Inventory study to measure under-reporting: sample size calculation based on precision

\[ N = N_{SRS} \times \text{DEFF} \]

\[ N = \left[ 1.96^2 \frac{1 - \pi_g}{t^2 \pi_g} \right] \times \left[ 1 + (\bar{m} - 1) \frac{k^2 \pi_g}{1 - \pi_g} \right] \]

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>Number of TB cases to be included in the study</td>
</tr>
<tr>
<td>( \pi_g )</td>
<td>“Prior guess” of the true level of TB under-reporting (expressed as a proportion)</td>
</tr>
<tr>
<td>( t )</td>
<td>Relative precision (expressed as a proportion). Recommended 0.20 or 0.25</td>
</tr>
<tr>
<td>( \bar{m} )</td>
<td>Harmonic mean of cluster size (=number of TB cases found in health-care facilities)</td>
</tr>
<tr>
<td>( k )</td>
<td>Coefficient of between-cluster variation. Recommended to assume is in the range 0.4 – 0.6</td>
</tr>
</tbody>
</table>
1. A prior guess of the true population prevalence

- First step is to make a prior guess of the true level of TB under-reporting

- Use previous relevant study survey results (if surveys comparable)

- Always include a pilot phase in your inventory study with an interim analysis to confirm the assumption of your sample size calculation
  - Start with a small number of clusters, say 5-10, recalculate DEFF and total sample size N

- Must be done in close collaboration between a team including a statistician and local TB experts
2. The relative precision

- This (relative) precision refers to "how far away" we are allowing the survey's estimate of TB under-reporting to be from the true

- In statistical terms this translates into the width of the 95% confidence interval around the TB prevalence estimate we expect the survey to give us

- The higher the precision the larger the sample size

- Recommended precision is between 20% and 25%

  e.g. If under-reporting is 20%, then 95% CI (16%, 24%)
Required sample size \( (N_{SRS}) \) for different assumption about under-reporting \( (\pi_g) \) and relative precision \( (t) \)

<table>
<thead>
<tr>
<th>( \pi_g )</th>
<th>( t = 5 )</th>
<th>( t = 10 )</th>
<th>( t = 20 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>29 196</td>
<td>7 299</td>
<td>1,825</td>
</tr>
<tr>
<td>10</td>
<td>13 830</td>
<td>3 458</td>
<td>865</td>
</tr>
<tr>
<td>15</td>
<td>8 708</td>
<td>2 177</td>
<td>545</td>
</tr>
<tr>
<td>20</td>
<td>6 147</td>
<td>1 537</td>
<td>385</td>
</tr>
<tr>
<td>25</td>
<td>4 610</td>
<td>1 153</td>
<td>289</td>
</tr>
<tr>
<td>30</td>
<td>3 586</td>
<td>897</td>
<td>225</td>
</tr>
<tr>
<td>35</td>
<td>2 854</td>
<td>714</td>
<td>179</td>
</tr>
<tr>
<td>40</td>
<td>2 305</td>
<td>577</td>
<td>145</td>
</tr>
</tbody>
</table>
3. A prior guess of the "design effect"

- We sample groups of TB cases (clustered-random design), and not individuals (simple-random design).
- Clustered-sampled (CS) studies have more uncertain results compared to simple-random sampled ones (for given assumptions).
- To account for this we need an increased sample size (multiply sample size for SRS by a factor called the "design effect").
- We estimate it from previous studies OR the likely variation of under-reporting between clusters.
- Design effect gets bigger the bigger the:
  - difference in under-reporting between clusters,
  - harmonic mean of cluster size,
  - expected TB prevalence is
4. Number of geographical areas to sample

- Geographical areas (= clusters) should be selected in such a way that total sample size will not be reached only by the inclusion of a very limited number of them (e.g. BMU's)
- Cluster size (= number of all TB cases diagnosed by all providers in the geographical area) is unknown and cannot be considered fixed
  - Use case notifications as a proxy to understand cluster size
- The number of clusters ($c$) to sample should be decided on a case-by-case scenario. One suggestion is:

$$c = \frac{N}{\text{median}}$$

($N$ sample size / expected median value across all areas in the sampling frame)
5. Stratification

- TB under-reporting will typically vary across different geographical regions or types of facilities (*strata*)

- A *stratified* design should be used to increase the precision and representativeness of the overall country

- This means number of clusters from each stratum is proportional to population size in the stratum

- By design the approach of stratification allows the estimation of stratum-specific estimates of TB, but their precision is smaller compared to the overall nationwide estimate

- Think about which strata apply to your country!
  - Health-care facilities that diagnose children
  - Large non-NTP hospitals
  - High/low case notification rate