Age misreporting in censuses in developing countries: a record linkage study in health and demographic surveillance systems in Senegal

Gilles Pison (1, 2), Bruno Masquelier (3), Cheikh Tidiane Ndiaye (4), Lucie Lecomte (2), Ndèye Binta Dieme (4), Ibrahima Diouf (4), Valérie Delaunay (5)

(1) French National Museum of Natural History, Paris, France
(2) French National Institute for Demographic Studies (INED), Paris, France
(3) Université catholique de Louvain, Louvain-la-Neuve, Belgique
(4) Agence nationale de la statistique et de la démographie, Dakar, Sénégal
(5) Institut de recherche pour le développement, Dakar, Sénégal

Paper submitted to the European Population Conference (EPC2016)

Mainz (Germany), 31 August – 3 September 2016

Short summary

Ages reported in censuses in developing countries are subject to errors and bias resulting in uncertainties in population estimates and age distributions. We examine these biases in the case of the last two censuses of Senegal, conducted in 2002 and 2013, by matching at the individual level the information they collected with those held by the health and demographic surveillance systems in place since three decades in rural Senegal. The information on the ages held by these systems is of high quality and serves as a reference to examine the quality of the ages reported in the censuses. We quantify the differences between reported and actual ages, check if the errors and bias have been reduced or not between the 2002 census and that of 2013, and examine the consequences of age misreporting on demographic estimates in the studied populations.

Extended summary

I - Introduction

Ages reported in censuses and surveys in developing countries are subject to errors and bias resulting in uncertainties in population estimates and age distributions (Ewbank, 1981). It is important to know more about these errors and biases in order to correct them, and also try to reduce them by improving the data collection.

These errors and bias in age determination are most often studied by conducting analysis at the aggregate level, examining the distortions they produce in the age distributions. It is rare, however, that they can be observed directly by comparing the reported ages to actual ages. In this study, we use the last two censuses of Senegal, conducted in 2002 and 2013, and link individual records with data collected in three health and demographic surveillance systems (HDSS) in rural areas of Senegal (Bandafassi Mlomp and Niakhar). The information on the ages held by these HDSS are of high quality because particular attention was paid to their collection when setting up the surveillance. In addition, the date of birth of individuals born after the initial census was registered during the follow-up, so it is known precisely. Information collected by the HDSS serves here as a reference for evaluating the quality of the ages reported in the censuses.

In this paper, we quantify the differences between actual and reported ages and examine the influence of various factors – such as gender, actual age, marital status, household status, level of education of the person and of the head of household, - on reporting errors. Since the 2013 census introduced a new data collection method supposed to improve the data quality with handheld computers (commonly known as PDA), we also examine whether the errors and biases have reduced between 2002 and 2013.

II - Age misreporting in Sub-Saharan Africa in the absence of complete civil registration: a review

Censuses and surveys in African countries where civil registration is incomplete face problems in assessing ages. Reported ages can depart quite substantially from real ages. The impact on the population estimates can be serious because, to a general imprecision is often added systematic errors or biases, plaguing reports for certain ages or specific population groups. Conventionally, age statements are affected by two types of errors: age heaping, which is a tendency to round ages to specific digits (0 or 5), and the tendency to systematically exaggerate or, conversely, under-estimate ages (Ewbank 1981; Pullum 2006). These problems are exacerbated when individuals are reporting on the ages of other members of their household or family, as is frequent in household interviews.

a) Age heaping

Age distributions derived from African censuses and surveys can exhibit large irregularities when the reported ages are considered by single years of age; typically, ages ending by a multiple of five will “attract” more reports than surrounding ages (Gendreau and Nadot 1967). Census enumerators and survey interviewers may face difficulties in determining the exact age of a person with a precision of one year, and sensitive to heaping, preferentially attribute ages ending in 0 or 5. But there are also counter-examples. In some cases, such as in a survey in Guinea in 1955, investigators were strongly warned against the heaping, and this resulted in an opposite pattern showing deficit at these ages. In other cases, such as in the census of Senegal in 1988, there was a distinct heaping on ages ending in 9 or 4 (Pison et al. 1995). This was due to the conversion by investigators of ages into years of birth, themselves subsequently converted into ages, resulting at the end for some people in a reduction of one year of their age because of this conversion (Pison and Ohadike 2006). Several indices have been developed to assess the amount of digit preference, such as the Myer’s index or Whipple’s index (Spoorenberg and Dutreuilh 2007).

b) Systematic underestimation or overestimation of age

Artificial “aging” or rejuvenation are common in sub-Saharan Africa, but there has been relatively few attempts to understand patterns of age misreporting. The literature on age misstatement is somewhat outdated and little is known on age misreporting in recent surveys and censuses. Ewbank (1981) discusses a number of important issues related to age errors and selective underenumeration and refers to some direct evaluation studies. In the 1973 census in the Gambia for example, men tended to exaggerate their age by 2.5 years on average between 30 and 49 years of age, and at least 5 years beyond age 50 (Gibril 1979). On the contrary, ages of women generally tended to be understated when they were under 40 years, and their age was overstated when they were aged more than 40.

Another trend, especially in West Africa, is to underestimate the age of children and adolescents, when they have not yet reached puberty or have never been married, and to overestimate it when they are pubescent or get married. This results in a “hole” in the pyramids around ages 15 to 20, the age groups from 10 to 24 years being underrepresented in favor of the two neighbouring groups of 5-9 years and 25-29 years (Ewbank 1981; Gendreau and Nadot 1967; Hertrich and Lardoux 2014).

With the progress of education, ages should be better known, especially for younger generations, but it is possible that the reported ages continue to be affected by errors. While most people interviewed or identified in current large national surveys have some sort of ID, birth dates that they contain may not necessarily be reliable. They may have been purposely biased, for example if reporting a younger age is needed for a child too old to be enrolled in a class.

The determination of ages can be improved using particular techniques. Pison (1979) lists some of the most widely used: historical calendars (Blanc 1962; Scott and Sabagh 1970); collection of relative ages by ranking a whole population by birth order (Gubry 1975; Howell 1979; Pison 1980); checking the internal consistency of the reports within the questionnaire ; or the use of calendar of events - a systematization of the internal consistency of the questionnaire by establishing a file of all events recorded on an individual and reconstructing with him the order in which these events have occurred (Ferry 1976; Freedman et al. 1988; Helleringer et al. 2014).

Whatever the method used to improve the age determination, in most cases we do not know the direction and magnitude of the error, because the “true” age remains unknown. To measure this error, we must know the true age from registration-type data, such as information from parish registers, or civil registration. However, even in population well covered by civil registration, direct evaluation studies are difficult to conduct since the existence of an effective civil registration system obviously alters the awareness of ages, thus limiting the inference that can be made for settings without vital registration.

Health and demographic surveillance systems (HDSS) are a special case where the analyst can know the precise date of birth of almost all individuals in a population, while these individuals themselves might not necessarily know it. In Niakhar, Senegal, we were able to compare actual ages with ages reported in a census for children under age 15 in 1977, because the HDSS had collected accurate dates of birth for them independently over the previous 15 years. We showed that ages reported for children aged 14 or 15 understated their actual age by two years on average (Pison 1979).

Here we present a similar study to evaluate the reporting of ages in the last two censuses of Senegal (2002 and 2013). We extend the analysis to the three rural areas of Bandafassi, Mlomp and Niakhar in Senegal, whose populations have been followed up for decades. We match the information collected by the last two national censuses with those held by these HDSS to analyze the quality of age reporting in the censuses. We present in this extended summary the populations studied and the methods used for the record linkage. We also provide some preliminary results for two of the three HDSS. The full paper will present complete results from the three sites and lessons learned from the study.

II - Populations and methods

a) The health and demographic surveillance systems (HDSS) in Senegal

The study is based on three health and demographic surveillance systems (HDSS) in place in rural areas of Senegal: Bandafassi, Mlomp and Niakhar[1]. These HDSS have followed up the population of each area for several decades, collecting regularly information on demographic events through repeated rounds. After an initial census, villages were visited at regular intervals (once a year in recent times). During each visit, households are reviewed, the list of people in each household at the preceding visit is checked, and information on births, marriages, migration and deaths (including the cause of death) which occurred since are collected.

The three Senegalese HDSS are located in three different regions (see map below):

- Bandafassi. The site is located in Southeastern Senegal; the population has been followed up since 1970 with two successive population expansions in 1975 and 1980 (Pison et al. 2014) ; it had nearly 14,000 inhabitants in 2013,

[1] These health and demographic surveillance systems are member of the *INDEPTH* network (International network of field sites with continuous demographic evaluation of populations and their health in developing countries) (http://www.indepth-network.org/).

- Mlomp. The site is located in Southwestern Senegal; the population has been followed up since 1985 (Pison et al. 2002) ; it had nearly 8000 inhabitants in 2013,

- Niakhar. The site is located in the southwest; the population has been followed up since 1963 with a reduction of the study area in 1968 followed by an extension in 1984 (Delaunay et al. 2013) ; it had nearly 43,000 inhabitants in 2013.

Map: Localisation of the three health and demographic surveillance systems (HDSS) in Senegal

In each HDSS, the date of birth of the persons born after the start of surveillance is accurately known (often to the nearest month, and for some people, to the day). For those already present at the initial census, their birth date was determined from the reported ages at that time, often corrected afterwards using various sources (administrative censuses, maternity registers, child growth monitoring registers, etc.). Ages could be determined in some cases indirectly without asking age, by ranking people according to their age as mentioned above (Pison 1980). In total, the ages of the individuals monitored are known with unusual precision for rural populations in Africa.

We used here the dates of birth held by the HDSS as a reference for assessing the quality of the ages reported in the censuses.

b) Census data from 2002 and 2013

Senegal has organized four national censuses of population in 1976, 1988, 2002 and 2013. Here we use the last two censuses, since they are the only ones for which detailed nominative information was available for this study. In Bandafassi, we used only the 2013 census because some of the census questionnaires from 2002 could not be found, and names were not recorded in the database. For the other two HDSS, we recovered all names from the questionnaires.

The 2002 census was paper-based and conducted between Dec 8th and Dec 22. The 2013 census was conducted with PDAs between Nov 19 and Dec 9. It is worth noting that the HDSS and censuses do not use the same definition of residence, resulting in important differences in the reported populations. In most cases, the censuses will only consider as resident those who have stayed or intent to stay in a compound for 6 months or more. In the case of HDSS, the definition is broader, including migrants who might have resided elsewhere for more than 6 months. Hence populations in HDSS tend to be larger than according to the census.

b) The record linkage between censuses and HDSS data

The record linkage was conducted in two steps.

First, we performed an automatic search based on names of villages and hamlets, and first and last names of all individuals. List of villages and hamlets from the census were compared with that of the HDSS, and localities were harmonized. In some cases, names or delimitations did not correspond, requiring a detailed examination of the lists of households by locality to establish correspondence. Then, we extracted from the 2002 and 2013 censuses the data covering the three HDSS. The spelling of the most common first and last names was harmonized between census and HDSS.

An automatic search was then performed by comparing the full names of all individuals registered in one or the other data source and residing in the same village (or a village nearby). We used the Jaro-Winkler distance that measures the similarity between two chains of characters (Jaro 1989; Winkler 1999). Two records were considered provisionally similar if the Jaro-Winkler score was greater than 0.9. Two households were then provisionally matched by the automatic method if at least half of their members had a similar first and last name[2]. An additional condition for this temporary match was that the sizes of the two households should not be too different (a doubling size from one source to the other resulting in not matching the households).

In a second step, we checked all provisional matches manually, and examined non-matched cases, based on kinship charts and field verifications with local informants and HDSS interviewers. Two records with approximately the same name and the same place in the household structure were considered as the same individual. Rather than representing the household as a list of individuals, kinship charts were drawn based on the relationship to the household head in the case of the census, and in the case of HDSS, based on genealogical data. Linkages based on the automatic search were thus used as a guide only. In addition, we also conducted research directly into the databases, but the kinship charts turned out to be more useful. These charts were updated based on manual

[2] This automatic search was limited to the 10 or 20 oldest household members to reduce computing time.

linkages made in the office and later used for field verifications. In Mlomp, due to the small population size, we went directly to the field and did not perform manual linkages in the office beforehand.

Note that we did not use the age or date of birth of individuals for this record linkage, in order to be able to evaluate the quality of reported ages. Through the use of kinship charts, we did use – albeit loosely - some correlates of age misstatement, such as marital status and the number of children ever born.

III – Preliminary results

This section presents some preliminary results for Bandafassi and Mlomp, since the record linkage is still ongoing in Niakhar (it will be finalized before the 2016 EPC). We present matching rates, and examine the influence of some factors on the difference between actual age and age reported in the census: gender, actual age, marital status, relationship to the household head, level of education of the person and the head of household, etc.. Since the 2013 census has used a new method of data collection with the use of handheld computers (PDA), we will also examine whether the errors and biases have narrowed between the 2002 and 2013 censuses.

a) Matching rates

Table 1 present the linkage rates by site, starting from the population enumerated in the censuses. These rates vary by HDSS and by census, but we can match at least 75% of individuals in Bandafassi and Mlomp.

	Bandafassi 2013	Mlomp 2002	Mlomp 2013	Niakhar 2002	Niakhar 2013
	Automatic search based on names				
Compounds	27.6	33.2	31.7	89.3	86.8
Households	-	24.6	28.2	82.9	77.0
Individuals	11.8	15.0	17.7	43.5	56.1
	Manual linkage based on kinship charts				
Compounds	58.0	-	-	95.6	94.2
Households	-	-	-	94.7	90.9
Individuals	45.1	-	-	74.9	77.8
	Field verifications				
Compounds	87.7	95.2	86.9		
Households	-	92.2	86.4	*in progress*	
Individuals	74.8	87.6	83.4		

Table 1: Matching rates by site and census

Matched and unmatched cases were examined in Mlomp and Bandafassi to detect significant deviations from the general population (Table 2). We did not find any significant differences by gender. Matching rates varied, however, according to the relationship to the head. In Mlomp, spouses of the head of household were more likely to be matched to an individual in the HDSS than the head himself, but this difference was not significant in Bandafassi. In Mlomp, children and in-laws were also more likely to be matched. Finally, grand-children, other kins or household members not directly related to the household head were, as expected, less likely to have a matched case in the HDSS in both Mlomp and Bandafassi. Matching rates also varied significantly by age, but the pattern observed in Bandafassi was different from Mlomp. In Bandafassi, matching rates were slightly higher among adults aged 15-49 (76%) than among children aged less than 15 (74%), but older adults had a significantly higher chance of being matched. By contrast, in Mlomp, young adults were less likely to be matched than other age groups. The differences observed between Mlomp and Bandafassi can be ascribed to migration patterns: Mlomp has exceptionally high rates of temporary migrations. Rice cultivation is the main activity during the rainy season (from June to October). However, during the dry season (November to May), more than 40% of the adult population migrates to earn money through fishing, taking part in the production of palm wine, or working as temporary workers and domestic servants in the cities. Special efforts are made by HDSS interviewers to collect information on seasonal migrations and it is likely that most of them are identified as such. However, it is possible that some adults identified in the census could not be linked to the HDSS because they were not considered as resident in the HDSS, once the stricter rule of the census was used to define residency. Extending the search to individuals in migration could help to improve the linkage rates here.

	Bandafassi 2013				Mlomp 2002				Mlomp 2013			
	Freq	Prop. match.	OR[3]	Cis	Freq	Prop. match.	OR	Cis	Freq	Prop. match.	OR	Cis
Sex												
Males	6182	0.75	1		2593	0.87	1		2490	0.83	1	
Females	6226	0.75	1	(0.92-1.08)	2492	0.88	1.06	(0.9-1.25)	2488	0.84	1.08	(0.93-1.26)
Age group												
15-49	4896	0.76	1		615	0.85	1		562	0.88	1	
0-4	2045	0.69	0.68	(0.61-0.77)	1518	0.88	1.29	(0.98-1.69)	1377	0.85	0.75	(0.56-1)
5-14	3930	0.73	0.85	(0.77-0.93)	1905	0.87	1.19	(0.91-1.53)	2094	0.82	0.61	(0.46-0.8)
50-69	1218	0.84	1.69	(1.43-2.01)	756	0.92	2.12	(1.5-3.01)	585	0.87	0.88	(0.62-1.25)
70+	319	0.82	1.38	(1.04-1.86)	291	0.81	0.77	(0.54-1.12)	360	0.82	0.61	(0.42-0.89)
Relationship to head												
Head	1369	0.84	1		1049	0.89	1		1011	0.84	1	
Spouse	1768	0.85	1.05	(0.86-1.27)	586	0.93	1.57	(1.09-2.31)	563	0.9	1.67	(1.22-2.32)
Children (+in laws)	5996	0.8	0.76	(0.65-0.89)	1988	0.92	1.41	(1.09-1.82)	1956	0.88	1.38	(1.11-1.71)
Grand-children	1260	0.51	0.2	(0.16-0.23)	666	0.79	0.45	(0.34-0.59)	539	0.78	0.67	(0.52-0.88)
Other	1763	0.64	0.34	(0.29-0.4)	533	0.79	0.45	(0.34-0.6)	662	0.76	0.58	(0.45-0.74)
Unrelated	252	0.2	0.05	(0.03-0.07)	263	0.71	0.29	(0.21-0.41)	247	0.72	0.49	(0.35-0.67)

Table 2: Matching rates by background characteristics

b) Age heaping

The Myers' blended index (Swanson, Siegel and Shryock 2004) was used as a first attempt to evaluate the quality of reported ages. This index can be understood as the percentage of

---

3 Odds ratios presented here are univariate odds ratios obtained from a logistic regression.

cases which should be reclassified to obtain a smooth distribution of ages. (Here it can be computed for Niakhar too since it does not require linking cases at the individual level).

	2002			2013		
	Mlomp	Bandafassi	Niakhar	Mlomp	Bandafassi	Niakhar
HDSS	2.6	2.7	2.8	2.4	1.9	1.9
RGPH	7.7	16.1	19.0	5.3	6.5	8.8

Table 3: Values of Myers' blended index by site and data source

Table 3 clearly illustrates the significant improvement in the quality of data on ages from 2002 to 2013. It may be tempting to conclude that the use of PDAs contributed to this improvement, but this could also be due to a growing awareness of ages as the overall level of educational attainment is raised. The difference in data quality between the three HDSS is also striking, especially with Mlomp. Compared with the other two sites, the educational infrastructure is more developed in this HDSS: more than half of women aged 15-49 have already been to school for at least one year.

c) Systematic age misstatement

Table 3 present for Bandafassi and Mlomp the age differences between the HDSS and the censuses, according to the age reported in the census. We consider here only the population enumerated in the census which we could match to the HDSS. Overall, there is a distinct tendency to under-estimate the ages in the census among adolescents and adults. For example, as much as 30% of adults reported as being in their 20s have a true age which is 2 to 5 years higher according to the HDSS in Bandafassi in 2013. There is a better agreement between sources in Mlomp in 2002, but intriguingly the extent of under-statement seems more pronounced in 2013 in this HDSS.

	Age (from census)	HDSS < Census			=	HDSS > Census		
		[10,…]	[5-10[	[2-5[	[0-2[	[2-5[	[5-10[	[10,…]
Bandafassi 2013	0-9	0,0	0,6	3,4	73,1	12,9	4,1	5,9
	10-19	0,9	2,0	3,3	53,4	**26,3**	**10,0**	4,1
	20-29	3,4	2,1	4,9	36,6	**29,2**	**17,4**	6,5
	30-39	2,4	2,4	6,2	42,2	**22,7**	**15,8**	8,2
	40-49	3,8	3,8	6,3	32,0	**23,5**	**19,8**	10,9
	50-59	6,3	6,3	10,4	32,8	**16,8**	**17,5**	9,9
	60-69	11,9	8,9	10,6	20,8	**18,6**	**21,5**	7,7
	70+	**22,6**	**16,9**	**11,5**	21,4	12,8	9,9	4,9
Mlomp 2002	0-9	0,0	0,1	1,4	86,7	6,6	1,4	3,7
	10-19	0,3	1,4	2,9	62,9	**20,8**	**7,8**	3,8
	20-29	1,2	1,7	2,9	66,4	**16,7**	**8,7**	2,4
	30-39	1,5	1,2	3,5	51,8	**23,0**	**14,2**	4,8
	40-49	1,4	3,1	5,5	47,4	**17,3**	**17,0**	8,3
	50-59	4,7	6,5	6,8	38,0	**20,8**	**13,3**	10,0
	60-69	4,5	6,5	11,2	46,0	**17,5**	**8,2**	6,0
	70+	**8,6**	**15,9**	**13,4**	42,2	10,8	6,9	2,2

	0-9	0,0	0,0	0,9	87,5	7,1	1,4	3,1
	10-19	0,4	0,6	1,8	55,3	**30,0**	**9,7**	2,2
	20-29	2,4	1,9	1,7	46,5	**30,4**	**13,9**	3,2
	30-39	2,5	0,6	3,0	49,9	**25,2**	**15,2**	3,6
Mlomp 2013	40-49	2,8	2,0	1,4	45,5	**25,9**	**14,5**	8,0
	50-59	3,0	3,0	5,3	47,7	**19,2**	**13,2**	8,6
	60-69	6,0	6,0	6,5	40,1	**21,7**	**15,2**	4,6
	70+	9,9	7,8	13,4	46,6	14,1	7,1	1,1

Table 4: Distribution of age differences between HDSS and census, by age reported at census: Bandafassi and Mlomp.

It should be borne in mind that some extreme cases, such as when ages differ by more than 20 years or so, could correspond to mismatches, that is, cases where two records were inappropriately linked together. They will need to be examined in detail each time to establish whether it is indeed a mismatch or a real important difference between the ages, for reasons to be determined on a case-by-case basis.

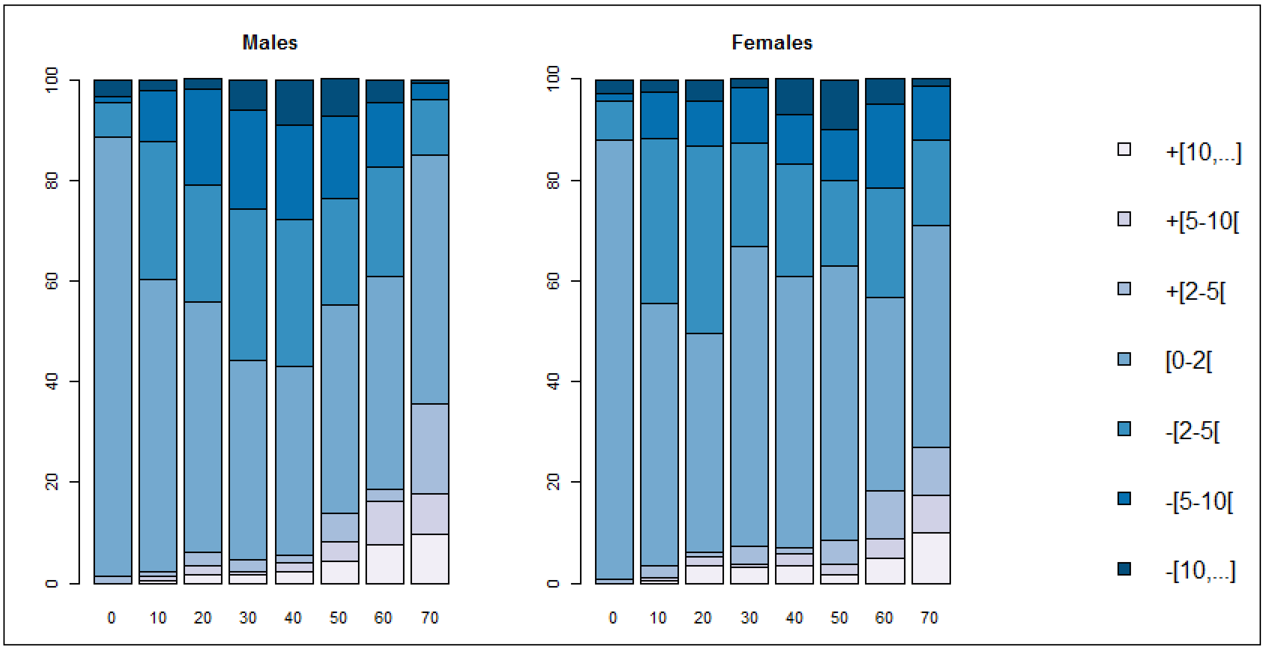

Figure 1: Distribution of age differences between HDSS and 2013 census in Mlomp, by age reported at census.

Such tendency to under-estimate the ages in the censuses are likely to distort demographic estimates. For example, they will result in younger age distributions of recent births, introducing biases in age-specific fertility rates which result in a fertility calendar younger than the true one as illustrated for Mlomp in 2002 (figure 2) and 2013 (figure 3).

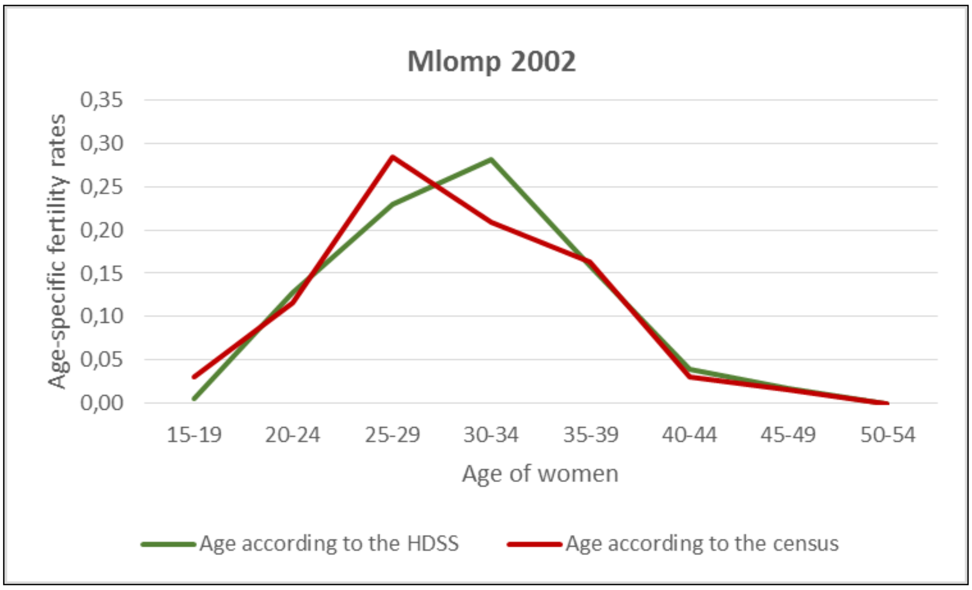

Figure 2: Age-specific fertility rates according to the census or the HDSS. Mlomp 2002

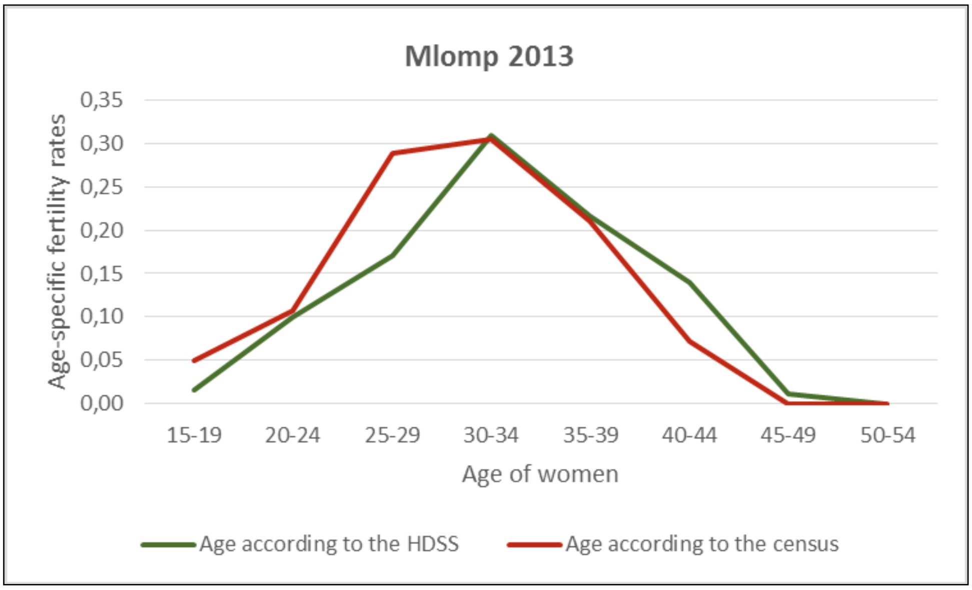

Figure 3: Age-specific fertility rates according to the census or the HDSS. Mlomp 2013.

In most developing countries, reports on the number children ever born and surviving are also used to infer indirect estimates of child mortality, using the Brass method. The proportion of deceased children is converted into probabilities of dying, and the age of the mother is used to approximate the ages of children. Since the proportions of children who died before the census increases with age, any under-statement of ages will result in an over-estimation of child mortality, because reports of deceased children will be transferred down to younger ages. This is illustrated in figure 4 for Bandafassi in 2013, using the two sources of data on ages. Errors are amplified is age misreporting varies by parity.

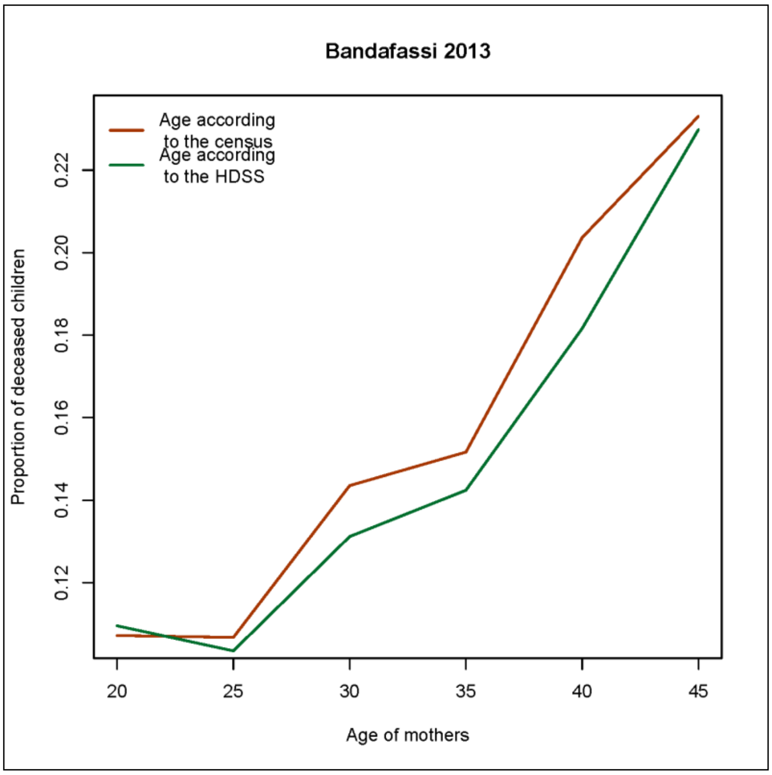

Figure 4: Proportion of deceased children by age-group of the mother according to the census or the HDSS. Bandafassi, 2013 census.

Similarly, another Brass-type approach consists in estimating adult mortality from reports on orphanhood. Again, under-stating the ages will result in upward biases in mortality rates estimated indirectly.

Note: As mentionned above, the final paper will expand from this extended abstract and provide more information on the study method, detailed analyses using the whole dataset including the data from the Niakhar site, and will examine in details the consequences of age misreporting on demographic estimates in the studied populations.

Références

Blanc, R. 1962. *Manuel de recherche démographique en pays sous-développé*. Paris: Presses universitaires de France.

Delaunay, V., L. Douillot, A. Diallo, D. Dione, J.-F. Trape, O. Medianikov, D. Raoult, and C. Sokhna. 2013. "Profile: the Niakhar Health and Demographic Surveillance System." *International Journal of Epidemiology* 42(4):1002-1011.

Ewbank, D.C. 1981. "Age misreporting and age-selective underenumeration: sources patterns and consequences for demographic analysis."

Ferry, B. 1976. "Le fichier-événement: une nouvelle méthode d'observation rétrospective." Pp. 137-163 in *L'observation démographique dans les pays à statistiques déficientes. Chaire Quételet 1976.*: Université Catholique de Louvain.

Freedman, D., A. Thornton, D. Camburn, D. Alwin, and L. Young-DeMarco. 1988. "The life history calendar: A technique for collecting retrospective data." *Sociological methodology* 18(1):37-68.

Gendreau, F.and R. Nadot. 1967. *Afrique noire, Madagascar, Comores... démographie comparée. 9-10. Structures par âge, actuelle et future*. Paris: Délégation générale à la recherche scientifique et technique.

Gibril, M. 1979. "Evaluating census response errors: a case study for the Gambia." *Development Centre Studies*.

Gubry, P. 1975. "Une confrontation entre deux méthodes de détermination des âges au Cameroun." *Cahiers ORSTOM, Série Sciences Humaines* 12(4):325-337.

Helleringer, S., G. Pison, B. Masquelier, A.M. Kanté, L. Douillot, G. Duthé, C. Sokhna, and V. Delaunay. 2014. "Improving the Quality of Adult Mortality Data Collected in Demographic Surveys: Validation Study of a New Siblings' Survival Questionnaire in Niakhar, Senegal." *PLoS Med* 11(5):e1001652.

Hertrich, V.and S. Lardoux. 2014. "Áge ā l'entrée en union des femmes en Afrique. Les données des enquȩtes et des recensements sont-elles comparables?" *Population*(3):399-432.

Howell, N. 1979. *Demography of the Dobe! Kung*. New York: Academic Press.

Jaro, M.A. 1989. "Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida." *Journal of the American Statistical Association* 84(406):414-420.

Pison, G. 1979. "Age déclaré et âge réel : une mesure des erreurs sur l'âge en l'absence d'état civil." *Population* 34(3):637-648.

—. 1980. "Calculer l'âge sans le demander. Méthode d'estimation de l'âge et structure par âge des Peul Bandé (Sénégal Oriental)." *Population* 35(4-5):861-892.

Pison, G., L. Douillot, A.M. Kante, O. Ndiaye, P.N. Diouf, P. Senghor, C. Sokhna, and V. Delaunay. 2014. "Health & Demographic Surveillance System Profile: Bandafassi Health and Demographic Surveillance System (Bandafassi HDSS), Senegal." *International Journal of Epidemiology*.

Pison, G., A. Gabadinho, A. Wade, and C. Enel. 2002. "Mlomp demographic surveillance system, Senegal." Pp. 271-278 in *Population and health in developing countries : Population health and survival at INDEPTH sites*, edited by INDEPTH. Ottawa: International development research centre.

Pison, G., K. Hill, B. Cohen, and K. Foote. 1995. *Population dynamics of Senegal*. Washington, DC: National Academy Press.

Pison, G.and P.O. Ohadike. 2006. "Errors and manipulations in age assessment." Pp. 313-336 in *Human Clocks: The Bio-Cultural Meaning of Age*, edited by C. Sauvain-Dugerdil, H. Leridon, and N. Mascie-Taylor. Berne: Peter Lang.

Pullum, T.W. 2006. "An assessment of age and date reporting in the DHS Surveys 1985-2003."

Scott, C.and G. Sabagh. 1970. "The historical calendar as a method of estimating age: The experience of the Moroccan multi-purpose sample survey of 1961–63." *Population Studies* 24(1):93-109.

Spoorenberg, T.and C. Dutreuilh. 2007. "Quality of age reporting: extension and application of the modified Whipple's index." *Population (english edition)*:729-741.

Swanson, D., J.S. Siegel, and H.S. Shryock. 2004. *The methods and materials of demography*: Elsevier.

Winkler, W.E. 1999. "The state of record linkage and current research problems." Presented at Statistical Research Division, US Census Bureau.