Measures of disease frequency

Counts

Numbers of events are the simplest way of describing disease frequency. The overall numbers of events or people with a condition are useful for planning health service provision. However, the use of counts alone may be misleading for comparisons. For example, Rachel Carson in her environmental polemic *Silent Spring* drew attention to the rise in number of deaths from leukaemia in the United States between 1950 and 1960. During this decade the number of deaths per year rose from 8845 to 12,725. This change could largely be explained by the increased population size and changed population structure between the two dates.¹

Therefore when comparisons between different populations are being made, it is usual to control for population size and structure to ensure that any difference observed is not simply due to these factors.

Crude rates

The first precaution is to express numbers of events as a crude rate, i.e. the number of events in a population divided by the number of people in that population (usually per 1000). Use of crude rates allows comparisons to be made accounting for differences in population size, but does not take account of differences in age structure.

Age-specific rates

To allow for variation in event or death rate by age, rates can be calculated for each age group. These age-specific rates are defined as the number of events or deaths in the age group per 1000 population in the same age group.

Bland gives an example.² In 1901 and 1981, the crude mortality rates were similar for adult males in England and Wales (15.7 deaths per 1000 in 1901 and 14.8 deaths per 1000 in 1981). However, the age-specific rates show that mortality rates were higher for every age band in 1901 than 1981 (Fig. 1); the similarity in overall crude mortality rates reflects the much older population in 1981 (Fig. 2).

Tables of age-specific rates provide the most complete information about a population, but the large number of figures require much effort to interpret, and can be overwhelming (see Box 1). To produce a simple summary of the death or disease experience of a population, accounting for different population structures, it is usual to calculate a standardised rate or ratio.
Box 1: 
Different methods of standardisation compared using East of England local authority data

From the age-specific mortality rates shown in Table i, we have calculated the SMR and directly standardised mortality rate for selected local authorities in the East of England using different standard populations, to determine the effect of choice of standard population. The rates, ratios and their rankings are compared in Table ii.

Interpretation

The first three columns in Table ii show the directly standardised mortality rates, as calculated using the Segi (“World”; Ref. 2) and Scandinavian (“European”; Ref. 2) standards, and England & Wales 1999 populations3 as the standard populations. The World and European standard populations are hypothetical populations and both are much younger than the England & Wales population. These hypothetical populations tend to place greater emphasis on deaths occurring at younger ages. For example, Norwich has a relatively high mortality in middle age and low mortality at older ages. It therefore has a low rank when standardised to World population, high rank when standardised to the England & Wales population and intermediate when standardised to the European population. Neighbouring Broadland has the opposite pattern and a possible explanation is that older

Table i
Age-specific mortality rates for local authorities in East of England, 1999 (from Ref.1)

<table>
<thead>
<tr>
<th>Local authority</th>
<th>Age-specific mortality rate per 1000 Males</th>
<th>0-4</th>
<th>5-9</th>
<th>10-14</th>
<th>15-19</th>
<th>20-24</th>
<th>25-29</th>
<th>30-34</th>
<th>35-39</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedford</td>
<td></td>
<td>1.91</td>
<td>0.21</td>
<td>0.43</td>
<td>1.26</td>
<td>0.51</td>
<td>0.33</td>
<td>0.79</td>
<td>1.11</td>
</tr>
<tr>
<td>Broadland</td>
<td></td>
<td>0.30</td>
<td>0.00</td>
<td>0.26</td>
<td>0.58</td>
<td>0.92</td>
<td>0.00</td>
<td>0.24</td>
<td>1.37</td>
</tr>
<tr>
<td>Luton</td>
<td></td>
<td>1.88</td>
<td>0.13</td>
<td>0.30</td>
<td>0.16</td>
<td>0.84</td>
<td>0.96</td>
<td>1.02</td>
<td>0.62</td>
</tr>
<tr>
<td>Norwich</td>
<td></td>
<td>1.49</td>
<td>0.00</td>
<td>0.00</td>
<td>0.53</td>
<td>0.36</td>
<td>1.83</td>
<td>1.67</td>
<td>0.82</td>
</tr>
<tr>
<td>South Cambs.</td>
<td></td>
<td>0.53</td>
<td>0.00</td>
<td>0.00</td>
<td>0.70</td>
<td>1.06</td>
<td>0.19</td>
<td>0.19</td>
<td>1.61</td>
</tr>
<tr>
<td>Stevenage</td>
<td></td>
<td>1.06</td>
<td>0.00</td>
<td>0.32</td>
<td>0.00</td>
<td>1.61</td>
<td>0.28</td>
<td>0.28</td>
<td>1.72</td>
</tr>
<tr>
<td>Watford</td>
<td></td>
<td>1.80</td>
<td>0.00</td>
<td>0.37</td>
<td>0.00</td>
<td>1.35</td>
<td>0.23</td>
<td>0.98</td>
<td>0.52</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Local authority</th>
<th>Age-specific mortality rate per 1000 Females</th>
<th>0-4</th>
<th>5-9</th>
<th>10-14</th>
<th>15-19</th>
<th>20-24</th>
<th>25-29</th>
<th>30-34</th>
<th>35-39</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedford</td>
<td></td>
<td>0.46</td>
<td>0.00</td>
<td>0.23</td>
<td>0.22</td>
<td>0.00</td>
<td>0.76</td>
<td>0.36</td>
<td>0.77</td>
</tr>
<tr>
<td>Broadland</td>
<td></td>
<td>1.27</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.48</td>
<td>0.00</td>
<td>0.21</td>
<td>0.00</td>
</tr>
<tr>
<td>Luton</td>
<td></td>
<td>1.75</td>
<td>0.00</td>
<td>0.14</td>
<td>0.00</td>
<td>0.22</td>
<td>0.38</td>
<td>0.36</td>
<td>1.09</td>
</tr>
<tr>
<td>Norwich</td>
<td></td>
<td>1.50</td>
<td>0.00</td>
<td>0.00</td>
<td>0.71</td>
<td>0.52</td>
<td>0.26</td>
<td>0.83</td>
<td>0.22</td>
</tr>
<tr>
<td>South Cambs.</td>
<td></td>
<td>0.78</td>
<td>0.24</td>
<td>0.00</td>
<td>0.25</td>
<td>0.77</td>
<td>0.22</td>
<td>0.95</td>
<td>0.75</td>
</tr>
<tr>
<td>Stevenage</td>
<td></td>
<td>1.16</td>
<td>0.33</td>
<td>0.00</td>
<td>0.00</td>
<td>0.50</td>
<td>0.29</td>
<td>0.30</td>
<td>1.18</td>
</tr>
<tr>
<td>Watford</td>
<td></td>
<td>1.09</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.39</td>
<td>0.24</td>
<td>0.00</td>
<td>1.14</td>
</tr>
</tbody>
</table>

Table ii
Directly standardised mortality rates and indirectly standardised ratios (SMRs) for local authorities in East of England, 1999 (from Refs 1–4).

<table>
<thead>
<tr>
<th>Local authority</th>
<th>Directly standardised mortality rates, by standard population</th>
<th>World standard (Segi)</th>
<th>European standard (Scandinavian)</th>
<th>England &amp; Wales 1999</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DSR per 100,000</td>
<td>Rank*</td>
<td>DSR per 100,000</td>
<td>Rank*</td>
</tr>
<tr>
<td>Bedford</td>
<td>439</td>
<td>34</td>
<td>698</td>
<td>30</td>
</tr>
<tr>
<td>Broadland</td>
<td>402</td>
<td>13</td>
<td>666</td>
<td>19</td>
</tr>
<tr>
<td>Luton</td>
<td>488</td>
<td>46</td>
<td>762</td>
<td>46</td>
</tr>
<tr>
<td>Norwich</td>
<td>448</td>
<td>37</td>
<td>687</td>
<td>27</td>
</tr>
<tr>
<td>South Cambs.</td>
<td>359</td>
<td>1</td>
<td>578</td>
<td>2</td>
</tr>
<tr>
<td>Stevenage</td>
<td>462</td>
<td>41</td>
<td>738</td>
<td>41</td>
</tr>
<tr>
<td>Watford</td>
<td>502</td>
<td>47</td>
<td>829</td>
<td>47</td>
</tr>
</tbody>
</table>

* Rank based on all East of England local authorities (48 in total): 1 represents best mortality experience, 48 represents worst.

Sources: World and European standard populations, Ref.2; England & Wales population, Ref. 3; Afghanistan and Japan populations, Ref. 4.
people when they become frail move into care in Broadland before dying. For other areas with a consistently high mortality at all ages (e.g. Watford) or low mortality at all ages (e.g. South Cambridgeshire) the choice of standard population has very little effect on ranking.

For indirect standardisation, populations with very different mortality experiences have been chosen for comparison. Afghanistan has a high mortality at all ages but especially among young children. This tends to increase the expected number of deaths in populations with a large proportion of children such as Luton, Stevenage and Bedford. The Afghanistan mortality rates are sufficiently different, that the ranking of SMRs calculated in this way is poorly correlated with other methods of standardisation.

Japan has a consistently low mortality at all ages, better than the best local authority population in the East of England. However the rates are not consistently lower at any particular age so there is little difference in the ranking compared with England & Wales rates.

The highest rank correlation occurs between the DSRs and SMRs standardised to the England & Wales 1999 population and rates ($R = 0.98$). It is likely that choice of standard population is more important in ensuring comparability than whether direct or indirect standardisation is chosen.

### References

Standardisation

There are two basic methods of standardisation – direct and indirect. Each method has advantages and disadvantages. Both methods use a study population (the population of interest, e.g. local population) and a standard population (e.g. national population or European standard population – see ‘The choice of standard population’ below) to generate weighted averages of age-specific rates, but they use different weighting schemes. Either method can produce rates or ratios, as shown in Table 1.

Indirect standardisation

The indirectly standardised mortality rate is the mortality rate expected in the study population if the age-specific rates of a standard population had applied. The standardised mortality ratio (SMR) is the ratio of the number of deaths observed in the study population to the number that would be expected if the age-specific rates of a standard population had applied. This ratio is usually multiplied by 100. An SMR of 120 means that the mortality in the study population is 20% greater than would have been expected, had the age-specific rates in the standard population applied.

\[ SMR = \frac{\text{Observed deaths in study population}}{\text{Expected deaths in study population}} \times 100 \]

or

\[ SMR = \frac{d}{\sum_{i=1}^{k} n_i R_i} \]

where \(d\) is the number of deaths in the study population, \(n_i\) is the number of people in the \(i^{th}\) group of the study population, \(R_i\) is the crude death rate in the \(i^{th}\) group of the standard population and \(k\) is the number of groups.

Table 1.
Summary measures of disease frequency and mortality.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Direct</th>
<th>Indirect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate</td>
<td>Directly standardised rate (DSR)</td>
<td>Indirectly standardised rate (ISR)</td>
</tr>
<tr>
<td>Ratio</td>
<td>Directly standardised ratio e.g. Comparative mortality figure (CMF)</td>
<td>Indirectly standardised ratio e.g. Standardised mortality ratio (SMR)</td>
</tr>
<tr>
<td>Note:</td>
<td>(DSR = \text{CMF} \times \text{crude rate in standard population})</td>
<td>(ISR = \text{SMR} \times \text{crude rate in standard population})</td>
</tr>
</tbody>
</table>

When SMRs are calculated for several populations with significantly differing population structures, any difference may be due to the different population structures rather than different age-specific rates. For this reason they cannot be compared directly with each other, only with the standard population. In spite of this inherent bias, SMRs and standardised limiting long-term illness ratios (similar to SMR, but using prevalence of limiting long-term illness instead of number of deaths) are used in resource allocation and target setting in the United Kingdom, as well as in the Compendium of Clinical and Health Indicators (formerly the Public Health Common Data Set).
Box 2: Calculating confidence intervals

**Confidence interval for standardised mortality ratio (SMR)**
For observed number of deaths greater than 25 or so, an approximate 95% confidence interval is given by

\[
\text{SMR} - 1.96 \times \text{SE(SMR)} \text{ to } \text{SMR} + 1.96 \times \text{SE(SMR)}
\]

where \( \text{SE(SMR)} = \sqrt{\frac{O}{E}} \)

\(O\) is the observed number of deaths in the study population and \(E\) is the expected number of deaths (see ‘Indirect standardisation’ in main text for calculation of expected number of deaths).\(^1\)

**Confidence interval for comparative mortality figure (CMF)**
Calculation of confidence intervals for the CMF is more cumbersome. The 95% confidence interval is given by

\[
\text{CMF} \times \exp \left[ 1.96 \times \text{SE(CMF)} \right] \text{ to } \text{CMF} \times \exp \left[ -1.96 \times \text{SE(CMF)} \right]
\]

where \( \text{SE(CMF)} = \sqrt{\sum_{i=1}^{k} \frac{N_{i}^2 \cdot d_{i}}{n_{i}^2 \cdot N_{i}^2}} \)

and \(N_{i}\) is the number of people in the \(i^{th}\) group of the standard population, \(d_{i}\) is the number of deaths in the \(i^{th}\) group of the study population, \(n_{i}\) is the number of people in the \(i^{th}\) group of the study population, \(N\) is the total number of people in the standard population, \(k\) is the number of groups. This estimate uses a method described by Breslow and Day\(^2\) to calculate the standard error, and assumes that the numbers of events in each age group follow a Poisson distribution. The estimate deteriorates for small rates.\(^3\)

**Confidence interval for directly standardised rate (DSR)**
The 95% confidence interval for a DSR is approximated by

\[
\text{DSR} - 1.96 \times \text{SE(DSR)} \text{ to } \text{DSR} + 1.96 \times \text{SE(DSR)}
\]

where \( \text{SE(DSR)} = \sqrt{\frac{\sum_{i=1}^{k} N_{i}^2 \cdot r_{i}}{N^2 \cdot n_{i}}} \)

and \(N_{i}\) is the number of people in the \(i^{th}\) group of the standard population, \(r_{i}\) is the death rate in the \(i^{th}\) group of the study population, \(N\) is the total number of people in the standard population, \(n_{i}\) is the number of people in the \(i^{th}\) group of the study population and \(k\) is the number of groups. This estimate uses a method described by Breslow and Day\(^3\) to calculate the standard error, and assumes that the numbers of events in each age group follow a Poisson distribution. The estimate deteriorates for small rates.\(^3\)

**References**
The choice of standard population

The choice of standard population is essentially arbitrary. It may depend on the point being made: the Registrar General William Farr, in his Sixteenth annual report of 1856, chose healthiest areas as the standard, implying that this was the standard to which other areas should aspire.9

Many standard populations currently in use were chosen several decades ago, when populations tended to be much younger. Examples include the European standard population and the 1940 United States population.10 Use of these younger populations tends to place greater emphasis on deaths occurring at younger ages.

Hypothetical populations are often chosen for direct standardisation. The European standard population is widely used for international comparisons and time trends. The European standard population approximates to the European population in about 1970. This standard population is similar for males and females, so standardised rates may be compared within each sex, and between sexes.

Changes in the standard population to reflect the age structure of modern populations may change the relative ranking of different groups. For example, when the United States moved from using the 1940 to the 2000 standard population, the CMF for ischaemic heart disease for blacks compared with whites changed from 1.07 to 0.96 (Ref. 8).

Whether indirect or direct methods are used, there are some general principles for choosing a standard population:

1. Always choose a standard population similar to the study populations. For example, if the units of interest are all the local authorities in England, then use the national population; if the units of interest are two wards in a local authority, then the local authority population or the average of the two wards may be appropriate.

2. Do not standardise the study populations to different standard populations.

3. Examine the age structures of the study populations as well as the standard population. If the study age structures differ widely from each other and from the proposed standard, then indirect standardisation will generate misleading results.

Which method and what data?
The method of choice will depend on the data available. If age-specific mortality rates for the study population are not available, then indirect methods are the only option. Table 2 presents the options for direct and indirect methods.
### Table 2.
Data requirements and features of direct and indirect methods of standardisation.

<table>
<thead>
<tr>
<th>Question?</th>
<th>Direct (DSR and CMF)</th>
<th>Indirect (SMR)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data requirements</strong></td>
<td>Study population: Age-specific rates Study population: Age structure</td>
<td>Study population: Age structure Standard population: Age-specific rates</td>
</tr>
<tr>
<td><strong>Interpretation</strong></td>
<td>DSR is a summary figure which has no intrinsic meaning. If there are two populations A and B with DSRs of 100 and 200 respectively, then B has twice the rate of A. The CMF is a summary ratio where a figure over 1 (or 100%) indicates higher than expected mortality.</td>
<td>The SMR is a summary ratio where a figure over 1 (or 100%) indicates higher than expected mortality.</td>
</tr>
<tr>
<td><strong>Precision</strong></td>
<td>Less precise when there are small numbers of events in a given age band in study population.</td>
<td>More precise so narrower confidence intervals.</td>
</tr>
<tr>
<td><strong>I want to compare myself with my neighbour(s)</strong></td>
<td>Direct methods are suitable as long as the same standard population is used.</td>
<td>If an external standard is used, the calculated standardised rates or ratios will not be strictly comparable. This will be important if the age structures of local populations differ.</td>
</tr>
<tr>
<td><strong>Which standard population should I use?</strong></td>
<td>General advice is to use a standard that is close to your population of interest; e.g. the European standard population is now relatively young in comparison with UK – may be better to use national or regional population.</td>
<td></td>
</tr>
<tr>
<td><strong>How do I work out confidence intervals?</strong></td>
<td>See Box 2</td>
<td></td>
</tr>
<tr>
<td><strong>I want to rank rates – which method…?</strong></td>
<td>Direct methods are better because each estimate is adjusted to the same population. If a ratio is required, the CMF is preferable to the SMR because CMFs can be compared with each other.</td>
<td>Indirect methods do not allow strict comparisons between rates – they measure performance relative to the standard, and ranking may mislead.</td>
</tr>
<tr>
<td><strong>What about small area methods…?</strong></td>
<td>If the data are available and the lack of precision is tolerable, DSRs are best – it may require aggregation of years and areas.</td>
<td>Usually the data are not available so SMRs are often used.</td>
</tr>
</tbody>
</table>

### Summary

1. Standardisation produces a simple summary statistic for the disease experience of a population. One should always standardise rates (i.e. adjust for age differences) when comparing event rates between two or more populations where age structures are likely to differ (as is usually the case).

2. Because standardised rates and ratios are summary measures, they may conceal variations in mortality experience between the age groups. To understand variation in standardised rates, it may be necessary to review the underlying age-specific rates.

3. Mortality rates for different areas using different methods of standardisation are usually closely correlated if the populations in question have population structures or age-specific rates similar to the standard population.

4. If populations have a very different age structure or different age-specific rates from the standard populations, the standardised rates will not be closely correlated.

5. To minimise bias, standard populations should be chosen to be as close as possible in structure and age-specific mortality rates to the populations being compared.

6. The European standard population, used for cancer registries, is a much younger population than currently found in Europe and therefore introduces bias, placing greater emphasis on events occurring at younger ages.

7. If local age-specific mortality data are available, either method can be used but directly standardised rates and ratios allow legitimate between-area comparisons.

8. If local age-specific mortality data are not available, indirect standardisation is the only option.

9. If there are small numbers of events, try to aggregate time periods, age groups or areas. Indirect methods give more precise estimates but are not strictly comparable unless population structures are similar.

10. If you must rank, rank directly standardised rates or ratios.

11. Consider other methods of summarising population mortality experience that use only age-specific rates and do not depend on standard population structure, such as life expectancy.
Serving your needs

As a result of the Steering Group/ Stakeholder Consultation on 10 March 2005, erpho’s work priorities henceforth have been defined as:

1. Small area data development.
2. Giving policy steers on when/how (not) to collect local data.
3. Lobbying locally to address inequity of intelligence resources.
4. Development of groups to complement the Steering Group (e.g. Heads of Public Health Intelligence).
5. Systematic telephone (and email) enquiry service.
6. Balancing ad hoc, reactive work with planned, proactive work.
7. Named person within erpho to liaise with counties/ Strategic Health Authorities.
8. Working to more systematic frameworks and methodological approaches.
9. Defining boundaries of action and responsibility between PCTs, Shared Services and erpho.
10. Website topic pages on specific subjects.
11. Mapping and scoping public health capacity and skills throughout the region.
12. Professional development of the public health intelligence workforce.

Find the full document on our website at Quick Link 12333.

References


INpho

Census 2001: Children
Census 2001: Education and employment
Overview of health inequalities in the East of England
Inequalities in all-cause mortality in the East of England
Inequalities in CHD in the East of England

INphoRM

Health inequality: alternative measures
Equity audits – a practical guide

Series on Risks and Determinants

Healthy eating in the East of England
Alcohol misuse in the East of England
Sexual activity in the East of England