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

Bruno Lankoande on behalf of the WP3 team

Colloque DEMOSTAF

Paris - October 16, 2019

- In Sub-Saharan Africa, because of the absence of full-fledged CRVS system in most countries, mortality levels and trends are largely derived from large-scale surveys and censuses ;

- Surveys have collected birth and sibling histories ;

- Censuses have included questions on the survival of children, parents, and **recent household members** ;

- A complete life table can be obtained from data on the number of deaths in each household preceding the enumeration;

- But these data are affected by various errors including :
    - 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 ;

- Estimates have sometimes been evaluated in simulated environments ;

- Few attempts to compare them to high quality data from Health and Demographic Surveillance Systems (HDSSs) except in Senegal ;

Using the Nouna HDSS as the reference, we evaluate the reliability of mortality indicators derived from the last national census of Burkina Faso, conducted in 2006

- 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 ;

Background
Data and methods
Key findings
Conclusion

Data
Methods

- Data collected in the Nouna HDSS since 1992.

- Extract of Individual-level data of the population under surveillance in the HDSS from the census database.

Background
Data and methods
Key findings
Conclusion

Data
Methods

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.

Background
Data and methods
Key findings
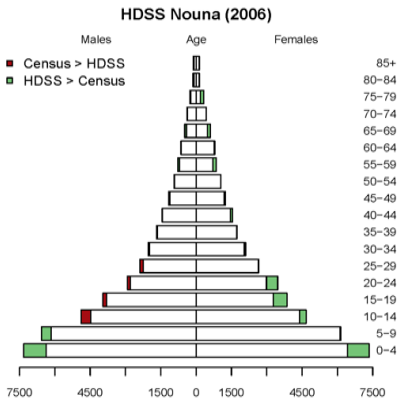Conclusion

Data
Methods

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.

Background
Data and methods
Key findings
Conclusion

Data
Methods

### Analysis at individual level

- Logistic regressions on the probability to be matched using socio-demographic characteristics.

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

- Computing a life table from the census using ages reported in the HDSS.

Background
Data and methods
Key findings
Conclusion

Comparisons at the aggregate level
Individual-level analysis

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

HDSS 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

Background
Data and methods
Key findings
Conclusion

Comparisons at the aggregate level
Individual-level analysis

Figure 2 : Number of deaths reported by month in 2006 in Nouna according to the HDSS and the census, by age group
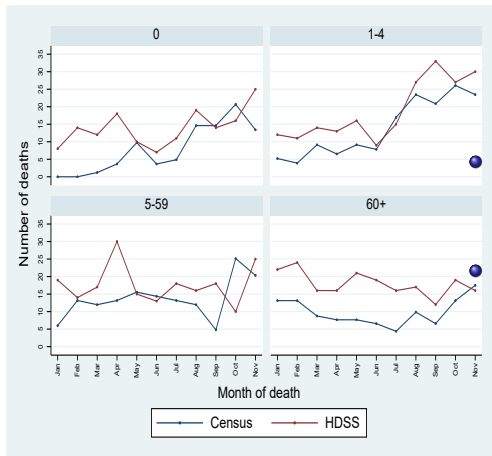
● '18% fewer deaths were collected in men and 29.6% in women

● Fewer deaths were particularly collected below 15 and above 60

Background
Data and methods
Key findings
Conclusion

Comparisons at the aggregate level
Individual-level analysis

Figure 3 : Age specific mortality rates (ASMR) inferred from the census and the HDSS data, Nouna, 2006

Background
Data and methods
Key findings
Conclusion

Comparisons at the aggregate level
Individual-level analysis
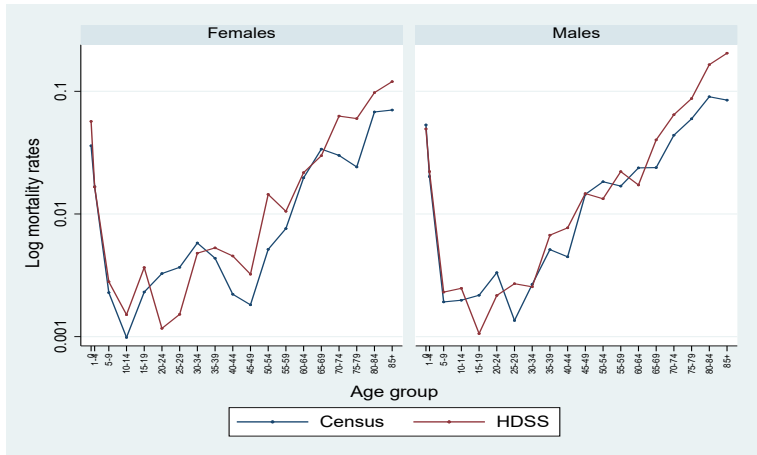
Table 1 : Direct estimates of mortality in Nouna according to the HDSS and the reporting of deaths that occurred in households during the last 12 months in 2006 census

Indices	Males				Females			
	Census	HDSS	Rela. diff	Contri.	Census	HDSS	Rela. diff	Contri.
$_5q_0$	124	128	-3%	0.2	97	115	-16%	1.4
$_{10}q_5$	19	24	-18%	0.2	16	21	-24%	0. 3
$_{45}q_{15}$	291	306	-5%	0.3	166	218	-24%	1.2
$_{20}q_{60}$	532	652	-18%	2.5	417	584	-29%	3.7
e_0				Diff.				Diff.
	61.0	57.8	6%	3.2	68.4	61.8	11%	6.6

Background
Data and methods
Key findings
Conclusion

Comparisons at the aggregate level
Individual-level analysis

Table 2 : Effects of age misstatement in the census on mortality indicators

Variables	Survivors		Deceased	
Matching rates	*58%*		*36%*	
Sex	Males	Ref.	Males	Ref.
	Females	0.977	Females	0.86
Age group	0-4	Ref.	0-4	Ref.
	5-14	0.765***	5-14	1.335
	15-29	0.553***	15-59	0.918
	30-39	0.722***	60-79	1.223
	40-49	0.847***	80+	0.913
	50-59	0.775***		
	60-69	0.721***		
	70-79	0.696***		
	80+	0.787***		
Observations	71,706		589	

*Statistical significance : *** $p < 0.01$; ** $p < 0,05$; * $p < 0.1$

Background
Data and methods
Key findings
Conclusion

Comparisons at the aggregate level
Individual-level analysis

Figure 4 : Age differences in men and women between the census and the HDSS in 2006 using the HDSS as a reference

Background
Data and methods
Key findings
Conclusion

Comparisons at the aggregate level
Individual-level analysis
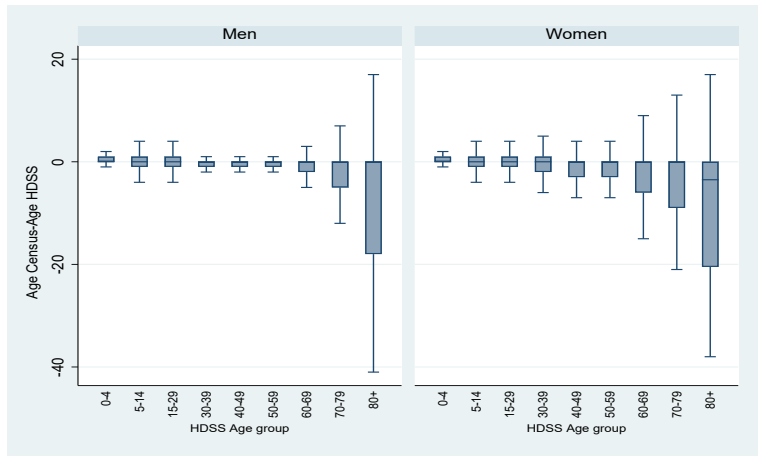
Figure 5 : : Age differences of deceased persons between the census and the HDSS in 2006 using the HDSS as a reference

Background
Data and methods
Key findings
Conclusion

Comparisons at the aggregate level
Individual-level analysis
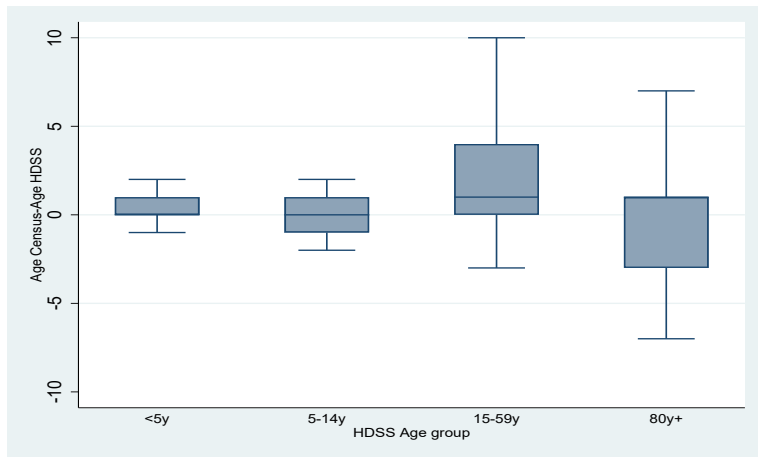
Table 3 : Effects of age misstatement in the census on mortality indicators in men

Indices	Census	Cen. corrected	HDSS	Rela. Diff[1]	Rela. Diff[2]
5q0	124	124	128	-3%	-3%
10q5	19	21	24	-21%	-12%
45q15	291	300	306	-5%	-2%
20q60	532	540	652	-18%	-17%
e0	61.0	60.2	57.8	6%	4%

(1) Relative difference, uncorrected estimates vs HDSS

(2) Relative difference, corrected estimates vs HDSS

Background
Data and methods
Key findings
Conclusion

Comparisons at the aggregate level
Individual-level analysis

Table 4 : Effects of age misstatement in the census on mortality indicators in women

Indices	Census	Cen. corrected	HDSS	Rela. Diff[1]	Rela. Diff[2]
5q0	97	96	115	-16%	-16%
10q5	16	16	21	-24%	-26%
45q15	166	222	218	-24%	2%
20q60	477	462	584	-29%	-21%
e0	68.4	67.7	61.8	11%	10%

(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, Enumarated population.

- Age errors mat be larger among individuals we failed to matched compared to those who were successfully linked

- It is likely that mortality rates 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 adjustements made to census estimates.