Background paper number 5

Assessing the quality and coverage of surveillance data in the UK: how is it done and how can HPA help to develop the Task Force framework for assessment of surveillance data, particularly the standards/process required for certification/accreditation and definition of the role of capture-recapture studies?

Ibrahim Abubakar

Question to the Task Force

1. How do you think the assessment of the data quality in the UK relates to the WHO Task Force framework?
Background Paper No. 5

Assessing the quality and coverage of surveillance data in the UK: how is it done and how can HPA help to develop the Task Force framework for assessment of surveillance data, particularly the standards/process required for certification/accreditation and definition of the role of capture-recapture studies?

Purpose
This paper summarises the context and history of tuberculosis (TB) surveillance, describes measures currently used to assure the quality of surveillance data and provides an initial assessment of extant systems in the UK using the WHO data quality model. This paper should be read in conjunction with the PowerPoint presentation to the WHO Global Task Force on TB Impact Measurement Meeting held in Geneva, March 2010.

Background
The incidence of tuberculosis in the UK has increased steadily over the last two decades with rates of tuberculosis exceeding 14 per 100,000 and over 8500 cases reported each year. Cases of tuberculosis are unevenly distributed geographically with the majority in major urban centres (figure 1) and over two thirds are diagnosed in the non-UK born population. The Chief Medical Officer (CMO) for England published a tuberculosis action plan in 2004, which forms the basis of the national policy to control tuberculosis. The action plan has 10 key areas, one of which is strengthening surveillance. In 2008, the Department of Health published a commissioning toolkit to support the planning and implementation of tuberculosis services. This toolkit outlined the minimum quality standards for surveillance (extract in appendix 1).

Figure 1. Map showing tuberculosis incidence in the UK
Surveillance systems
The UK has a long history of surveillance of tuberculosis. In the 1800’s, statutory recording of deaths was instituted in England leading to one of the oldest reliable sources of tuberculosis trends. A clinical diagnosis of tuberculosis became statutorily notifiable in 1917 in England. Data on cases of tuberculosis have therefore been available for nearly a hundred years. The statutory notification only collected minimal information necessary to monitor overall trends and inform public health action.

In 1994, a UK wide network of reference laboratories and the national tuberculosis surveillance team initiated a national system for monitoring all laboratory confirmed cases of tuberculosis in a central database. This collaboration, known as the UK Mycobacterial Network (MycobNet) now holds over 15 years of data on culture confirmed cases of *Mycobacterium tuberculosis* complex with information recorded on species and drug susceptibility results.

Prior to 1999, detailed information on tuberculosis cases was monitored through five yearly surveys the last of which was undertaken in 1998. From 1999, a national enhanced tuberculosis surveillance system was initiated using a distributed system of databases at the district, regional and national level. This system collected detailed clinical and demographic data. Data are annually matched to the national laboratory *M. tuberculosis* (MycobNet) database.

Following the publication of the national tuberculosis action plan, a decision was taken to replace the distributed model of enhanced surveillance with a national web based enhanced surveillance system. This system integrates the laboratory and clinical surveillance systems and has successfully being rolled out nationally. In London, a web based register has been in existence since 2002.

Further planned developments of the system, which are currently being undertaken, include a strain typing and contact tracing application to inform the public health management of tuberculosis clusters.

Quality systems

Tuberculosis surveillance, similar to other aspects of national surveillance are subject to periodic audit at the national level by the Health Protection Agency Internal Audit department, the last of which was undertaken in 2009. In addition, all HPA surveillance systems were evaluated in 2008.

The processes for running national surveillance are governed through a system of standard operating procedures.

The tuberculosis section at the HPA Centre for Infections, is responsible for the national surveillance of tuberculosis. In addition to the collection, analysis and reporting of tuberculosis data, the tuberculosis section also quality assures national tuberculosis surveillance.
Elements of quality that are continually monitored include the completeness, accuracy, timeliness and validity of data. Systems are continually improved and the effects of such changes monitored.

**Completeness**
The tuberculosis section utilises three approaches to monitor the completeness of surveillance data. The first category relies on inventory methods based on matching with other data sources including the laboratory (MycobNet) database, national HIV surveillance, death registration and bespoke surveys. These systems in general show between 5 to 17% undernotification to the enhanced surveillance system. A detailed audit of unmatched laboratory M. tuberculosis isolates recorded in the MycobNet database, however, revealed that the proportion is closer to 5% than 17%, as several cases reported as failing to match, are either present in the enhanced surveillance system with few identifiers, or may relate to organisms that should not have been reported such as those from non-human cases. The matching systems have been continually improved and now rely on a probabilistic programme (see appendix 2 for validation of matching programme). Any matching programme is, however, only as good as the quality of identifiers reported by local clinical teams.

The second approach relies on capture recapture based assessment of completeness using hospitalisation records (health care administration data), laboratory (MycobNet) records and death registrations. This allows the estimation of records that are not notified to any surveillance system. Previous studies using this approach have suggested undernotification to the surveillance systems of about 17%. This approach is, however, not without significant limitations due to the inherent dependencies between the data sources and quality of the matching as outlined above.

The most recent approach for estimating completeness that is under development by the Health Protection Agency will be based on “back-calculation” as widely applied to HIV and Hepatitis surveillance. The planned approach will use mortality data, knowledge of time to event, and Bayesian estimates of uncertainty to back project the expected number of tuberculosis cases in any time period. The proportion of deaths identified during post mortem examination will allow the determination of missed cases who presented to the health care system but were misdiagnosed.

**Accuracy**
Audits of accuracy of records are undertaken annually based on the criteria suggested in Department of Health’s commissioning toolkit (appendix 1). Other measures to improve accuracy include the removal of duplicate records by comparing data reported from different regions, automated systems to ensure that clinicians are reminded to denotify cases who turn out not to have tuberculosis, checks for the validity of fields in the web-based system such as NHS number validity, a requirement to have a minimum number of fields completed (this, however, requires a trade-off with completeness). Other logical checks in the web-
based system include those for dates and sites of disease. In addition, the system
derives a number of administrative variables from postcodes reducing data entry
errors.

**Timeliness**
Previous audits of timeliness suggested that nearly all regions are able to provide
data by the agreed deadline for the national collation of figures, February for
provisional data and July for finalised data.

The implementation of the web-based surveillance system now means
instantaneous availability of these data to all levels. However, the finalisation of
national datasets still requires significant work at the national level to improve final
data quality.

**Validity**
The programme of audit by the HPA internal audit department described earlier is
designed to ensure that a random selection of regions are assessed with records at
all levels compared to ensure that data recorded at the national level are valid. The
matching audit also provides a means to monitor the validity of data.

Further measures to improve validity include a data dictionary for the web-based
system, training events provided for new users and a webcast recorded to help new
users.

**Assessing trends and conclusion**
Comparison of trends over time using a variety of sources, for example laboratory
(MycobNet) data and statutory notifications, suggests that although the
completeness of notification may not be 100%, the data provides a reliable means
for monitoring trends (figure 2). A similar level of consistency is also observed when
overall trends are compared to immigration patterns. It is possible, that countries
with smaller populations and comparative low TB incidence may achieve better data
completeness levels. It is, however, unlikely that similar low incidence countries
with large populations will have better levels of data completion if they apply similar
rigorous evaluation criteria. Based on this, we would recommend that data
completeness thresholds for all, especially high burden, countries should not aim for
100% completeness but at broad measures of quality (completeness, accuracy,
validity and timeliness) with realistic targets and checks that data quality remains
consistent over time to allow monitoring. Measures of accuracy, validity and
timeliness of data are all reasonably high in the UK. Based on this we have concluded
that UK tuberculosis surveillance provides a good measure of trends.
Figure 2 Comparison of enhanced surveillance, statutory notifications and laboratory reports of tuberculosis cases in England, 2000 to 2008.
Appendix 1

Reporting new cases by clinical teams/local TB services (case definitions are given in the appendix on page 39)

• All cases should be reported by the clinical team to the local health protection unit.
• At least 95% of cases should be reported within two weeks of diagnosis or decision to treat with a full course of anti-TB drugs.
• At least 95% of reported cases should include complete data for the key variables (see appendix on page 40 for the key variables).
• At least 95% of all originally notified cases of TB that are subsequently denotified, should be reported within two weeks of the date of the non-TB diagnosis.

High priority fields:
• name;
• date of birth;
• sex;
• ethnic group;
• born/not born in the UK;
• postcode (with option for ‘no fixed abode’);
• date of notification;
• previous TB treatment;
• site of disease (pulmonary/extra-pulmonary); and
• sputum smear status (only needs to be completed for pulmonary cases).

Appendix 2

Validation of probabilistic matching algorithms