Patama Vapattanawong & Pramote Prasartkul

Introduction
Mortality is a demographic indicator not only of health outcomes in various population groups, but also of a country’s state of development. Death data are fundamentally crucial for identifying health problems and for monitoring health programmes. They are also used to measure and compare mortality rates in subpopulations within the same country or across different countries. Accordingly, the quality of mortality data is the key to reliable indicators and should therefore be carefully assessed before the data are used.

In Thailand, the most important source of mortality data is the vital registration system, which was initiated in 1916. Civil registration law mandates that every vital event (births and deaths) be registered at the offices of the district or municipality registrars, which are under the interior ministry. According to this law, births must be registered within 15 days of delivery, while deaths must be registered within 24 hours. For deaths, the registration process is not straightforward because two steps are involved: (i) notifying an authorized person about the death, and (ii) validating and registering the death. After both registration steps are completed, the name of the deceased is deleted from the household roster and the details concerning the death are entered into the registration system. Before the end of each year the interior ministry counts the births, deaths, and total population and makes the figures publicly available.

In addition to the vital registration system, Thailand has a second important source of population data: the population census, which is conducted every 10 years. Theoretically, these two sources of demographic data should provide very similar data in Thailand’s civil registration system. Its objectives were to measure the under-registration of vital events. Thailand’s most recent SPC was conducted in 2005–2006.

Although estimating under-registration is a main objective of each SPC, the methods used in the surveys have not been the same. The differences in such methods are described in detail elsewhere. In brief, the estimation methods used in the first three SPCs (1964–1966, 1974–1976, 1984–1986) consisted of a dual records procedure, while those used in the 1995–1996 and 2005–2006 SPCs were based directly on survey questions, whether vital events were registered or not. According to the SPC results, during the 1960s and 1970s approximately 60% of the deaths in the Thai population were registered.

The very large increase in vital event registration observed in the 1995–1996 SPC relative to previous SPCs raised doubts, and the completeness of death registration in Thailand remains a controversial issue requiring further scrutiny.

This study aimed to investigate the quality of mortality data in Thailand’s civil registration system. Its objectives were to estimate the magnitude of under-registration of deaths and its associated age and sex patterns in 2005–2006 by cross-matching data from two sources.

Methods
Data sources
The two sources of data used in this study were the SPC 2005–2006 and vital registration records. The SPC 2005–2006 was a national, longitudinal, multi-round household survey with a large sample: 2050 sample rural villages/urban blocks and 82 000 households. This survey was designed to determine changes...
in household composition. To overcome the errors generated by the respondents’ limited ability to correctly recall events that occurred over an extended period, each sample household was visited five times within a year, or about once every three months. The survey period began in July 2005 and ended in July 2006.

All households in the sample rural villages and urban blocks were visited by the enumerators to obtain information on the usual household residents. In the initial interview, all persons residing in each household were listed and their basic demographic characteristics were recorded. A total of 327,735 individuals were counted in this first interview. As an additional check, to obtain as complete a coverage of births and deaths as possible, the enumerator also asked in every household: (i) if any women were known to be pregnant, and, if so, the number of months of pregnancy; and (ii) if any individuals were known to be seriously ill. For the second and subsequent interviews, the enumerator determined and recorded any changes in household composition that had taken place since the previous visit on account of births, deaths or movements in and out. If such changes had taken place, the date when they occurred was recorded. An additional question was asked about any babies born or deceased since the enumerator’s last visit.

To obtain vital registration data, individual death records for 2005–2006 from the interior ministry were used. Each record was in electronic format and had a 13-digit unique personal identifier (PID). Death records from before and after the SPC survey period were excluded from the study.

The individual records of the SPC had been linked to death records from vital registration using a 13-digit unique personal identifier (PID). Death records from before and after the SPC survey period were excluded from the study. The individual records of the SPC had been linked to death records from vital registration using a 13-digit unique personal identifier (PID). Death records from before and after the SPC survey period were excluded from the study.

The 2 × 2 contingency table (Table 1) containing the results of cross-matching the deaths from the SPC with those from vital registration system, Thailand, 2005–2006 is calculated by constructing a 2 × 2 contingency table containing the results of cross-matching the deaths from the SPC with those from vital registration system. The formula of \( p_1 \) gives the estimated completeness of death registration. Thus, the under-registration \( D_u \) could be estimated from:

\[
D_u = N - \left( p_1 \times N \right)
\]

Table 1. A 2 × 2 contingency table showing case-by-case cross-matching of deaths from the Survey of Population Change (SPC) and from the vital registration system, Thailand, 2005–2006

<table>
<thead>
<tr>
<th>Death in SPC</th>
<th>Vital registration</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>( M )</td>
<td>( NR )</td>
</tr>
<tr>
<td>No</td>
<td>( NS )</td>
<td>( MB )</td>
</tr>
<tr>
<td>Total</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

\( M \), matched (i.e. in both systems); \( MB \), missing from both systems; \( N \), number of events (estimated total); \( NR \), not in registry (i.e. in survey only); \( NS \), not surveyed (i.e. in vital registry only).

\[ Var(N) = N \frac{pq}{p_1 \cdot p_2} \]

where:

\[ p_1 = \frac{M + NS}{N} \]

\[ p_2 = \frac{M + NR}{N} \]

and

\[ p_1 + q_1 = p_2 + p_3 = 1 \]

The 95% confidence intervals (CIs) of \( N \) were calculated using the formula:

\[ 95\% \text{ CI} = N \pm 1.96 \sqrt{Var(N)} \]

The formula of \( p_1 \) gives the estimated completeness of death registration. Thus, the under-registration \( D_u \) could be estimated from:
Under-registration of deaths in Thailand

and the variance of \( q_i \) was calculated as:\(^7\)

\[
\text{Var}(q_i) = \frac{(1 - p_i)p_i}{N(1 - p_i)}
\]

**Results**

**Cross-matching of deaths**

The diagram of the cross-matching process between deaths listed in the SPC and vital registration data is shown in Fig. 1. Of 290,625 individual records from the SPC, 1882 were deaths that could be matched to vital registration data; 179 were deaths that could not be matched; 556 were people listed as survivors in the SPC who could be matched to vital registration death records; and 288,008 were people listed as survivors in the SPC who could not be matched to vital registration death records (Table 2).

As shown in Table 2, the SPC listed a total of 2061 deaths as having occurred during 2005–2006. Over that period, 2438 deaths were included in the vital event registry. The figures in this table also reflect counting errors, since 556 people listed in the death registry were reported as being alive during the survey. The table also shows the magnitude of the under-registration: 179 people reported to be dead during the survey were not registered in the vital event registry.

**Estimation of under-registration of deaths**

A second 2 × 2 contingency table (Table 3) was created with the death data in Table 2 to determine how many people had been missed by both systems (the cell with ‘???’). The estimated total number of deaths occurring during the 2005–2006 survey and its 95% CI were:

\[
N = \frac{(1882 + 556) \times (1882 + 179)}{2670} = 2061
\]

where: the result was rounded to the nearest integer. In addition, the estimated under-registration of deaths, and its 95% CI could be calculated as:

\[
\text{Du} = \left( \frac{1}{100} \times (1882 + 556) \right) - 1882
\]

To estimate the under-registration of deaths by age and sex, we repeated the above calculation using data classified by age and sex instead. Table 4 presents the age- and sex-specific numbers of deaths in 2005–2006 as recorded by the SPC and vital registration records, and Table 5 shows the estimated percentage of under-registration of deaths. In 2005–2006 under-registration was 8.69% (95% CI: 8.65–8.72) for both sexes combined and, for all ages overall, it was 9.00% (95% CI: 8.95–9.05) for males and 8.36% (95% CI: 8.31–8.41) for females.

When we stratified by broad age groups, the highest under-registration was found among children 1–4 years old: 54.55% (95% CI: 54.25–54.84) in males and 71.43% (95% CI: 71.11–71.75) in females. Under-registration of male and female deaths decreased as age increased, with the sole exception of deaths among females less than 1 year.
of age, which showed much lower under-registration than deaths among male infants (8.33%; 95% CI: 7.87–8.80 versus 34.78%, 95% CI: 34.48–35.08, respectively). Furthermore, deaths among older males and females showed less than 10% under-registration.

**Discussion**

The aim of this study was to estimate the magnitude of under-registration of deaths by making use of two valuable sources of data that already existed in Thailand. Although the method used herein has been known for more than half a century, it is still useful, particularly in countries having several sources of mortality data, such as Egypt and the Philippines. This method is frequently used in studies and is often referred to in epidemiology and ecology as capture–recapture.

In Thailand, the dual records procedure or a similar method based on cross-matching of data from two different sources has been practised since the mid-1960s, and in that early period it was applied in a manner very similar to the one shown in this study in the sense that data from independent sources – in our case the survey and the vital event registration system – were cross-matched. However, some practical differences do exist. Whereas decades ago the deaths reported in the survey were manually cross-matched with the ones included in the registration system on the basis of name and other characteristics common to both systems, such as sex, age of deceased, place of death and place of residence, in this study manual cross-matching was replaced by computerized cross-matching using a 13-digit unique PID. In addition, to ensure correct cross-matching, we double checked the age and sex of each pair of cross-matching cases. Thus, this study was stronger than earlier studies. Moreover, the dual records procedure made it possible to generate estimates for population subgroups. Thus it is possible to estimate under-registration of deaths at a subnational level or among subgroups of the population if the data are available.

The estimated under-registration of deaths for both sexes combined and for males and females separately was 8.69%, 9.00% and 8.36%, respectively. These figures are approximately 1.8, 1.7 and 1.9 times higher than the percentages estimated directly by Thailand’s National Statistical Office from the questions in the SPC.

Table 4. Number of deaths in 2005–2006, by age and sex, obtained by cross-matching data from the Survey of Population Change (SPC) and the vital registration system

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Male deaths</th>
<th></th>
<th>Female deaths</th>
<th></th>
<th>Overall deaths</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;1 year</td>
<td>1–4 years</td>
<td>5–14 years</td>
<td>15–59 years</td>
<td>60–74 years</td>
<td>75+ years</td>
</tr>
<tr>
<td>Male deaths</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In both systems (M)</td>
<td>15</td>
<td>5</td>
<td>11</td>
<td>289</td>
<td>222</td>
<td>408</td>
</tr>
<tr>
<td>In SPC only (NR)</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>29</td>
<td>18</td>
<td>24</td>
</tr>
<tr>
<td>In vital registry only (NS)</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>112</td>
<td>74</td>
<td>89</td>
</tr>
<tr>
<td>Estimated total no. (N)</td>
<td>28</td>
<td>20</td>
<td>27</td>
<td>441</td>
<td>320</td>
<td>526</td>
</tr>
<tr>
<td>Female deaths</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In both systems (M)</td>
<td>11</td>
<td>2</td>
<td>4</td>
<td>138</td>
<td>205</td>
<td>572</td>
</tr>
<tr>
<td>In SPC only (NR)</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>24</td>
<td>17</td>
<td>37</td>
</tr>
<tr>
<td>In vital registry only (NS)</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>75</td>
<td>73</td>
<td>117</td>
</tr>
<tr>
<td>Estimated total no. (N)</td>
<td>16</td>
<td>11</td>
<td>5</td>
<td>250</td>
<td>301</td>
<td>734</td>
</tr>
</tbody>
</table>

CI, confidence interval.

These figures are not the sum of the values in each age group.

Table 5. Estimated percentage of under-registration of deaths, by age and sex, Thailand, 2005–2006

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Male deaths</th>
<th>Female deaths</th>
<th>Overall deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>95% CI</td>
<td>%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1</td>
<td>34.78</td>
<td>34.48–35.08</td>
<td>8.33</td>
</tr>
<tr>
<td>1–4</td>
<td>54.55</td>
<td>54.25–54.84</td>
<td>71.43</td>
</tr>
<tr>
<td>5–14</td>
<td>45.00</td>
<td>44.72–45.28</td>
<td>20.00</td>
</tr>
<tr>
<td>60–74</td>
<td>7.50</td>
<td>7.39–7.61</td>
<td>7.66</td>
</tr>
<tr>
<td>75+</td>
<td>5.56</td>
<td>5.47–5.64</td>
<td>6.08</td>
</tr>
<tr>
<td>Overall</td>
<td>9.00</td>
<td>8.95–9.05</td>
<td>8.36</td>
</tr>
</tbody>
</table>

CI, confidence interval.
our study captured more deaths, particularly 556 deaths that were in the vital event registration system but that had not been counted in the survey. Two factors could account for the substantial under-reporting of deaths in the SPC (11.3% or more) for all deaths in both systems: (i) the assumption that the population interest was closed and (ii) the quality of the counting of deaths in the survey. The life/death status of the 556 people who were in the death registry but not in the SPC had been carefully reassessed in the fifth (last) visit paid to the household during the survey period (data not shown). Of these people, 78.8% were found to be regular residents, 17.4% had moved away and 3.8% were temporarily absent or transient residents. Thus, the quality of the counting of deaths may have been the main reason that deaths were under-reported in the SPC, followed by the mistaken assumption that the population surveyed was a closed one. Similar findings have not been reported in other countries.\(^1\)\(^2\)\(^3\)

Our figures may be an underestimation of the actual rate of under-reporting. One reason is that not all SPC 2005–2006 sample data could be analysed because some information (11.3%) was incomplete; either the PID was missing (10.3%) or it had fewer than 13 digits (0.2%) or was duplicated (0.8%). Another reason is that only one unique identifier was used as a linkage key. If some identifiers were incorrectly recorded during the interview or when entered into the system, the result of the cross-matching could have been affected. Thus, this study arguably provides a minimum estimate of the rate of under-registration of deaths in Thailand during 2005–2006. Its findings lend support to the notion that the use of different estimation methods accounts for the large improvement in the completeness of death registration observed after the mid-1990s.

Conclusion

In Thailand, overall under-registration of deaths in 2005–2006 was around 8.69%, but the percentage varied by age and sex. The age- and sex-specific estimates of under-registration from this study can be used as correction factors to improve the accuracy of the death statistics in the country’s registration system. Several studies have calculated the magnitude of under-registration overall, but few have made age- and sex-specific estimates. By using mortality data from two important sources – the SPC 2005–2006 and the vital registration system – this study provides some degree of assurance that the results are reliable. Age- and sex-specific numbers of deaths, after adjustment by the correction factors generated in this study, can be used as numerators to make various mortality estimates. The newly-adjusted age- and sex-specific death rates can even be used, for example, to construct life tables of Thailand’s general population and subpopulation groups.

Acknowledgements

We thank Thailand’s National Statistical Office and the Ministry of Interior for allowing us to use the Survey of Population Change and the mortality data in the vital registration system for this study.

Competing interests: None declared.

Method

The goal of this study was to evaluate the completeness of death registration, using the survey of population change, and to provide age- and sex-specific estimates of under-registration. The cross-matching of death and age-specific numbers can be used as correction factors to improve the accuracy of the death statistics in the country’s registration system. Several studies have calculated the magnitude of under-registration overall, but few have made age- and sex-specific estimates. By using mortality data from two important sources – the SPC 2005–2006 and the vital registration system – this study provides some degree of assurance that the results are reliable. Age- and sex-specific numbers of deaths, after adjustment by the correction factors generated in this study, can be used as numerators to make various mortality estimates. The newly-adjusted age- and sex-specific death rates can even be used, for example, to construct life tables of Thailand’s general population and subpopulation groups.

Our figures may be an underestimate of the actual rate of under-reporting. One reason is that not all SPC 2005–2006 sample data could be analysed because some information (11.3%) was incomplete; either the PID was missing (10.3%) or it had fewer than 13 digits (0.2%) or was duplicated (0.8%). Another reason is that only one unique identifier was used as a linkage key. If some identifiers were incorrectly recorded during the interview or when entered into the system, the results of the cross-matching could have been affected. Thus, this study arguably provides a minimum estimate of the rate of under-registration of deaths in Thailand during 2005–2006. Its findings lend support to the notion that the use of different estimation methods accounts for the large improvement in the completeness of death registration observed after the mid-1990s.

Conclusion

In Thailand, overall under-registration of deaths in 2005–2006 was around 8.69%, but the percentage varied by age and sex. The age- and sex-specific estimates of under-registration from this study can be used as correction factors to improve the accuracy of the death statistics in the country’s registration system. Several studies have calculated the magnitude of under-registration overall, but few have made age- and sex-specific estimates. By using mortality data from two important sources – the SPC 2005–2006 and the vital registration system – this study provides some degree of assurance that the results are reliable. Age- and sex-specific numbers of deaths, after adjustment by the correction factors generated in this study, can be used as numerators to make various mortality estimates. The newly-adjusted age- and sex-specific death rates can even be used, for example, to construct life tables of Thailand’s general population and subpopulation groups.

Acknowledgements

We thank Thailand’s National Statistical Office and the Ministry of Interior for allowing us to use the Survey of Population Change and the mortality data in the vital registration system for this study.

Competing interests: None declared.

Method

The goal of this study was to evaluate the completeness of death registration, using the survey of population change, and to provide age- and sex-specific estimates of under-registration. The cross-matching of death and age-specific numbers can be used as correction factors to improve the accuracy of the death statistics in the country’s registration system. Several studies have calculated the magnitude of under-registration overall, but few have made age- and sex-specific estimates. By using mortality data from two important sources – the SPC 2005–2006 and the vital registration system – this study provides some degree of assurance that the results are reliable. Age- and sex-specific numbers of deaths, after adjustment by the correction factors generated in this study, can be used as numerators to make various mortality estimates. The newly-adjusted age- and sex-specific death rates can even be used, for example, to construct life tables of Thailand’s general population and subpopulation groups.

Acknowledgements

We thank Thailand’s National Statistical Office and the Ministry of Interior for allowing us to use the Survey of Population Change and the mortality data in the vital registration system for this study.

Competing interests: None declared.
Subregistro de muertes en Tailandia en el periodo comprendido entre 2005 y 2006: resultados de la obtención de datos de pruebas cruzadas procedentes de dos fuentes

Objetivo Estimar la magnitud del subregistro de muertes, por edad y sexo, en Tailandia.

Métodos Los datos incluidos en este estudio proceden de dos fuentes: el Estudio tailandés de variaciones de población (EVP) de 2005–2006, un estudio de hogares con visitas múltiples consecutivas realizado en un periodo de 12 meses y los registros civiles de Tailandia. Las entradas del EVP para los decés de personas de todas las edades que han sido comparadas con las entradas de los registros civiles de Tailandia. Los principios de un sistema de double enregistrement que se pueden utilizar para ajustar los datos de mortalidad en Tailandia, por edad y sexo, y la fórmula de Chandrasekaran-Deming.

Resultados El subregistro global de muertes durante el periodo comprendido entre 2005 y 2006 fue del 9,00% (intervalo de confianza del 95%, IC: 8,95–9,05) para hombres y del 8,36% (IC del 95%: 8,31–8,41) para mujeres. Para ambos sexos, el subregistro de muertes disminuía con el tiempo. Para mujeres, el subregistro de muertes disminuyó a menos del 10% después de los 5 años. Para hombres, el subregistro de muertes disminuyó a menos del 10% después de los 4 años.

Conclusiones Estos resultados proporcionan factores de corrección que se pueden utilizar para ajustar los datos de mortalidad en Tailandia, por edad y sexo.