TB Data Quality Assurance: United Kingdom

Ibrahim Abubakar, MBBS, PhD, FFPH
Tuberculosis Section
HPA Centre for Infections
• 8,655 cases in 2008
• 14.1 per 100,000
• Increase over last 2 decades
TB Surveillance systems

INCIDENT CASES
- ETS
- NOIDS

TREATMENT OUTCOMES

FINGERPRINTS
National strain typing database

ISOLATES
MycobNet

Death Registrations

INCIDENTS & OUTBREAKS
TB surveillance – data flow

- Clinician & nurse: case of TB
- Specimen
- Feedback
- Local laboratories
- Species, sensitivities
- Culture
- Reference Labs
- Matching
- MycobNet
- Web-based
- HPA national database
- Regional Coordinator
- CCDC in HPU
DATA QUALITY
- Completeness
- No duplications, no misclassifications
- Internal and external consistency

TRENDS
Do surveillance data reflect trends in TB incidence and mortality?
- Analyse time-changes in notifications and recorded deaths alongside changes in case-finding, case definitions, HIV prevalence and other determinants of changes in TB incidence and TB mortality

ARE ALL TB CASES AND DEATHS CAPTURED IN SURVEILLANCE DATA?
- "Onion" model
- Inventory studies
- Capture re-capture studies
- Prevalence surveys
- Innovative operational research

IMPROVE surveillance system

EVALUATE trends and impact of TB control

UPDATE estimates of TB incidence and mortality

TB notifications ≈ TB incidence
TB deaths in VR system ≈ TB mortality

If appropriate, CERTIFY TB surveillance data as direct measure of TB incidence and mortality
Quality Scheme

• Components
  – Completeness
  – Accuracy
  – Validity
  – Timeliness

• Single Quality System - QPulse
Assessing completeness

- “Inventory Method”
  - Comparison of laboratory and clinical reports
  - TB-HIV matching
  - Bespoke audits: British Paediatric Surv Unit

- Capture Recapture
  - Overall data
  - Mortality

- Modelling from mortality
Completeness method 1: Matching Clinical and Lab Register, by region, 2003 and 2004

<table>
<thead>
<tr>
<th>Region</th>
<th>2003</th>
<th>2004</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>East Midlands</td>
<td>19/290</td>
<td>7</td>
</tr>
<tr>
<td>East of England</td>
<td>28/229</td>
<td>12</td>
</tr>
<tr>
<td>London</td>
<td>356/2131</td>
<td>17</td>
</tr>
<tr>
<td>North East</td>
<td>16/104</td>
<td>15</td>
</tr>
<tr>
<td>North West</td>
<td>90/360</td>
<td>25</td>
</tr>
<tr>
<td>South East</td>
<td>139/387</td>
<td>36</td>
</tr>
<tr>
<td>South West</td>
<td>16/132</td>
<td>12</td>
</tr>
<tr>
<td>Wales</td>
<td>24/140</td>
<td>17</td>
</tr>
<tr>
<td>West Midlands</td>
<td>58/510</td>
<td>11</td>
</tr>
<tr>
<td>Yorkshire &amp; the Humber</td>
<td>58/332</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>804/4615</td>
<td>17</td>
</tr>
</tbody>
</table>
Completeness Method 1: Reasons for not matching and solutions implemented

**Shouldn’t be on Laboratory Register (MycobNet)**
- False positives (cross contamination) – not denotified
- Not UK resident
- Non-human isolate
Solution: Automated identification of such isolates

**In Clinical Register (ETS) but not matched successfully**
- Case report more than 1 year from specimen date
- Missed by matching method (misspelt, insufficient completed variables to match etc.)
Solution: Improve matching algorithm with stochastic modelling, use of Soundex codes, local responsibility for matching

**Should be on Clinical Register (ETS), but isn’t**
- Treatment started but ETS not completed
- Patient lost to follow up (Treatment never started despite +ve culture)
- Patient died and ETS never completed

Solution: Identification of districts with particular issues, local audits
Completeness Method 2: Source of information on TB/HIV co-infection England and Wales

- TB reports from clinicians (Enhanced TB surveillance)
- TB reports from laboratories (MycobNet)
- HIV/AIDS patients dataset
  - Clinical HIV reports collecting demographic and epidemiological information
Completeness Method 2: Proportion of patients with active TB reported only to the HIV surveillance database

<table>
<thead>
<tr>
<th>Year</th>
<th>Not reported as active cases of tuberculosis to tuberculosis surveillance database</th>
<th>Total co-infected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number (% of total co-infected)</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>40 (23.7)</td>
<td>169</td>
</tr>
<tr>
<td>2000</td>
<td>52 (20.6)</td>
<td>253</td>
</tr>
<tr>
<td>2001</td>
<td>54 (17.2)</td>
<td>314</td>
</tr>
<tr>
<td>2002</td>
<td>78 (17.0)</td>
<td>459</td>
</tr>
<tr>
<td>2003</td>
<td>99 (18.1)</td>
<td>548</td>
</tr>
<tr>
<td>Total</td>
<td>323 (18.5)</td>
<td>1,743</td>
</tr>
</tbody>
</table>

Ahmed A, Thorax 2007
Completeness Method 3: Bespoke Paediatric TB Survey

- BPSU 320, ETS 446, Matched: 222
- Total: 557; 98 (18%) reported only to BPSU, and 237 (42%) reported only to ETS
- 19% under-reporting to ETS
- Matching on year, partial postcode, gender

BPSU – British Paediatric Surveillance Unit, ETS – Enhanced Tuberculosis Surveillance

Teo S et al. ERJ. 2008
Completeness Method 4: NATURALIST'S NOTEBOOK

Saturday

17 spotted butterflies caught, marked and released.

Monday

13 spotted butterflies caught, of which 3 have marks from Saturday.

Calculation

$\frac{3}{13} = 23\%$ of all butterflies were caught and marked on Saturday.

$23\% = 17$, therefore $100\% = 17 \times 100 \div 23 = 74$.

There are 74 spotted butterflies in the neighbourhood.
Completeness Method 4: RESULTS

Four years' incidence of TB in England

Data:

\[
\begin{align*}
&7860 \quad 6311 \quad 879 \\
&6437 \\
&4192 \quad 479 \\
&5848
\end{align*}
\]

Results: \( n = 32,006 \)

\[
\text{CRC } n_0 = 20,523 \ (18,109, \ 23,259)
\]

Under-notification = 15.9%

Van Hest et al, Epid and Inf 2008
Completeness Method 4: Capture recapture - Mortality

Treatment Outcome 853, 77%

NHS Register 888, 80%

Death Certificates 527, 48%

61 additional deaths estimated

Crofts J, IJTLD 2008
PROBLEMS

*List dependency:* You are more likely to get on to Laboratory Register if you are on Hospital Admission Data (HES).

*Case dependency:* You are more likely to be listed if your mother is listed.

*Case heterogeneity:* You are less likely to be diagnosed if you have no GP.

*Overcount:* Some people listed do not meet the case definition......especially true for Hospital Admission Data (HES)
Capture Recapture Recommendations

• Should never take precedence over routine surveillance strengthening
• Best when adequately resourced
• Should have at least three data sources
• May provide useful information on data sources and how they relate to each other
• Provides a broad estimate of under-notification that should always be interpreted with local knowledge
Assessing completeness estimates: modelling from mortality

• Back-calculation from mortality
• Bayesian approach allowing incorporation of prior knowledge on progression to death with and with/without treatment
• This combined with mortality data, hospital admissions, and surveillance data allows estimates of incidence and credible intervals to be produced.
Improving completeness: users are prompted to match isolates to case reports or to make new case reports.
Improving completeness

- Local audits in the South East region compared to prescription data and other sources
- Dedicated resources including additional staff
- New electronic system
Assessing accuracy

• Annual audit of 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).
## Assessing accuracy

<table>
<thead>
<tr>
<th>Information</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>100</td>
</tr>
<tr>
<td>Date of birth</td>
<td>99.9</td>
</tr>
<tr>
<td>Sex</td>
<td>99.8</td>
</tr>
<tr>
<td>Ethnic group</td>
<td>95.4</td>
</tr>
<tr>
<td>Born in UK</td>
<td>91</td>
</tr>
<tr>
<td>Postcode</td>
<td>99.4</td>
</tr>
<tr>
<td>Date of case report</td>
<td>100</td>
</tr>
<tr>
<td>Previous TB treatment</td>
<td>75</td>
</tr>
<tr>
<td>Site of disease</td>
<td>98.3</td>
</tr>
<tr>
<td>Sputum smear status</td>
<td>82</td>
</tr>
<tr>
<td>Treatment Outcome</td>
<td>95</td>
</tr>
</tbody>
</table>
Improving accuracy

- Training
  - Training event
  - Web-cast
- Automated checks in the web based system
  - NHS number validity
  - Minimum number of fields – trade off
  - Dates
  - Sites of disease
  - Derived fields from postcodes
Improving accuracy: Duplicates identified / Not TB denotified
Assessing validity

• Audit by national internal audit department comparing local, regional and national data

• Programme of audit by TB section
  – Matching audit
  – Next one will be on DOT
Improving validity

- Audit
- Data dictionary
- Training events
- Web casts
Timeliness

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.
Timeliness

• Assessing timeliness: annual audit of reporting dates against national standards

• Improving timeliness
  – High priority identified nationally
  – Use of web-based system – however.......
  – Cleaning, validation, record linkage, audits take a lot of time
Applying the Onion Model

- Cases recorded in TB notification data
- Diagnosed by NTP but not notified
- Cases presenting to health services but not diagnosed
- Cases with access to health services that do not present
- Cases with no access to healthcare

- Free healthcare for all via the National Health Service
- Contact tracing, mobile x-ray screening and ‘find and treat’ initiatives
- Clinical expertise, guidelines and quality assured laboratory services
- Minimal private sector in the UK

Focus for UK surveillance
Trends show consistency with immigration patterns

*Immigrants from non-EU and non-commonwealth countries

Sources: Enhanced Tuberculosis Surveillance, Office for National Statistics Long-Term International Migration
Trends show consistency across data sources

Sources: Enhanced Tuberculosis Surveillance, Enhanced Surveillance of Mycobacterial Infections, Statutory Notifications of Infectious Diseases (England and Wales only), UK Mycobacterial Surveillance Network
Resources

- **Staff**
  - National TB Epidemiology Unit (see organogram below)
  - Regional units (London *5, other units about 0.3 to 1 person per region)
Conclusion

• Complex national system for historical reasons
• Recent improvement in systems
• Trends reflect real changes over time
• Good routine surveillance requires resources