Using Parish Registers for Demographic Research in African Settings:

A Record Linkage Study on Age Data Quality in Ouagadougou

Ouedraogo Soumaila[*1], Compaoré Yacouba[2], Lankoandé Bruno[2], and Helleringer Stephane[1]

[1]New York University Abu Dhabi
[2]Institut Superieur des Sciences de la Population

February 12, 2024

Abstract

Accurate age data are essential for studying aging in African countries, particularly in a comparative perspective. Unfortunately, data on age are often affected by large errors, particularly at older ages, notably in the form of age heaping or exaggerations. Without verification and validation, these errors are likely to bias age-related demographics. In this manuscript, we tested the feasibility of verifying the age of older adults through linkages with parish registers in Ouagadougou, the capital of Burkina-Faso, a low-income country in West Africa.

Keywords: *Age data, linkage, parish registers, Africa*

[*]Corresponding author: so2276@nyu.edu

1 Background

Sub-Saharan Africa is a region in transition from a demographic perspective, with a very young population. According to the latest World Population Prospects, only 4% of sub-Saharan adults were aged 60 or over (Nations, 2022). However, the slow transformation of its age structure hides the rapid growth in the older adult population (Masquelier et al., 2017), which could triple over the next thirty years (Nations, 2022). These statistics point to a silent ageing of the population. Not only do they alert about future implications in terms of social policies and social protection systems. They also call for monitoring this demographic ageing that is underway in African populations.

Aging is an important component that help understanding demographic transition (Fernandes et al., 2023). In view of the multiple implications that may arise from this, particularly in relation to health, the World Health Organization, in a forward-looking statement in 2015, called for urgent actions to tackle measurement issues, and improve monitoring and knowledge about the health of older adults (WHO, 2015). Most demographic indicators are age-related, and age is one of the main risk factors for many diseases. To this end, accurate data on age is essential to studying aging in African countries, especially in a comparative perspective. Similarly, inaccurate ages limit our ability to identify individuals most at risk from aging-related diseases, and complicate targeting prevention and intervention (Fransquet et al., 2019). Unfortunately, Civil Registration and Vital Statistics systems are partial, deficient and incomplete in most sub-Saharan African countries. Age data produced in the region mainly comes from censuses and large-scale household surveys. Given the declarative nature of these data, age error issues commonly encountered in such data sources are well known. Both digit preferences and age exaggerations have obvious drawbacks for demographic estimation. According to the literature, age exaggerations are known to be particularly pronounced at older ages (Dechter and Preston, 1991; Palloni et al., 2016; Preston et al., 1996, 1999). When it comes to ages at death, these exaggerations lead to higher than expected survival rates and result in an underestimation of mortality.

Different strategies have been developed to tackle poor data quality on ages among older adults. Addressing racial crossover in the study of African-American mortality, Preston et al. (1996) found inconsistencies between ages on death certificates and ages recorded in early censuses and Social Security records. Likewise, Palloni et al. (2016, 2021) used an integrated method based on simulations using Costa Rica as a standard to correct for age misreporting and adjust older adult mortality estimates in Latin America. Rather than resorting to statistical corrections and adjustments, new methods for age ascertainment have being developed, building on technological innovations. As an illustration, ranking methods from machine learning algorithm have been proposed so that to project older person with higher rank value than the younger one based on face image (Onifade and Akinyemi, 2015; Cao et al., 2012). In the same vein, Helleringer et al. (2019) ascertained ages of women of childbearing age using computer vision algorithm from machine learning technique. Besides, epigenetic biomarkers of aging have being used as epigenetic clock based on the measurement of DNA methylation to assess chronological age and study longevity (Dec et al., 2023; Bernabeu et al., 2023).

Regardless of the strategy used, age verification requires confirmed ages for a sample of individuals within the population of interest. With technology, reliable images of older adults are needed to train the algorithm for computer vision, just as the epigenetic clocks needs a reliable sample of individuals of known to accurately predict chronological age. Regarding linkage with administrative and historical registration records, including early censuses, such records are rare in sub-Saharan Africa. When they do exist, they do not have a long track record that could include birth records of older adults at the time of their births. The only historical sources capable of containing such information are parish registers. They existed only in a few areas where the headquarters of the colonizers were established. Examining parish registers from East and Central Africa, Walters (2016) recalled the use of parish registers for family reconstruction to church records. She also highlighted the potential of these registers in Africa to describe what happened to individual agency from a historical demography perspective.

Beyond these issues, we proposed to use these records as an alternative source for verifying the age of older adults. In this paper, we tested the feasibility of verifying the ages of older adults through linkages with parish registers in Ouagadougou, the capital city of Burkina Faso, a low-income country in West Africa.

2 Data and Methods

2.1 Study setting

According to the latest population census carried out in 2019, Catholics accounted for around 20% of the population. This proportion remained almost the same as in 2006. In the capital Ouagadougou, the share of Catholics is higher, at approximately 30% (INSD, 2022). The country is home to more than 60 ethnic groups, which can be grouped into several broad categories based on cultural proximity and similar religious beliefs. The distinctive religious feature of the mossi group, largely located in the center of the country, lies in their strong belief in the invisible and the supernatural, and the profound veneration of their ancestors. This predisposition made them potentially permeable to the evangelization message of the White Fathers missionaries. After an initial mission in the Center-East of the country, which led to the establishment of the parish of Koupela in 1900, it was finally in 1901 that the plateau mossi, of which Ouagadougou is the epicenter, received its first White Father missionaries. The mission founded the Cathedral as the first parish of what would later become the archdiocese of Ouagadougou. They recruited and instructed the first catechumens and recorded the first baptisms as early as 1905. Evangelization efforts were crowned by the erection of Ouagadougou as a diocese in 1921(Tiendrebeogo, 2008). This was followed by the expansion of Catholicism with the creation of new parishes in and around Ouagadougou. By December 2022, the archdiocese of Ouagadougou had 34 parishes, in which baptism registers for children and adults, as well as confirmation, marriage and death registers, have been regularly kept since their creation. In their organization, a duplicate of each register recorded in parishes is forwarded to the archdiocese to facilitate follow-up. With some registers lost, stolen or damaged in parishes, this strategy has made it possible to save a large number.

2.2 Methodological approach

Sampling and survey implementation

In this study, we focused on Catholics aged 60 and over. By opting to start from the age of 60, we ensured that we covered the bulk of the older adult population among Catholics. We used data from the Ouagadougou Health and Demographic Surveillance System (OHDSS) as sampling frame. The OHDSS has been established in 2008 by the Institut Supérieur des Sciences de la Population (ISSP), in which around 100,000 individuals were under surveillance in 2019. As selection criteria, we targeted older Catholics aged 60 or over living in one of the five districts covered by the OHDSS. Assuming that most baptisms take place at the place of birth of the baptized individual, we prioritized first those who declared that they were born in Ouagadougou (category 1), followed by those who were born in other places in Burkina Faso (category 2). We did not find it useful to add to these two categories those who were born outside Burkina Faso. Out of a population of 2059 older Catholics living in the OHDSS and aged 60 or over, 262 had stated to have been born in Ouagadougou, of whom 238 were aged 60-79 and only 24 were aged 80 or over. We therefore selected all the 262 individuals. The only people left were those born in other places in Burkina Faso, a total of 1,095 of whom 103 were aged 80 or over. To compensate for the low proportion of people aged 80 or over in category 1, we opted to consider all the remaining 103 aged 80 or over, bringing our sample to a total of 365. Table 1 below shows the final composition of the sample, disaggregated by age group within the 5 districts of the area.

Table 1: Sampling allocation

Districts	District status	Sample			OHDSS		
		60-79 years	80 years +	Total	60-79 years	80 years +	Total
KILWIN	Formal	44	38	82	544	63	607
NIOKO2	Informal	59	15	74	233	28	261
NONGHIN	Informal	9	16	25	332	43	375
POLESGO	Informal	19	6	25	107	8	115
TANGHIN	Formal	107	52	159	611	90	701
Total		**238**	**127**	**365**	**1827**	**232**	**2059**

Source : Tabulation compiled on 15/09/2022 using the OPO data, Round 11

We developed a questionnaire and displayed it on tablets to carry out interviews with the sampled individuals. We collected a range of information about both the individual and the household to which they belonged. At the individual level, notwithstanding the age data from the OHDSS, we collected the dates of birth and ages in two different questions to ensure consistency between stated ages and dates of birth. When the stated age did not match the stated date of birth, a question was asked to ascertain whether the error stemmed from the stated date of birth or the stated age, in order to make the necessary correction in accordance with the respondent. In addition to age data, we recorded individual socio-demographic characteristics such as gender and level of education as well as parent names for the linkage purpose. Then, individuals were asked if they had a baptismal booklet, in which case they were asked to see the booklet in order to collect the year and parish of baptism, as well as the number of the baptismal record. Even if they said to have no booklet, they were still asked if they had been baptised and the details

of the baptism were collected. We also collected information on people who stated to have been married in church. This ensured that we had at hand the required information to search for individuals in the appropriate registers. When baptismal information failed to find an individual, marriage information were used to find them. Consequently, individuals who were not baptized in Ouagadougou can also be found when they were married in Ouagadougou.

Linkage strategy

From the survey, 311 out of 365 from the sample were interviewed. Although they were all supposed to be at least 60 years old according to the ages stated in the OHDSS, we nevertheless found five individuals who did not meet this criterion to be interviewed. In addition, 6 other people who claimed to be Christians had never been baptized as such. This brings us to a total of 300 baptized from whom we had obtained written consent to search and verify their records on parish books. For each of these individuals, we had information on the year and parish of baptism and/or marriage. This made it possible to identify the appropriate baptism book in which to look for the individual details. If the individual had a baptismal number, mainly for those who had their baptismal booklet, all we had to do was find the number in the parish book and verify the concordance of the name and the parents names. This was made simpler with the help of the archivists at the archdiocese of Ouagadougou. For people for whom we did not have the baptismal numbers but we did have information on the year and the baptismal register, we had to go through the register number by number using different criteria: the name and the name of the parents. When a person was found, their details were recorded in a file and their references photographed. For various reasons (neither baptism nor marriage registered in Ouagadougou, registers damaged, or erroneous information on year and parish of registration), 248 were found in the parish registers, giving a matching rate of 83%.

3 Results

3.1 Statistical association with eligibility for linkage

The analysis of the statistical associations between the eligibility for linkage and individual characteristics is portrayed in Table 2 below. From this summary, it appears that the proportion of linked individuals differs statistically and tends to decrease with age, falling from 92% in the < 65 age group to 77% in the ≥ 85 age group. Otherwise, it is more difficult to link older generations than younger ones. Not surprisingly, there was also a statistical difference between places of baptism in terms of linkage status, as those baptized in Ouagadougou were more likely to be linked (94%). The statistical association with marital status was somewhat weak, with a probability of 0.055. Indeed, 84% of ever-married eligibles were linked, compared to 62% among the others (single and cohabiting). However, no statistical evidence could be established for the variables gender, educational attainment and self-assessment of reliability of dates of birth (DOB) on national IDs. Admittedly, slight differences were noted. For example, 84% of eligible women were linked,

compared with 81% of eligible men. The proportion of linked eligibles among non-schooling and those with primary education was 82%, compared with 85% among those with secondary education or higher. Regardless of the category into which respondents were classified according to self-assessment of the reliability of dates of birth on their IDs, an almost equally proportion were linked to the registers. The differences observed between the different categories of these variables were not marked enough to denote statistical associations.

Table 2 : linkage status by eligible characteristics

Characteristic	Overall N = 300[1]	Linked N = 248[2]	Not linked N = 52[2]	p-value[3]
Gender				0.366
Men	128	104 (81%)	24 (19%)	
Women	172	144 (84%)	28 (16%)	
Age group (self-reported)				<0.001
<65y	90	83 (92%)	7 (8%)	
65-74y	99	89 (90%)	10 (10%)	
75-84y	81	53 (65%)	28 (35%)	
≥85y	30	23 (77%)	7 (23%)	
DOB on ID card				0.755
DOB accurate	209	170 (81%)	39 (19%)	
DOB inaccurate	63	54 (86%)	9 (14%)	
No ID card	28	24 (86%)	4 (14%)	
Marital status				0.055
Ever-married	287	240 (84%)	47 (16%)	
Never married	13	8 (62%)	5 (38%)	
Educational attainment				0.912
No schooling	188	155 (82%)	33 (18%)	
Primary	72	59 (82%)	13 (18%)	
Secondary & higher	40	34 (85%)	6 (15%)	
Place of baptism				<0.001
Ouagadougou	250	236 (94%)	14 (6%)	
Outside of Ouagadougou	50	12 (24%)	38 (76%)	

[1]n
[2]n (%)
[3]Pearson's Chi-squared test; Fisher's exact test

3.2 Change in age at baptism

Looking at the cumulative distribution curves of ages at baptism for linked individuals (see Figure 1 below), there are clearly differences between their age groups. Among people aged < 65, almost 40% were baptized before their first birthday. This proportion falls sharply to almost 15% in the 65-74 age group, and to around 11% in the 75-84 age group. Among those in the ≥ 85 age group, the first baptisms occurred from age 15 onwards. A closer examination of the curves shows that the medians age at baptism of these four distributions are 12 years, 15 years, 23 years and 47 years for the < 65, 65-74, 75-84 and ≥ 85 age groups respectively.

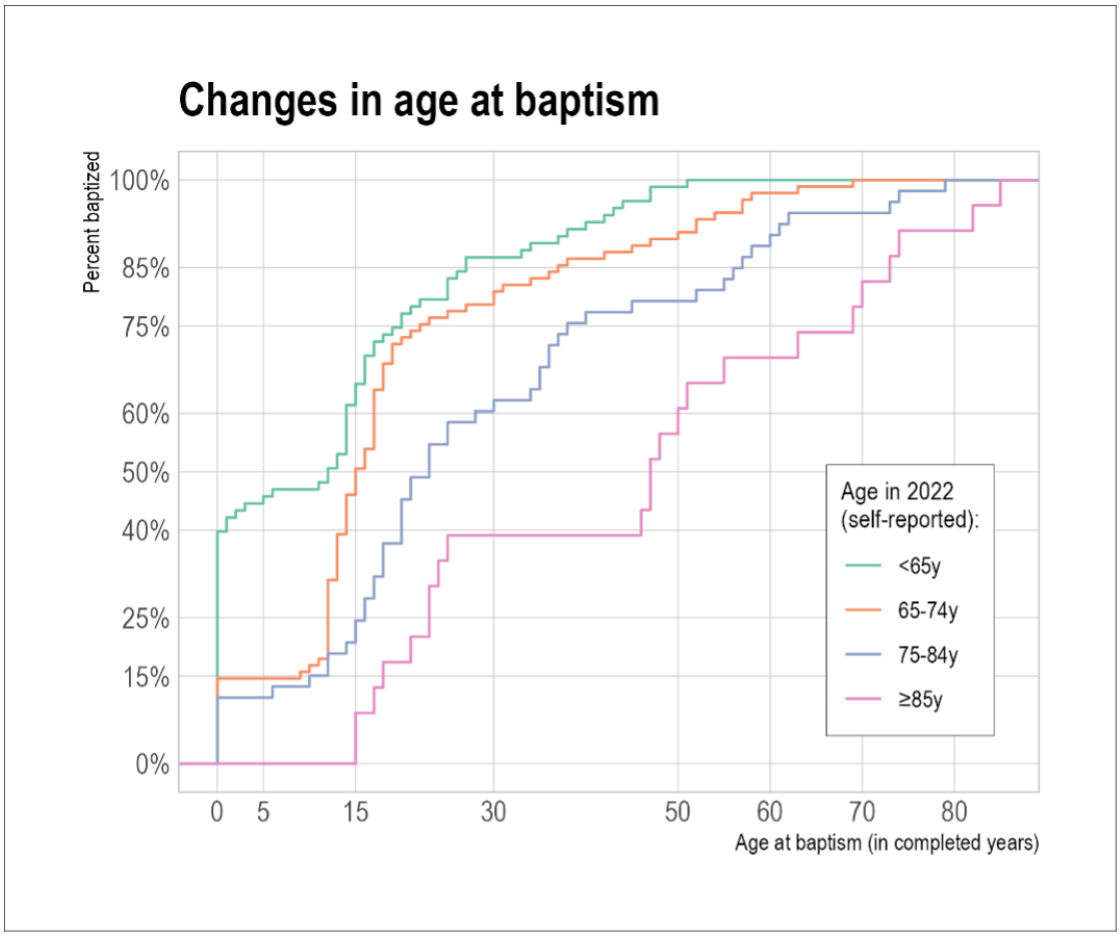

Figure 1: Change in age at baptism

3.3 Reliability of ages at baptism

In Figure 2 below, we observe the distribution of differences between self-reported ages and ages in registers for different age groups from baptism registers. It appears clearly that people who were baptized as children (0-4 years) tend to report ages that are consistent with those in their baptismal records. The distribution of age differences looks normal, peaked and centered around 0. Among those baptized between ages 5-14, the distribution of age differences is also normal and centered around 0, but less peaked than the previous one. This indicates greater variability around 0, although the ages are consistent for the maximum number of individuals in this group. For the young adult age groups (15-49), the distribution of age differences is less regular, with two modal values, the highest of which is 0 and the second -1. Compared to the value of 0, we can see that this distribution is skewed to the left, i.e. towards negative values. This implies that many individuals in this group overstate their age in comparison with that recorded in the registers. In the last group, notably those aged 50 or over, we observe the disappearance of the modal value at 0 and the full constitution of the modal value around -1. The clear shift of the curve below 0 highlights the exaggeration of ages at these ages. Even though

most of the shift remains between -5 and 0, there are still gaps of more or less 10 years.

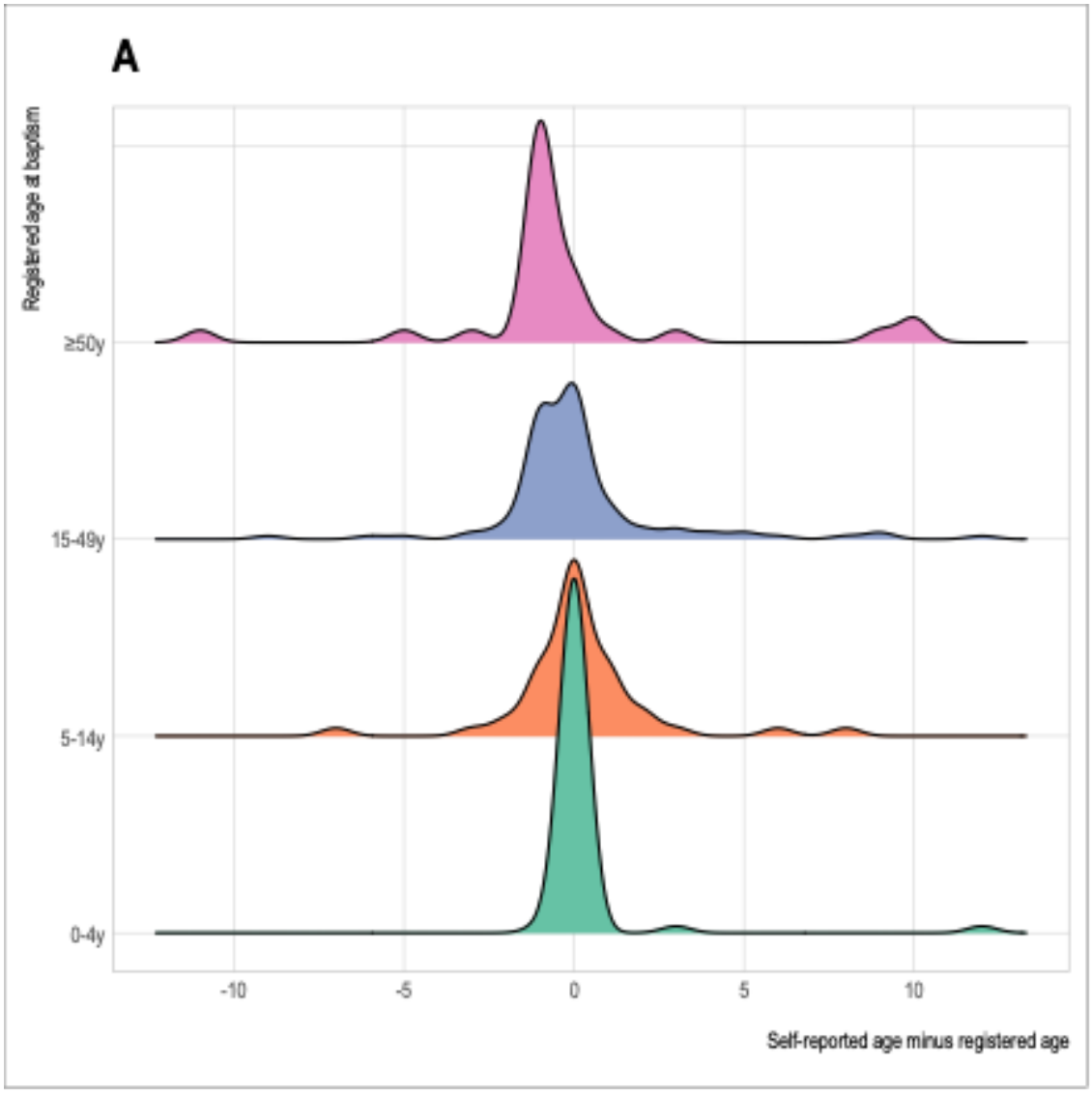

Figure 2: Reliability of age at baptism

4 Conclusion

The present study addresses this age reporting quality in older adults, whether self-reported, from IDs or even from other sources such as parish registers. However, we have seen that dates of birth recorded at baptisms in childhood look more reliable than those recorded at baptisms in adulthood. Regarding the analysis of ages at baptism, the fact that there is a high proportion of 0-year baptisms among the ¡65 age group compared with the older age groups is suggestive of a change in the population of 0-year baptisms. This change could result in a greater increase in the rate of 0-year-olds baptized in the future, including among Catholics over 65. If

we assume that those who are baptized before their first birthday, or even under 5 years of age, are more likely to have more accurate ages or birthdates recorded at the time of their baptism in the registers, this could help to improve the quality of age reporting. This is certainly the rationale behind the fact that in the group of baptized individuals over age 15, there was less concordance between the self-reported ages and the ages in the registers, while the concordance was good for those baptized under age 5.

References

Bernabeu, E., McCartney, D. L., Gadd, D. A., Hillary, R. F., Lu, A. T., Murphy, L., Wrobel, N., Campbell, A., Harris, S. E., Liewald, D., Hayward, C., Sudlow, C., Cox, S. R., Evans, K. L., Horvath, S., McIntosh, A. M., Robinson, M. R., Vallejos, C. A., and Marioni, R. E. (2023). Refining epigenetic prediction of chronological and biological age. *Genome Medicine*, 15(1):12.

Cao, D., Lei, Z., Zhang, Z., Feng, J., and Li, S. Z. (2012). Human Age Estimation Using Ranking SVM. In Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., Zheng, W.-S., Sun, Z., Wang, Y., Chen, X., Yuen, P. C., and Lai, J., editors, *Biometric Recognition*, volume 7701, pages 324–331. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Lecture Notes in Computer Science.

Dec, E., Clement, J., Cheng, K., Church, G. M., Fossel, M. B., Rehkopf, D. H., Rosero-Bixby, L., Kobor, M. S., Lin, D. T., Lu, A. T., Fei, Z., Guo, W., Chew, Y. C., Yang, X., Putra, S. E. D., Reiner, A. P., Correa, A., Vilalta, A., Pirazzini, C., Passarino, G., Monti, D., Arosio, B., Garagnani, P., Franceschi, C., and Horvath, S. (2023). Centenarian clocks: epigenetic clocks for validating claims of exceptional longevity. *GeroScience*, 45(3):1817–1835.

Dechter, A. R. and Preston, S. H. (1991). Age misreporting and its effects on adult mortality estimates in Latin America. *Population Bulletin of the United Nations*, (31-32):1–16.

Fernandes, F., Turra, C. M., and Rios-Neto, E. L. (2023). World population aging as a function of period demographic conditions. *Demographic Research*, 48:353–372. Publisher: JSTOR.

Fransquet, P. D., Wrigglesworth, J., Woods, R. L., Ernst, M. E., and Ryan, J. (2019). The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis. *Clinical Epigenetics*, 11(1):62.

Helleringer, S., You, C., Fleury, L., Douillot, L., Diouf, I., Ndiaye, C. T., Delaunay, V., and Vidal, R. (2019). Improving age measurement in low-and middle-income countries through computer vision. *Demographic Research*, 40:219–260. Publisher: JSTOR.

INSD (2022). Monographie de la commune de Ouagadougou. Monographies régionales et communales, Institut National de la Statistique et de la Démographie, Ouagadougou, Burkina Faso.

Masquelier, B., Eaton, J. W., Gerland, P., Pelletier, F., and Mutai, K. K. (2017). Age patterns and sex ratios of adult mortality in countries with high HIV prevalence. *AIDS*, 31(1):S77–S85.

Nations, U. (2022). World Population Prospects 2022, Online Edition. Technical report, Department of Economic and Social Affairs, Population Division.

Onifade, O. F. and Akinyemi, D. J. (2015). GWAgeER-A GroupWise Age Ranking Framework for Human Age Estimation. *International Journal of Image, Graphics and Signal Processing*, 7(5):1. Publisher: Citeseer.

Palloni, A., Beltrán-Sánchez, H., and Pinto, G. (2021). Estimation of older-adult mortality from information distorted by systematic age misreporting. *Population Studies*, 75(3):403–420. Publisher: Routledge _eprint: https://doi.org/10.1080/00324728.2021.1918752.

Palloni, A., Pinto, G., and Beltrán-Sánchez, H. (2016). Estimation of life tables in the Latin American Data Base (LAMBdA): Adjustments for relative completeness and age misreporting. In *United Nations Expert Group Meeting on the methodology and Lessons learned to evaluate the completeness and quality of vital statistics data from civil registration. Retrieved February*, volume 21, page 2021.

Preston, S. H., Elo, I. T., Rosenwaike, I., and Hill, M. (1996). African-American mortality at older ages: Results of a matching study. *Demography*, 33(2):193–209.

Preston, S. H., Elo, I. T., and Stewart, Q. (1999). Effects of age misreporting on mortality estimates at older ages. *Population studies*, 53(2):165–177.

Tiendrebeogo, W. M. (2008). *Le développement de l'Église-Famille de Dieu au Burkina Faso: le cas de l'archidiocèse de Ouagadougou*. M.A dissertation, Université Laval Canada.

Walters, S. (2016). Counting Souls: Towards an historical demography of Africa. *Demographic research*, 34:63–108.

WHO (2015). *World report on ageing and health*. World Health Organization, Geneva. Section: 246 p.