Assessing TB incidence by comparing TB notification data with vital registration data: an example from Brazil

Ana Bierrenbach
WHO / STOP-TB / TME
The Onion Model

- All incident cases
- Access to health facilities
- Presenting to health facilities
- Diagnosed by public or private providers
- Diagnosed by NTP or collaborating providers
- Recorded in notification data

TB deaths

- Undiagnosed cases
- Diagnosed but not notified cases
- Notified cases
Background

• On-going initiative in Brazil to reduce the number of death records with ill-defined causes of deaths

  – Probabilistic record-linkage between the mortality database and various health information databases: Health insurance, cancer, high-cost procedures (QTX, dialysis) and notified chronic diseases

  – Validation of the cause of death by verbal autopsy and/or examination of full medical records in hospital and out-patient clinics

• Brazil had some doubts about the WHO CDR estimates
## TB burden 2005

<table>
<thead>
<tr>
<th></th>
<th>Notification data</th>
<th>WHO estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number incident cases (all forms)</td>
<td>74,113</td>
<td>111,050</td>
</tr>
<tr>
<td>TB all forms rates</td>
<td>41 per 100 K</td>
<td>60 per 100 K</td>
</tr>
<tr>
<td>Case detection rate</td>
<td>x</td>
<td>69%</td>
</tr>
</tbody>
</table>
Objectives

• Estimate TB case fatality rate

• Estimate TB incidence rates using the following equation:

\[
\text{incidence}_{TB} = \frac{\text{deaths}}{\text{proportion cases dying (case fatality rate)}}
\]
Record linkage burden

TB NOTIFICATION
403,318 unique records (2001 - 2005)

TOTAL DEATHS

TB deaths
19,766

Need to verify if any of the 400,000 notifications appear in 4 million death records; and how many of the 20,000 TB deaths can be matched with 400,000 TB notification records
TB notifications

Linked TB deaths

Not-linked TB deaths

All Deaths
Key assumptions to estimate TB incidence from VR data and record-linkage

- Vital registration data **better** than TB notification (coverage and quality)
  - Comparison of VR with census and population counts projections
  - Ill-defined <10%
  - Recent paper (França *et al*) in IJE 2008;37:891-901

- TB cases not registered in the TB notification registry would have similar case fatality rates as those registered

- Good linkage procedures: reduced matching errors (no false-positives or false-negatives)
Probabilistic record-linkage methods

- Probabilistic linkage X Deterministic linkage

- Link-plus (CDC) and in-house software

- Name of patient, name of mother, date of birth as matching variables

- Address, notifying unit, … used to verify matches

- Data in both databases cleaned prior to linkage, with removal of invalid duplications
Results linkage
1.2% of all death found in TB notification data
8.2% of TB notifications found in VR mortality data (up to 4 years of follow-up)

5% of deaths from all causes in DOTS cohorts (2005, 6 months of follow-up)
TB death records linked to TB notifications

54% of TB deaths found in TB notification

Overall probability of a TB death record being linked to a TB notification
TB death records linked to TB notifications

TB notifications

TB Deaths

96.5% 3.5%

Overall probability of a death from TB among linked records

3.5% of TB notification found in TB death data
WHO estimate for CFR 2005 = 13%
Reasons for non-linkage

• Under-notification to TB notification (↑ CFR)
  – Diagnosis post-mortem or just before death?
  – Diagnosed by non-NTP but not notified
  – TB notification prior to 2001

• Low sensitivity linkage (↓ CFR)
  – Missing data in matching or verifying variables
  – Common names
  – Homonyms

• Wrong diagnosis of cause of death (↑ or ↓ CFR)
Survival analysis – Kaplan-Meyer curve

Probability of TB record being linked to TB notification

Last point in the curve is the prob. at 4 years of follow-up: used in subsequent calculations
Calculating CFR

CFR = probability of death at 4 years of follow-up among linked records
    probability of linkage of TB death records

<table>
<thead>
<tr>
<th></th>
<th>Probability of death from TB among linked records (at 4 years of follow-up)</th>
<th>Probability of TB death record being linked to TB notification</th>
<th>CFR (at 4 years of follow-up)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>0.0475</td>
<td>0.0543</td>
<td>8.8%</td>
</tr>
</tbody>
</table>
TB CFR for each of the 27 Brazilian states, by sex and age group
Survival analysis – Kaplan-Meyer curve, by HIV status

Probability of death from TB among linked TB notifications

- **Positive**
- **Negative**
- **Unknown**

Analysis time (years)
Calculating TB incidence - 2005

- TB reported deaths – 7,510
- Death detection rate – 90%
- Estimated TB deaths – 8,344
- Estimated death rate – 4.5 per 100K

\[
\text{incidence}_{TB} = \frac{\text{deaths}}{\text{proportion cases dying (case fatality rate)}}
\]

\[
95,408 = 8,344 / 0.087
\]
## Revision of WHO estimates

<table>
<thead>
<tr>
<th></th>
<th>Notification data</th>
<th>WHO original estimates</th>
<th>WHO revised estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number incident cases (all forms)</td>
<td>74,113</td>
<td>111,050</td>
<td>95,408</td>
</tr>
<tr>
<td>TB all forms rates</td>
<td>41 per 100 K</td>
<td>60 per 100 K</td>
<td>51 per 100K</td>
</tr>
<tr>
<td>Case detection rate</td>
<td>x</td>
<td>69%</td>
<td>78%</td>
</tr>
</tbody>
</table>
Discussion points

• Methodological issues
  – Benefits of linkage with whole death record database X just TB death records
  – Years of follow-up

• Feasibility of repeating the study in Brazil
  – Just for TB deaths (underlying + sequelae + associate)
  – Investigation of why TB deaths are not in TB notification system

• Feasibility of doing the study elsewhere
  – Electronic databases
  – Advanced computational and data management capacity
  – Human resources capacity