
J. P. Crofts,* R. Pebody,† A. Grant,‡ J. M. Watson,* I. Abubakar*

Departments of * Respiratory Diseases, † Immunisation and ‡ Statistics Modelling & Bioinformatics, Health Protection Agency Centre for Infections, London, UK

SUMMARY

OBJECTIVE: To obtain a more accurate estimate of tuberculosis (TB) case fatality within 12 months of starting treatment or notification among TB cases reported to the national surveillance system.
METHODS: Records of deaths for all TB cases reported to the national surveillance system were identified using linkage to two other sources of mortality information: the National Health Service (NHS) central register and death registrations from the Office for National Statistics (ONS). These data were compared to reports of deaths ascertained by national TB surveillance through treatment outcome monitoring. Capture-recapture methodology was used to estimate any remaining unascertained deaths for the final calculation of the case fatality rate (CFR).

RESULTS: In total, 1169 deaths (95%CI 1140–1224) were identified among 13 176 cases (CFR 8.9%, 95%CI 8.7–9.3%). Data linkage with the NHS central register and death registrations from the ONS identified a further 255 deaths not recorded by national TB surveillance. Capture-recapture estimated 61 (95%CI 32–116) unascertained deaths.

CONCLUSION: Mortality among TB cases is underestimated by national TB surveillance. Real-time data linkage between case reports and death registrations would allow a more accurate and timely estimate of TB mortality.

KEY WORDS: tuberculosis; mortality; surveillance; capture-recapture

AN ACCURATE ESTIMATE of mortality and its causes among tuberculosis (TB) cases is important to understand the burden and cost of this disease to the public and the nation. It is also important for the effective assessment of TB management and control, because death is the most commonly reported reason for TB treatment failure in England and Wales.1,2

National information on deaths among TB cases is currently collected through treatment outcome monitoring.3 This includes cases diagnosed post mortem and deaths that occur before reported completion within 12 months of the start of treatment or notification date. Although reporting to the Health Protection Agency (HPA) has gradually improved since the inception of monitoring in 2002, outcomes remain unknown for approximately 20% of cases.4 When a death is recorded, the date of death and the contribution of TB to that death are often unknown. In addition, deaths that occur after completion of treatment are not recorded.

The Office for National Statistics (ONS) collates information from death certificates on all deaths in England and Wales. Information is also available on deaths from the National Health Service (NHS) central register. Neither of these sources of information is directly linked to the national TB surveillance system run by the HPA. The aim of this study was first to use these additional sources of mortality information to obtain a better estimate of the total number of deaths among TB cases within 12 months of starting treatment or notification date, and second, to estimate case fatality and describe the characteristics of these patients. Another objective was to report all the deaths identified by the study, including those that occurred after 12 months of follow-up.

METHODS

All TB case reports to the national TB surveillance system in 2001 and 2002, including information on treatment outcome, were linked to two other sources of mortality information: the NHS central register and routine death registrations from the ONS. Details on the Enhanced TB Surveillance system and treatment outcome monitoring have been reported elsewhere.3,5

Correspondence to: I Abubakar, Department of Respiratory Diseases, Health Protection Agency Centre for Infections, 61 Colindale Avenue, London, NW9 5EQ, UK. Tel: (+44) 20 8327 7649. Fax: (+44) 20 8327 7868. e-mail: ibrahim.abubakar@hpa.org.uk

**Linkage with the NHS central register**

The NHS central register is an administrative database in England and Wales of all residents and new entrants registered with a general practitioner and all births since 1991. The system is updated on a daily basis with information from the NHS and weekly with information on new births or deaths from the ONS. A service is available from the ONS to trace persons on the NHS central register, including if they have died. For this study, the following criteria were used to automatically trace TB case reports for the years 2001 and 2002 in England and Wales to the NHS central register for any deaths recorded up to June 2006:

1. Exact matches on NHS number: first three characters of surname, first character of forename and any two out of the three parts of the date of birth.
2. If the NHS number is unavailable, automatic matching is on full surname, full forename and full date of birth.

**Linkage with death registrations**

It is a statutory requirement to report all births and deaths within England and Wales to the ONS. Routine death registration includes date of death and the underlying and contributing causes of death, which are classified according to the International Classification of Diseases coding system (ICD 10th revision). National TB case reports for the years 2001 and 2002 were also linked to a mortality data set of all death registrations occurring between 2000 and 2005 with any mention of TB on the death certificate corresponding to ICD-10 codes A15–A19, which cover infections due to *Mycobacterium tuberculosis* and *M. bovis*, excluding congenital TB, pneumoconiosis associated with TB, sequelae of TB and silicotuberculosis.

Data linkage was conducted at the Centre for Infections using in-house software that considers potential matches between two data sets and ranks them using a scoring system based on the completeness of the match across personal identifiers (soundex of forename, soundex of surname, date of birth, NHS number and post code). High scores above an agreed level indicating exact to near exact matches are considered automatically matched, and scores below this level are operator assessed down to a level where it is evident that no further matches are likely to be found.

**Definitions**

All TB cases known to have died were classified as TB or non-TB deaths based on the reported contribution of TB to the event. A TB death was defined as any mention of TB on the death certificate or if treatment outcome monitoring had recorded the cause of death as caused or contributed by TB. All remaining deaths were classed as non-TB deaths.

Deaths were also classified into whether the TB diagnosis was likely to have occurred before or after death. Death before diagnosis was defined as any case that had a date of death earlier than or equal to the date of notification, unless the date of diagnosis or treatment were before the date of death. All other deaths were placed in the ‘death after diagnosis’ category, including those with no known date of death.

**Capture-recapture analysis**

A three-list saturated log-linear model was used to undertake capture-recapture. Log-linear models are designed to incorporate any list (data source) dependence (inclusion in one sample has a direct causal effect on a deaths inclusion in another sample) that may exist. For a standard three-list model, all possible combinations of capture are arranged as an incomplete contingency table. Data captured by any combination of the three lists are assumed to be Poisson distributed. The number of unascertained deaths is extrapolated by assuming that the fitted model also applies to the missing cell of the complete contingency table. The main and interaction effects, estimated when fitting the log-linear model, are used in this calculation.

**Case-fatality rate**

The total number of deaths obtained from the data linkage and the capture-recapture analysis were then used to calculate the observed and estimated all-cause mortality case-fatality rates (CFR = number of deaths/number of reported cases), within 12 months of start of treatment (or notification if the treatment date was unknown).

**RESULTS**

A total of 13,391 cases of TB were reported between 2001 and 2002 in England and Wales. Treatment outcome reports revealed that 215 cases were non-tuberculous (often atypical mycobacteria), leaving a total of 13,176 cases identified as dying within 12 months of starting treatment or notification. The overlap between the three sources of mortality is illustrated in the Figure. The total number of TB cases identified as dying within 12 months of starting treatment or notification was 1108 (CFR 8.4%, 1108/13,176).

Treatment outcome monitoring recorded 119 deaths that were not matched to the other two mortality sources, while the linkage with the NHS central register and the ONS data set identified 255 deaths that were not recorded by TB surveillance. Among the latter, 43 cases were identified having died after completing treatment. A further 596 deaths were identified with dates of death beyond the 12 month cut-off from start of treatment or notification date up to June 2006 (CFR 12.9%, 1108 + 596/13,176).
Table 1 gives the CFR among the TB cases by all deaths and TB deaths, broken down by demographic, clinical and microbiological characteristics. Within each category, the CFR was higher among males, patients aged ≥65 years, UK born, white ethnic group, patients with a previous diagnosis of TB and patients with pulmonary disease. Table 2 describes the 1108 deaths by cause of death using information from the ONS data set and treatment outcome monitoring. In total, 628 deaths (57%) were classed as TB deaths. Among the 1108 deaths, 262 (24%) were classified as dying prior to diagnosis, of which 131 (50%) were classified as TB-associated deaths. Among the deaths prior to diagnosis, 171 (68%) were among patients aged ≥65 years.

In the capture-recapture analysis, one interaction term was non-significant with a small estimated coeffi-

Table 1  Case-fatality rate among reported tuberculosis cases within 12 months of start of treatment or notification date, England and Wales, 2001–2002

<table>
<thead>
<tr>
<th>Category#level</th>
<th>Deaths</th>
<th>Case-fatality rate, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All cases</td>
</tr>
<tr>
<td>All cases</td>
<td>1108</td>
<td>628</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>704</td>
<td>395</td>
</tr>
<tr>
<td>Female</td>
<td>404</td>
<td>233</td>
</tr>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>5–14</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>15–44</td>
<td>144</td>
<td>88</td>
</tr>
<tr>
<td>≥65</td>
<td>730</td>
<td>413</td>
</tr>
<tr>
<td>Country of birth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK-born</td>
<td>539</td>
<td>315</td>
</tr>
<tr>
<td>Non-UK born</td>
<td>330</td>
<td>183</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>682</td>
<td>395</td>
</tr>
<tr>
<td>Black Caribbean</td>
<td>23</td>
<td>12</td>
</tr>
<tr>
<td>Black African</td>
<td>79</td>
<td>47</td>
</tr>
<tr>
<td>South Asian†</td>
<td>208</td>
<td>120</td>
</tr>
<tr>
<td>Other</td>
<td>34</td>
<td>16</td>
</tr>
<tr>
<td>Previous diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>121</td>
<td>76</td>
</tr>
<tr>
<td>No</td>
<td>691</td>
<td>389</td>
</tr>
<tr>
<td>Site of disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary‡</td>
<td>834</td>
<td>495</td>
</tr>
<tr>
<td>Extra-pulmonary</td>
<td>258</td>
<td>125</td>
</tr>
</tbody>
</table>

* Excludes unknowns.
† South Asian = Indian, Pakistani or Bangladeshi.
‡ With or without extra-pulmonary disease.
TB = tuberculosis.
ficient, but it was not removed from the model. As it was thought the conditional independence implied by omitting the term was unlikely to be real, it was considered better to let the uncertainty about the size of any interaction influence the final confidence limits. The results of the saturated log-linear model gave an estimate of 61 (95% confidence interval [CI] 32–116) deaths unascertained by the three mortality sources. The final estimate of the total number of deaths among reported cases within 12 months of starting treatment or notification is 1169 (1108 + 61, 95% CI 1140–1224; CFR 8.9%, 95% CI 8.7–9.3%).

DISCUSSION

This study found that mortality among TB cases was underestimated in national TB surveillance. An additional 255 deaths were identified by data linkage between the case reports and the two other sources of mortality information. Capture-recapture methodology suggests a further 61 unascertained deaths. The CFR (all-cause mortality) within 12 months of treatment or notification may therefore be more than a third higher than that estimated from treatment outcome monitoring.

An elevated risk of mortality remains among TB patients after completing treatment, and a small proportion (4%) of the deaths in the 12-month follow-up period was among such patients. These deaths would not be recorded in treatment outcome monitoring, as follow-up of patients after treatment completion is not specified by national surveillance. This figure may be higher because treatment outcomes are not reported for all cases. As most of the additional deaths were identified for cases without an outcome report, improving outcome reporting remains a priority. A substantial number of deaths (all-cause mortality) also occurred after 12 months.

Our methods give results that are consistent with other findings that approximately half of the deaths among TB cases are directly attributable to TB. The 10% CFR (all-cause mortality) among culture-confirmed pulmonary cases is within, but at the high end of, the reported range of between 5% (Ireland) and 12% (Germany) in other countries in Western Europe in 2005. Past estimates for England and Wales put the all-cause mortality rate among notified cases of pulmonary TB at 12.5% in 1979 and 12.9% in 1983.

Like many other low TB incidence countries, the greatest burden of mortality among TB cases in England and Wales falls on the elderly population. The majority of these patients are UK-born and suffer from reactivating pulmonary disease. Many of these deaths may be unavoidable due to other co-morbidities, but other factors that may contribute to the higher mortality in this group include late diagnosis and the effect of a low index of suspicion of TB among clinicians. The high proportion of deaths classified as occurring before TB diagnosis suggests that opportunities for earlier diagnosis are being missed and that at least some of these deaths may have been avoidable.

Other estimates of TB mortality at the national level in the UK have used death (TB reported as the underlying cause) to notification ratios (DNR). This methodology is useful for examining mortality trends, but it does not provide a very reliable point estimate of mortality among TB cases. The numerator and denominator are not linked and it is not possible to relate the deaths to when they were diagnosed with TB. Furthermore, statutory notifications offer a less precise estimate of the incidence of TB. Counting only those deaths with TB as the underlying cause of death may also be too strict a definition to use to estimate the mortality burden attributable to TB.

The accuracy of death certification with regard to TB is unclear. The results of this study indicate that for some patients TB is reported as a contributing factor to death in the national surveillance system but no mention of TB is made on the death certificate. Without a detailed audit of the medical records of these patients and death certificates it is difficult to draw any firm conclusions about this issue. Under-reporting of TB in death certification was observed in New York in the 1990s, and increasing age was one factor associated with the omission of a TB diagnosis from the death certificate. Such a situation may be common to countries with significant numbers of elderly cases such as the UK.

For some deaths, the information recorded by treatment outcome monitoring was contradicted by information on the death certificate, with TB classified as incidental to the cause of death by the former but classified as the underlying cause of death in the latter. A small audit of deaths among TB cases in London also revealed such inconsistencies and led to a number of deaths being reclassified. Accurate information on the role of TB in the death of a patient is important for the interpretation of treatment outcome results, as it determines whether or not such a death could be considered a failure of TB services.

Each source of mortality information used in this study has its limitations. The NHS central register may potentially not hold information on all the patients reported with TB, especially people not born in the UK. The latter group accounts for the majority of reported TB cases in England and Wales (65% in 2001 and 2002). The database of death registrations used was limited to only those with a mention of TB, and treatment outcomes are not known for approximately 20% of cases. However, the discrepancy between the three information sources used suggests that the data linkage compensated for most of the deaths not recorded by any one source. The results of the capture-recapture analysis gave a relatively small number of unascertained deaths, which suggests that the study
was successful in identifying the majority of the deaths among the reported cases.

The estimate for the number of unascertained deaths is dependent on the accuracy of the data linkage. Data linkage is only as good as the methods used and the quality of the suitable identifiers. The study used deterministic rather than probabilistic methods for the data linkage work. Tracing to the NHS register was not operator assessed, and only approximately 30% of TB cases have an NHS number recorded. It is therefore possible that this study might have missed or misclassified some deaths. Misclassification would affect the degree of overlap between the three sources of mortality and result in either an overestimate or an underestimate of the number of unascertained deaths by capture-recapture.

Our figures may also be an underestimate of the CFR due to underdiagnosis or under-reporting of TB in groups at higher risk of death, such as patients with human immunodeficiency virus infection or the elderly. Under-reporting in the latter may be influenced by the frequency with which post mortems are performed. In some European countries, including the UK, the practice of autopsy is declining, but there may also be increasing diagnosis of TB post mortem due to improved diagnostic techniques.

The confidence and prediction intervals given for the capture-recapture estimates are only valid if the model is true. The saturated log-linear model makes an assumption about the conditional dependence of sources, and is based on the idea that each death has an equal and independent chance of being ascertained by each method. As is usual in observational studies, it is doubtful that all these assumptions are justified, so the results have to be taken as an approximate guide only.

In conclusion, by using three sources of information it was possible to identify a considerable underestimation of TB mortality through routine surveillance. The NHS programme for information technology (Connecting for Health) envisages real-time linkage with death registrations and patient records, including the information currently collected by national TB surveillance. This should provide the opportunity for continuous and robust estimates of TB mortality and CFRs. In the meantime, data linkage studies provide a better estimate of mortality among TB cases.

Acknowledgements
The authors are grateful to the following professionals without whose contribution national surveillance for TB would not be possible: TB nurses, chest physicians, microbiologists, the Regional Tuberculosis Coordinators of the HPA and D Quinn, the national database manager.

References
1 Abubakar I, Crofts J, Gelb D, Story A, Andrews N, Watson J M. Investigating urban–rural disparities in tuberculosis treat-
2 Antoine D, French C E, Jones J, Watson J M. Tuberculosis treatment outcome monitoring in England, Wales and North-
3 Health Protection Agency. First annual report on tuberculosis treatment outcome surveillance in England, Wales and North-
5 Health Protection Agency. Annual report on tuberculosis cases reported in England, Wales and Northern Ireland in 2003. Lon-
12 Heldal E, Naalsund A, Kongerud J, Tverdal A, Bee J. Deaths from active tuberculosis: can we rely on notification and mor-
13 Euro TB and the national coordinators for tuberculosis surveil-
ance in the WHO European Region. Surveillance of tubercu-
14 Cullinan P, Meredith S K. Deaths in adults with notified pul-
15 Farah M G, Tverdal A, Steen T W, Heldal E, Brantsæter A B, Bjune G. Treatment outcome of new culture positive pulmo-
16 Borgdorff M W, Veen J, Kalisvaart N A, Nagelkerke N. Mor-
20 Health Protection Agency. Tuberculosis surveillance: notifica-
tions and enhanced surveillance. Communicable Disease Rep Wkly
TB case mortality in England and Wales


**Résumé**

**Contexte :** Angleterre et Pays de Galles, 2001–2002.

**Objectif :** Obtenir une estimation plus précise de la léthalité par tuberculose (TB) au cours des 12 premiers mois après le début du traitement ou de la déclaration dans les cas de TB signalés au système national de référence.

**Méthodes :** On a identifié les dossiers de décès de tous les cas de TB signalés au système national de surveillance en utilisant les liens avec deux autres sources d’information sur la mortalité : le Registre Central du Service National de la Santé (NHS) et les Registres de décès provenant de l’Office des Statistiques Nationales (ONS). Ces données ont été comparées aux déclarations de décès mises en évidence par la surveillance nationale de la TB grâce au suivi des résultats du traitement. Une méthodologie capture-recapture a été utilisée pour estimer tous les décès résiduels mais non vérifiables persistants en vue du calcul final du taux de léthalité.

**Résultats :** On a identifié au total 1169 décès (IC95% 1140–1224) sur 13 176 cas, soit un taux de léthalité de 8,9% (IC95% 8,7–9,3%). Grâce au lien des données avec le registre central du NHS et les déclarations de décès provenant de l’ONS, on a pu identifier 255 décès supplémentaires non enregistrés par la surveillance nationale de la TB. La méthodologie de capture-recapture a estimé à 61 (IC95% 32–116) le nombre de décès non vérifiés.

**Conclusion :** Le système national de surveillance de la TB sous-estime la mortalité au sein des cas de TB. Les liens des données en temps réel entre les dossiers des cas et les enregistrements des décès permettraient une estimation plus précise et plus précoce de la mortalité par TB.

**Resumen**

**Marco de referencia :** Inglaterra y Gales entre 2001 y 2002.

**Objetivo :** Obtener una estimación más precisa de la mortalidad por tuberculosis (TB) en los primeros 12 meses después del comienzo del tratamiento antituberculoso o de la fecha de notificación de los casos a la red nacional de vigilancia.

**Métodos :** Se recopilaron todos los registros de defunción de casos de TB declarados al sistema nacional de vigilancia mediante un vínculo con otras dos fuentes de información sobre mortalidad: el registro central del sistema nacional de salud (NHS) y el registro de defunciones del departamento nacional de estadística (ONS). Estos datos se compararon con los registros de defunciones verificados por el sistema nacional de vigilancia de la TB por conducto de la supervisión de los desenlaces terapéuticos. El cálculo final del índice de letalidad se obtuvo aplicando el método de captura y recaptura a fin de incluir todas las defunciones no confirmadas.

**Resultados :** Se encontró un total de 1169 defunciones (IC95% 1140–1224) en los 13 176 casos; la tasa de letalidad fue 8,9% (IC95% 8,7–9,3%). El vínculo con el registro central del NHS y el registro de defunciones del ONS puso en evidencia 255 defunciones adicionales, no detectadas por el sistema nacional de vigilancia de la TB. Mediante el método de captura y recaptura se calcularon 61 (IC95% 32–116) defunciones no verificadas.

**Conclusión :** El sistema nacional de vigilancia de la TB subestima la tasa de mortalidad por esta enfermedad. El vínculo en tiempo real entre las notificaciones de casos y los registros de defunciones permitiría un cálculo más preciso y oportuno de la mortalidad por TB.