Estimating Mortality from Census Data: A record linkage study in the Nouna Demographic and Health Surveillance System in Burkina Faso

Bruno Lankoande

United Nations Expert Group Meeting, NYU Abu Dhabi

25 August 2022
Data on **recent household members** are affected by various errors including (Bennett et al., 1984; Hill, 1987; Timæus, 1991):

- Underreporting of deaths;
- Transfers outside the reference period;
- Age mistatement;
- Under enumeration of some specific populations.
• The magnitude and direction of these errors are difficult to assess in the absence of a mortality gold standard;

• Different methods to adjust estimates have been developed and mostly evaluated in simulated environments (Hill et al., 2009; Murray et al., 2010; Verhulst, 2016)

• Few attempts to compare them to high quality data from Health and Demographic Surveillance Systems (HDSSs) except in Senegal (Masquelier et al., 2016);
Using the Nouna HDSS as the reference, we evaluate the reliability of mortality indicators derived from the 2006 census of Burkina Faso

- Capture the magnitude of mortality underestimation in the census and their variation by age group and sex;

- Link individual records to evaluate the quality of ages and their impact on mortality estimates.
Data and methods

- Data collected in the Nouna HDSS since 1992 (Sié et al., 2010).

- Extract of individual-level data of the population under surveillance in the HDSS from the census database.
Comparisons at the aggregate level based on the names of villages

- Relying on the same methodology to compare summary indices of mortality between census and HDSS estimates.

- Decomposition of the differences in life expectancies at birth into contributions of the major age groups (Arriaga, 1984).
Record linkages

- Automatic search based on first and last names was performed using Jaro-Winkler distance.

- Manual search based on kinship graphs derived from the census and the HDSS.
Analysis at individual level

- Compared ages of the surviving population as well as of the deceased in 2006 across data sources.

- Computed a life table from the census using ages reported in the HDSS.
Comparisons at the aggregate level

Key findings

Figure 1: Population pyramid according to the HDSS and the Census, Nouna, 2006

- The male population is only 2% larger in the HDSS
- The female population is 7% larger in the HDSS, as compared to the census
**Figure 2** : Number of deaths reported by month according to the HDSS and the census, by age group, Nouna, 2006

- 18% fewer deaths were collected in men and 29.6% in women.
- Fewer deaths were particularly collected below age 5 and above 60.
Figure 3: Age specific mortality rates (ASMR) inferred from the census and the HDSS data, Nouna, 2006
Table 1: Direct estimates of mortality in Nouna according to the HDSS and the reporting of deaths that occurred in households during the last 12 months in 2006 census

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Census</th>
<th>HDSS</th>
<th>Rel. diff</th>
<th>Contribution</th>
<th>Census</th>
<th>HDSS</th>
<th>Rel. diff</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>5q0</td>
<td>124</td>
<td>128</td>
<td>-3%</td>
<td>0.2</td>
<td>97</td>
<td>115</td>
<td>-16%</td>
<td>1.4</td>
</tr>
<tr>
<td>10q5</td>
<td>19</td>
<td>24</td>
<td>-18%</td>
<td>0.2</td>
<td>16</td>
<td>21</td>
<td>-24%</td>
<td>0.3</td>
</tr>
<tr>
<td>45q15</td>
<td>291</td>
<td>306</td>
<td>-5%</td>
<td>0.3</td>
<td>166</td>
<td>218</td>
<td>-24%</td>
<td>1.2</td>
</tr>
<tr>
<td>20q60</td>
<td>532</td>
<td>652</td>
<td>-18%</td>
<td>2.5</td>
<td>417</td>
<td>584</td>
<td>-29%</td>
<td>3.7</td>
</tr>
<tr>
<td>e0</td>
<td>61.0</td>
<td>57.8</td>
<td>6%</td>
<td>3.2</td>
<td>68.4</td>
<td>61.8</td>
<td>11%</td>
<td>6.6</td>
</tr>
</tbody>
</table>
**Figure 4**: Age differences of deceased persons between the census and the HDSS in 2006 using the HDSS as a reference.
Table 2: Effects of age misstatement in the census on mortality indicators in men

<table>
<thead>
<tr>
<th>Indices</th>
<th>Census</th>
<th>Cen. corrected</th>
<th>HDSS</th>
<th>Rela. Diff(^1)</th>
<th>Rela. Diff(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5q0</td>
<td>124</td>
<td>124</td>
<td>128</td>
<td>-3%</td>
<td>-3%</td>
</tr>
<tr>
<td>10q5</td>
<td>19</td>
<td>21</td>
<td>24</td>
<td>-21%</td>
<td>-12%</td>
</tr>
<tr>
<td>45q15</td>
<td>291</td>
<td>300</td>
<td>306</td>
<td>-5%</td>
<td>-2%</td>
</tr>
<tr>
<td>20q60</td>
<td>532</td>
<td>540</td>
<td>652</td>
<td>-18%</td>
<td>-17%</td>
</tr>
<tr>
<td>e0</td>
<td>61.0</td>
<td>60.2</td>
<td>57.8</td>
<td>6%</td>
<td>4%</td>
</tr>
</tbody>
</table>

(1) Relative difference, uncorrected estimates vs HDSS
(2) Relative difference, corrected estimates vs HDSS
Table 3: Effects of age misstatement in the census on mortality indicators in women

<table>
<thead>
<tr>
<th>Indices</th>
<th>Census</th>
<th>Cen. corrected</th>
<th>HDSS</th>
<th>Rela. Diff$^1$</th>
<th>Rela. Diff$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5q0</td>
<td>97</td>
<td>96</td>
<td>115</td>
<td>-16%</td>
<td>-16%</td>
</tr>
<tr>
<td>10q5</td>
<td>16</td>
<td>16</td>
<td>21</td>
<td>-24%</td>
<td>-26%</td>
</tr>
<tr>
<td>45q15</td>
<td>166</td>
<td>222</td>
<td>218</td>
<td>-24%</td>
<td>2%</td>
</tr>
<tr>
<td>20q60</td>
<td>477</td>
<td>462</td>
<td>584</td>
<td>-29%</td>
<td>-21%</td>
</tr>
<tr>
<td>e0</td>
<td>68.4</td>
<td>67.7</td>
<td>61.8</td>
<td>11%</td>
<td>10%</td>
</tr>
</tbody>
</table>

(1) Relative difference uncorrected estimates vs HDSS
(2) Relative difference corrected estimates vs HDSS
Some limitations

- Age misreporting may affect some groups of individuals in the HDSS: migrants, enumerated population.

- Age errors may be larger among individuals we failed to matched compared to those who were successfully linked.
It is likely that mortality rates were underestimated in the 2006 census, particularly in elderly and women.

Omissions of deaths play a larger role than age errors in explaining the gaps.

There is a crucial need to develop innovative ways to improve the reporting of demographic events.

Comparisons in other HDSSs sites of SSA may be a starting point to inform adjustments made to census estimates.
Thank you