observation period may influence the age structure and the distribution of active and inactive person-time. In the present study the recruitment of workers started in 1933 while the observation period began in 1962. This implied that 46% of the cohort members began employment before start of follow-up and 20% were only observed during the inactive period. As might be expected a slightly lower mortality was found for workers employed first time before 1962 (SMR 1.06, 95% CI 1.02–1.10) than among those employed in 1962 or later (SMR 1.13, 95% CI 1.05–1.21). Thirdly, if a plant has closed down the group of workers from this plant will change from a dynamic population to a fixed cohort. The proportion of inactive person-time and the average age of the cohort members are supposed to increase more rapidly in fixed cohorts than in dynamic populations where younger workers are continuously recruited.

Most of the above results are in good agreement with the corresponding results from the large US cohort (16). However, the active person-time SMR was only 0.39 for all deaths in that study. The difference is partly explained by the diverging stratification of active and inactive person-time in these two studies. In the present study active person-time was defined by adding 0.1 year to the last day of employment while one week was added in the US cohort study (16). When this definition of active person-time was applied in the present study an overall active person-time SMR of 0.49 (95% CI 0.44–0.55) and inactive person-time SMR of 1.23 (95% CI 1.19–1.27) was found. The original SMRs in four strata of duration of follow-up indicated a slight decrease of the HWE with increasing duration of follow-up. However, when the same analyses were performed with the same stratification of active and inactive person-time as in the US cohort the active person-time SMRs in the four strata of duration of follow-up became almost identical (range of SMRs 0.47–0.52). These results thus indicate that the HWE is not modified by duration of follow-up.

The applied weights are important for the SMRptw. An alternative to taking weights from the person-time distribution of the cohort is to use the age-distribution of the general population for a particular year or the average of the follow-up period. However, this has the disadvantage that for the youngest strata of age the general population contains a high proportion of individuals of e.g. 15-19 years while this age stratum attributes very few person-years in a study of an occupational cohort. Weighting by the person-time distribution of the general population should therefore not include weighting of the youngest age-strata of a cohort, e.g. < 20 years. A preferable alternative is to take the weights from the age distribution of a large national occupational cohort where no major exposure-effect association has been found. The
weights for adjustment of the distribution of active and inactive person-time should also be taken from a large national occupational cohort. From the large US cohort study where active person-time constituted 46% of all person-time (16) and the present cohort it seems in general reasonable to apply weights of approximately 45% active and 55% inactive person-time below retirement age which for the present cohort gives an SMRptw of 1.17. If the active person-time for the present cohort had been more strictly defined (as in the US cohort) the SMRptw would have increased.

This study has focused on the measure of the overall mortality in occupational cohort studies. From the above results it seems natural to recommend that an overall SMRptw is calculated in addition to or instead of the traditional SMR. The SMR (and the SMRptw) has its greatest advantage in the use of easily accessible and stable national mortality rates for estimation of expected values (1), and in being simple to understand and communicate with non-professionals. The SMRptw also allows application of external weights which thus is not influenced by the confounder distribution of the cohort. The overall SMR has traditionally been emphasised by researchers in the discussion of results from a mortality study. An observed positive trend of mortality (e.g. by increasing exposure) may be disregarded if the overall SMR or cause-specific SMR is not increased. There is therefore a need to emphasise that the traditional SMRs are influenced by several random conditions discussed above. It is thus important to include person-time weighting of age and employment status in the SMRptw. In studies with long follow-up periods it may be advisable to weight the calendar time specific rates by the person-years of observation, in particular if the rates change markedly over the follow-up period. The importance of person-time weighting of the calendar time specific rates has not been examined in the present cohort. In general, person-time weighting of age, employment status and calendar time is assumed to give a more valid measure of mortality, both for the overall mortality and for cause specific mortality. The SMRptw expresses the average mortality per person-year of observation. An advantage is that the results from different studies would be comparable when calculated in the same way with application of standard person-year weights for the age distribution (e.g. from a national occupational cohort) and for employment status below retirement age (e.g. 45% active and 55% inactive person-time).

The random age distribution of various cohorts is an important aspect also for computation of standardised incidence ratios (SIR), e.g. in studies of cancer incidence. The incidence of most cancers increases considerably by increasing age. Person-time weighting of age is thus recommended also in the estimation of SIRs for cancer. The employment status variable has not the same importance in cancer incidence as in mortality analyses (including cancer mortality). There are commonly only small differences in the cancer incidence rates during active and inactive person-time. However, estimation of SIRptw weighted by the same variables as recommended for SMRptw (age, employment status and calendar time) is assumed to give a better measure of cancer incidence.

**CONCLUSION**

Traditional SMRs are influenced by the distribution of age and active/inactive person-time of the study cohort. The present study shows that the proportion of person-time from individuals above 65 years may vary considerably in various cohorts. The proportion of active person-time, an important part of the HWE, may also show great variations across various populations. The results concerning the HWE are compatible with recent reports that the HWE is composed of the healthy hire effect and the HWSE, and that duration of follow-up do not affect the HWE. To reduce the influence from various random conditions associated with the study design and the age of the plants, etc. it is proposed to compute an SMRptw which is the sum of the stratum specific O/E-ratios weighted by the person-time distribution of age, employment status, and calendar time.

**REFERENCES**


