Assessing the validity of mobile phone surveys for mortality data collection: a case study from Ouagadougou, Burkina Faso.

Kassoum Dianou[1,2], Bruno Masquelier[1], Bruno Lankoandé[2], Abdramane Soura[2], Hervé Bassinga[2], Ashira Menashe Oren[1], Malebogo Tlhajoane[3], Georges Reniers[3]

Abstract

Mobile Phone Surveys (MPS) are increasingly being used, however, their comparability and validity in collecting mortality data is unclear. This study quantifies the impact of errors related to age, dating, and omissions in mortality rates from MPS. We conducted a validation study in the Ouagadougou Demographic and Health Surveillance System (OHDSS), interviewing adults over the phone, for whom we had high-quality data on the survival of their children, siblings, and parents. Individual records were matched, and we assessed the sensitivity and accuracy of mobile phone interviews in recording vital events, ages and dates. Our findings revealed that men tended to overestimate their age in MPS, while women tended to underestimate it. Our results also showed that data on the survival of children, siblings, and parents collected during MPS exhibited moderate specificity but high sensitivity in recording deaths. We applied an imputation method to analyze biases in survey estimates of child and adult mortality. Reporting errors in MPS data did not significantly impact estimates of under-five mortality rates but led to an underestimation of adult mortality. This study sheds light on the complex error patterns associated with MPS, emphasizing the need for thorough examination before considering their broader implementation.

[1] Center for Demographic Research, University of Louvain, Belgium
[2] Institut Supérieur des Sciences de la Population, Burkina Faso
[2] The London School of Hygiene & Tropical Medicine, UK

Introduction

In 2015, all United Nations member states unanimously embraced the 17 Sustainable Development Goals (SDGs) as part of the agenda for Sustainable Development by 2030. Among these objectives, Goal 3 stands out, dedicated to advancing well-being across all age groups. This ambitious goal delineates specific targets, such as reducing neonatal and under-5 mortality rates to a maximum of 12 and 25 per 1,000 live births, respectively. Additionally, it commits to lowering maternal mortality to fewer than 70 deaths per thousand women by 2030 (UN, 2023).

Achieving this goal necessitates precise data on mortality rates and trends across all age groups (Lawn et al., 2014; Moxon et al., 2015). These data play a pivotal role in refining the focus of interventions aimed at decreasing neonatal, adult, and elderly mortality. They also serve in monitoring the repercussions of alterations in the coverage of diverse interventions (Davey et al., 2017).

While developed countries monitor mortality through robust civil registration systems, low- or middle-income nations grapple with deficient civil registration. Estimation and monitoring in these regions rely on data from retrospective surveys, censuses, or projections issued by UN agencies. This challenge is particularly pronounced in most low- or middle-income countries, where civil registration rates are notably low, with a few exceptions (Mikkelsen et al., 2015). In these regions, a significant number of deaths occur at home, often going unreported to civil registration agencies. Moreover, deaths are frequently inadequately recorded, as underscored by multiple studies (Amo-Adjei & Annim, 2015; Fisker et al., 2019; Mikkelsen et al., 2015).

Traditional face-to-face surveys, renowned for their structured, yet flexible, approach, have historically served as the primary source of demographic and mortality data in these nations. Face-to-face surveys were considered superior to other remote methods (telephone, internet, web, etc.), benefitting from personal interaction and the ability to control survey environments. Despite their advantages, high costs, extended durations, and potential impracticality during crises prompted researchers and practitioners to seek alternative methods for swiftly collecting demographic information. The emergence of mobile phone surveys, leveraging the rapid expansion of cellular network coverage across the world, sparked enthusiasm for their potential to enhance data collection efficiency and feasibility during crises. Consequently, amid the COVID-19 and Ebola crises, coupled with associated containment measures, several mobile phone surveys were conducted in low- or middle-income countries to furnish updated data and insights into mortality trends (Zezza et al., 2021).

Existing literature reveals numerous studies comparing face-to-face surveys and phone surveys across various fields and countries, highlighting significant differences between the two methods (Lamanna et al., 2019; Mahfoud et al., 2015). However, these comparative analyses predominantly operate at aggregated levels, leaving a notable gap in individual-level insights.

For monitoring mortality trends, low- and middle-income countries predominantly rely on inquiries regarding the survival of close relatives (parents, siblings, and children) or recent household deaths, collected during retrospective surveys or censuses (Hill, 2013). Yet, utilizing information from these inquiries for mortality surveillance encounters two primary challenges, irrespective of the collection method: omission errors and reporting errors. Each type of error impacts mortality estimation uniquely in terms of significance and intensity. For instance, a failure to report a death can result in underestimating mortality, while errors in the ages of survivors and deceased individuals distort age-specific mortality patterns and affect aggregate indicators.

While numerous studies have scrutinized the impact of age, dating, and omission errors on mortality estimation in face-to-face surveys (Helleringer et al., 2014; Lankoandé et al., 2022), there is a notable gap in research specifically assessing the quality of individual-level data from phone surveys or examining the impact of age, dating, or omission errors in these surveys on the estimation of mortality across all ages.

Data and Methodology

To establish robust benchmarks for our comparative analysis, we leveraged data from the Ouagadougou Demographic and Health Surveillance System (OHDSS), operational since 2008. This system involves systematic data collection on crucial life events, including births, deaths, and migrations, encompassing all residents within a well-defined geographic area post the baseline census (Pison, 2005). Recognized for its data quality, population observatories have been employed as standards in other studies, evaluating the reliability of mortality indicators derived from alternative sources (Helleringer et al., 2014; Rerimoi et al., 2019). The OHDSS compiles detailed information on all residents in the specified area.

Furthermore, we utilized data from Burkina Faso's most recent general population and housing census conducted in 2019, capturing information on parental survival and deaths within the past twelve months. Additionally, we integrated data from Burkina Faso's latest Demographic and Health Survey conducted in 2021, documenting the reproductive history of women of reproductive age and the comprehensive survival history of siblings.

For the telephone survey component, we relied on data collected during the RaMMPS project's validation survey, conducted via phone from January to March 2022 with the same individuals tracked in the OHDSS. A sample of 3,517 individuals aged 18 to 64, residing in the OHDSS, was selected. Among them, 2,977 individuals were successfully contacted, and interviews were conducted with 2,777 people. Respondents answered questions about their personal characteristics and the survival of their close relatives (parents, siblings). Additionally, women of reproductive age were interviewed about their reproductive history over the last seven years. We deliberately oversampled women of reproductive age who experienced the death of a child under five, as well as individuals who faced the death of a sibling or parent in the last five years.

For the individual-level analysis, we applied both probabilistic and deterministic techniques to match households or individuals between the two data sources (Telephone Survey and OHDSS) (Felderer et al., 2019; Masquelier et al., 2021). Initial scrutiny involved examining age errors and omissions in matched individuals through descriptive statistics. Subsequently, for each individual, the age collected via the telephone survey was substituted with the age obtained face-to-face before re-evaluating mortality levels using telephone survey data. This approach aimed to quantify the contribution of age errors to observed mortality discrepancies. Concerning death omissions, sensitivity and specificity analyses, alongside Cohen's Kappa statistical tests, were employed to assess the agreement of information provided between the two collection methods. Finally, error patterns in ages and omissions were replicated through imputation on national data (census and DHS) to evaluate the impact of each error type on mortality estimates resulting from telephone surveys.

Results

Our analysis reveals high sensitivity and specificity in recording vital events through telephone surveys compared to OHDSS (Table 1). The degree of sensitivity and specificity varies depending on the type of event considered, with higher accuracy observed for recording birth history of women of reproductive age and the survival of respondents' parents compared to reports on the survival of siblings (Table 1). Furthermore, Cohen's Kappa tests indicate values ranging between 0.6 $<k<0.9$, suggesting almost perfect agreement between information from respondents in both collection modes (Table 2). In other words, no significant difference is observed in the quality of information provided by respondents via phone or face-to-face.

Our analyses on the reliability and concordance of information provided by respondents regarding their own ages or those of their close relatives (parents, siblings, and children) alive or deceased demonstrate a strong correlation (Table 3, Figur 1 and Figure 2). However, it is crucial to emphasize that the degree of correlation varies depending on the socio-demographic characteristics of respondents, increasing with the level of education and varying by place of residence and where the respondent answers the questions with the mobile phone survey (Table 3).

The assessment of the impact of age, dating, or omission errors on mortality estimation indicates that: (i) age errors lead to an underestimation of adult mortality without significantly affecting child mortality estimation (Figure 3 and Figure 4); (ii) they result in an underestimation of elderly mortality but to a relatively low degree (Figure 5); (iii) they alter mortality patterns at adult ages.

Table1: Sensitivity and specificity of mobile phone interviews in recording vital events

Events reported on :	**Sensitivity**	**Specificity**	**Kappa**
Father	92.1	82.6	80.3
Mother	97.5	71.9	81.6
Child	99.5	98.5	70.5

Table2: Consistency of information between the mobile phone survey and the face-to-face survey

Events	**Total**		**Child born during the last six years**	
	Agreement	*Kappa*	*Agreement*	*Kappa*
Sex of child	93.0	86.1	93.0	86.1
Survival status	97.5	88.5	98.7	93.8
Date of birth				
Day	67.2	66.1	70.7	69.7

Month	76.7	74.5	80.1	78.2
Year	70.2	66.2	74.6	69.3

Table3: Distribution of differences between ages in the HDSS and ages reported in the MPS

HDSS - MPS	***Age difference (HDSS- MPS)***							***Sample***
	<10	**10-6**	**5-3**	**0-2**	**3-5**	**6-10**	**>10**	
Gender	Design-based F(6.00, 16656.00) = 3.5660 **Pr = 0.002**							
Men	1.9	2.8	8.3	77.2	6.2	3.0	0.8	907
Women	1.9	2.2	5.8	74.5	9.9	4.8	0.9	1870
Place of residence	Design-based F(6.00, 16656.00) = 3.5405 Pr = 0.002							
Formal	2.3	2.4	6.1	78.9	6.4	3.3	0.6	1076
Informal	1.6	2.4	6.9	73.1	10.2	4.8	1.0	1701
Education	Design-based F(18.00, 49967.00) = 12.1247 **Pr = 0.000**							
None	3.2	3.8	8.1	60.6	14.6	7.9	1.7	972
Primary	1.7	1.7	7.5	77.7	7.5	3.5	0.3	636
Secondary	1.1	1.7	5.2	85.4	4.7	1.6	0.4	946
Superior	0.0	0.9	3.6	90.6	3.6	1.3	0.0	223
Place of interview	Design-based F(29.00, 80503.78) = 2.3141 **Pr = 0.000**							
Home	2.0	2.0	6.7	74.1	9.6	4.7	0.8	1780
Office	1.9	2.5	6.1	77.1	7.6	4.2	0.7	594
Public place	2.3	5.4	5.4	73.6	7.8	3.9	1.6	129
Traffic	4.2	0.0	12.5	79.2	4.2	0.0	0.0	24
Not at home	0.8	2.8	7.2	81.1	6.0	1.2	0.8	249
Others place	0.0	100.0	0.0	0.0	0.0	0.0	0.0	1
Total	1.9	2.4	6.6	75.4	8.7	4.2	0.8	2777

Appendix

1. *References*

Amo-Adjei, J., & Annim, S. K. (2015). Socioeconomic determinants of birth registration in Ghana. *BMC International Health and Human Rights*, *15*(1), 14. https://doi.org/10.1186/s12914-015-0053-z

Fisker, A. B., Rodrigues, A., & Helleringer, S. (2019). Differences in barriers to birth and death registration in Guinea-Bissau : Implications for monitoring national and global health objectives. *Tropical Medicine & International Health: TM & IH*, *24*(2), 166-174. https://doi.org/10.1111/tmi.13177

Helleringer, S., Pison, G., Masquelier, B., Kanté, A. M., Douillot, L., Duthé, G., Sokhna, C., & Delaunay, V. (2014). Improving the Quality of Adult Mortality Data Collected in Demographic Surveys : Validation Study of a New Siblings' Survival Questionnaire in Niakhar, Senegal. *PLOS Medicine*, *11*(5), e1001652. https://doi.org/10.1371/journal.pmed.1001652

Hill, K. (2013). *Direct estimation of child mortality from birth. In : Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I., Zaba, Z. (eds.) Tools for Demographic Estimation*. https://scholar.google.com/scholar_lookup?title=Direct+estimation+of+child+mortality+from+birth+histories&author=K+Hill&author=T+Moultrie&author=R+Dorrington&author=A+Hill&author=K+Hill&publication_year=2013&

Lamanna, C., Hachhethu, K., Chesterman, S., Singhal, G., Mwongela, B., Ng'endo, M., Passeri, S., Farhikhtah, A., Kadiyala, S., Bauer, J.-M., & Rosenstock, T. S. (2019). Strengths and limitations of computer assisted telephone interviews (CATI) for nutrition data collection in rural Kenya. *PloS One*, *14*(1), e0210050. https://doi.org/10.1371/journal.pone.0210050

Lankoandé, Y. B., Masquelier, B., Zabre, P., Bangré, H., Duthé, G., Soura, A. B., Pison, G., & Ali, S. (2022). Estimating mortality from census data : A record-linkage study of the Nouna Health and Demographic Surveillance System in Burkina Faso. *Demographic Research*, *46,* 653-680. https://doi.org/10.4054/DemRes.2022.46.22

Lyons-Amos, M., & Stones, T. (2017). Trends in Demographic and Health Survey data quality : An analysis of age heaping over time in 34 countries in Sub Saharan Africa between 1987 and 2015. *BMC Research Notes*, *10*(1), 760. https://doi.org/10.1186/s13104-017-3091-x

Mahfoud, Z., Ghandour, L., Ghandour, B., Mokdad, A. H., & Sibai, A. M. (2015). Cell Phone and Face-to-face Interview Responses in Population-based Surveys : How Do They Compare? *Field Methods*, *27*(1), 39-54. https://doi.org/10.1177/1525822X14540084

Masquelier, B., Kanyangarara, M., Pison, G., Kanté, A. M., Ndiaye, C. T., Douillot, L., Duthé, G., Sokhna, C., Delaunay, V., & Helleringer, S. (2021). Errors in reported ages and dates in surveys of adult mortality : A record linkage study in Niakhar (Senegal). *Population Studies*, *0*(0), 1-19. https://doi.org/10.1080/00324728.2020.1854332

Mikkelsen, L., Phillips, D. E., AbouZahr, C., Setel, P. W., de Savigny, D., Lozano, R., & Lopez, A. D. (2015). A global assessment of civil registration and vital statistics systems :

Monitoring data quality and progress. *The Lancet*, *386*(10001), 1395-1406. https://doi.org/10.1016/S0140-6736(15)60171-4

ONU. (2023). Objectif de Développement Durable—Santé et Bien-Être pour tous. *Développement durable*. https://www.un.org/sustainabledevelopment/fr/health/

Szolnoki, G., & Hoffmann, D. (2013). Online, face-to-face and telephone surveys—Comparing different sampling methods in wine consumer research. *Wine Economics and Policy*, *2*(2), 57-66. https://doi.org/10.1016/j.wep.2013.10.001

Zezza, A., Martuscelli, A., Wollburg, P., Gourlay, S., & Kilic, T. (2021). Viewpoint : High-frequency phone surveys on COVID-19: Good practices, open questions. *Food Policy*, *105*, 102153. https://doi.org/10.1016/j.foodpol.2021.102153

2. Graph

Figure1: Distribution of ages in the HDSS and ages reported in the MPS by sex of respondent (HDSS - MPS)

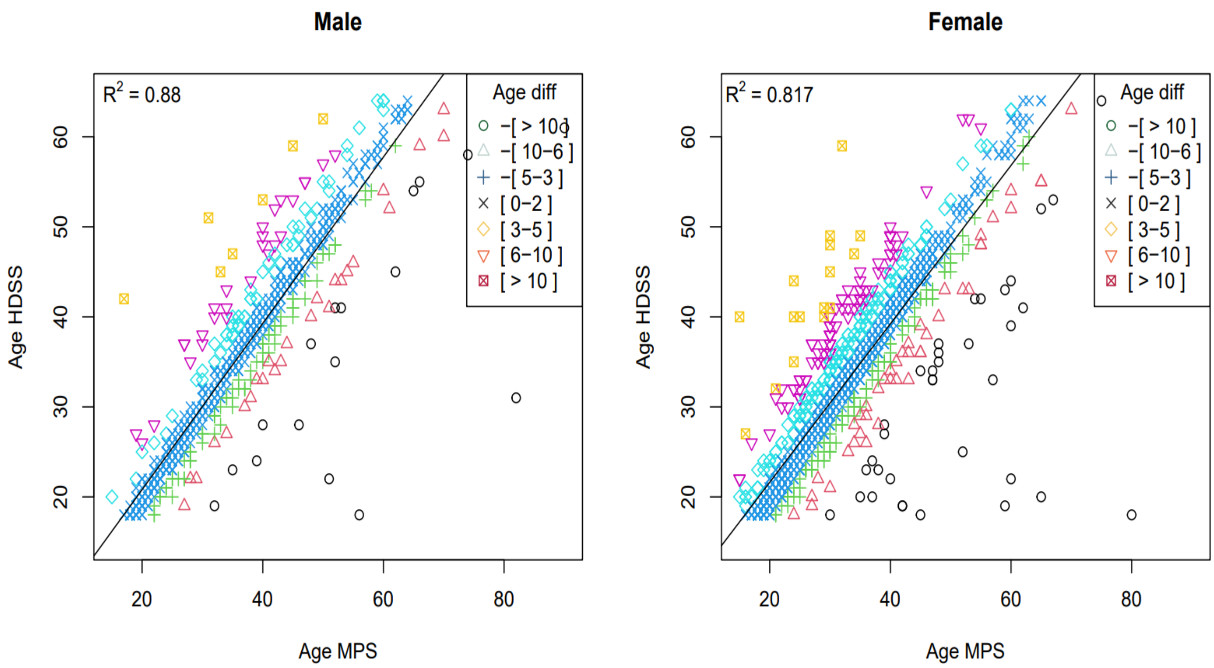
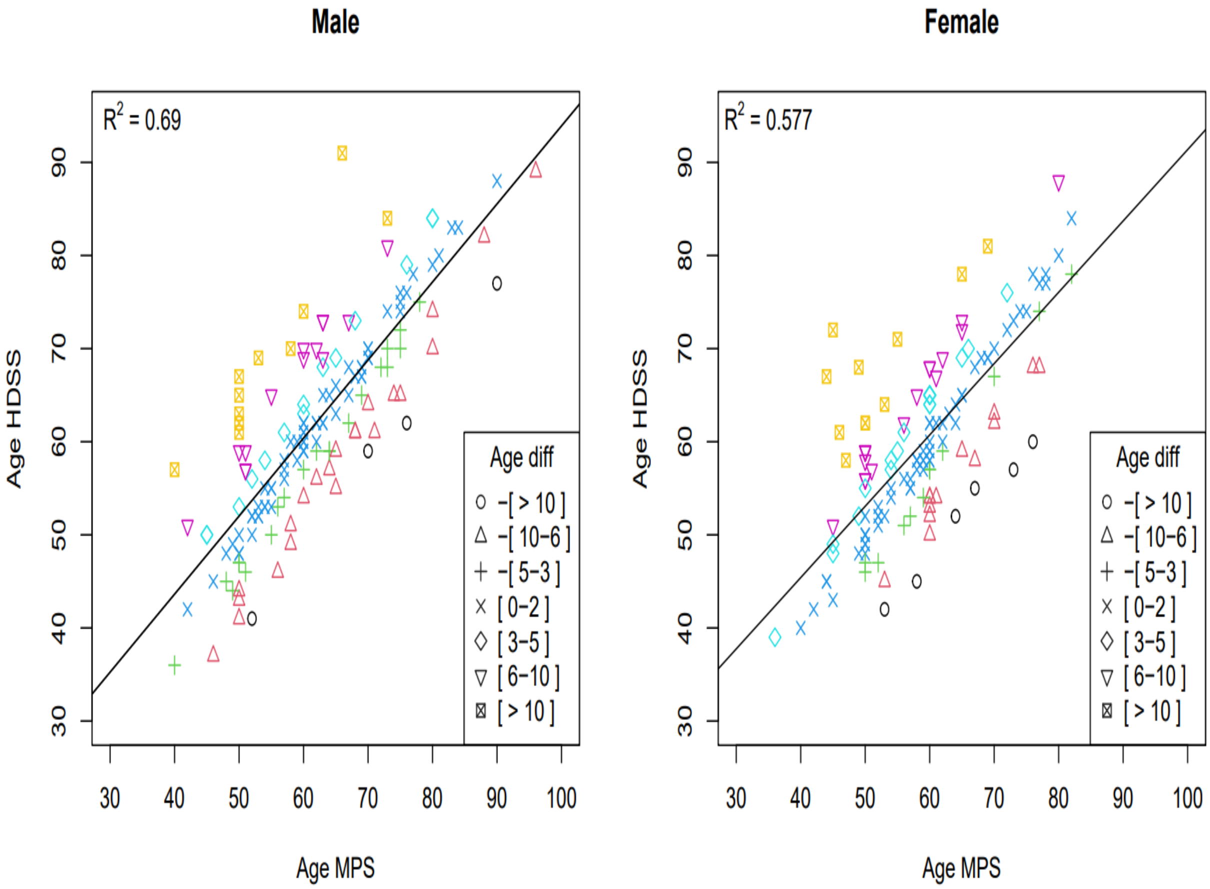


Figure2: Distribution of current ages of respondent's father in the HDSS and ages reported in the MPS by sex of respondent (HDSS - MPS)



3. Impact of inaccuracies in age reporting on the estimation of mortality levels

Figure 3: Risk of dying between age 0 and 5 (5q0): contrasting DHS data with DHS data with age errors from the RaMMPS validation survey, categorized by place of residence

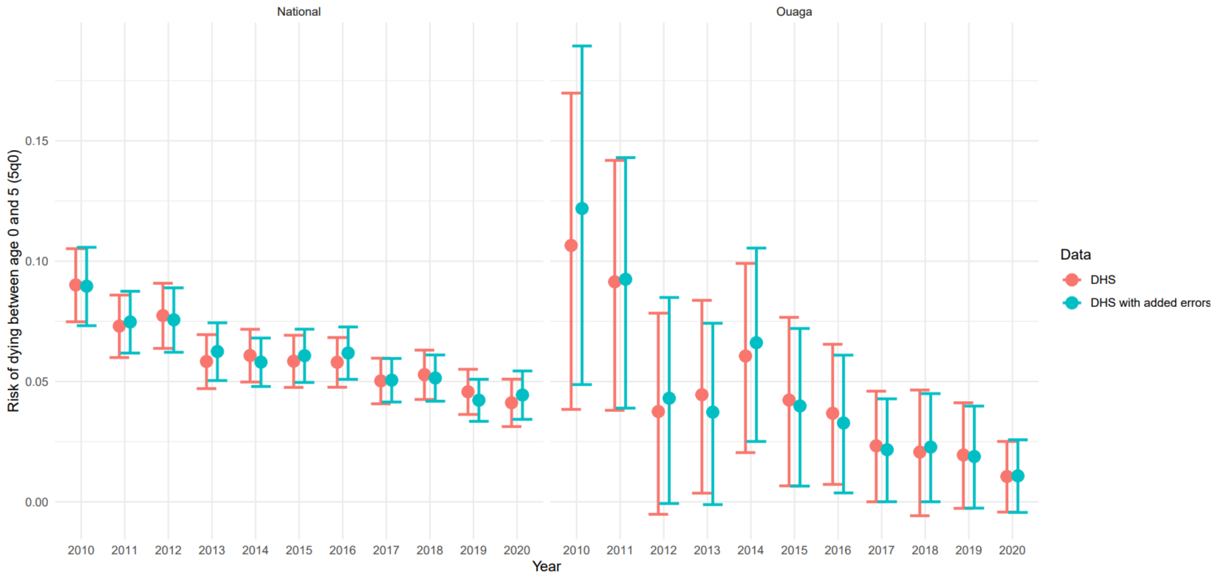


Figure 4: Risk of dying between age 15 and 50 (35q15): contrasting DHS data with DHS data with age errors from the RaMMPS validation survey, categorized by place of residence

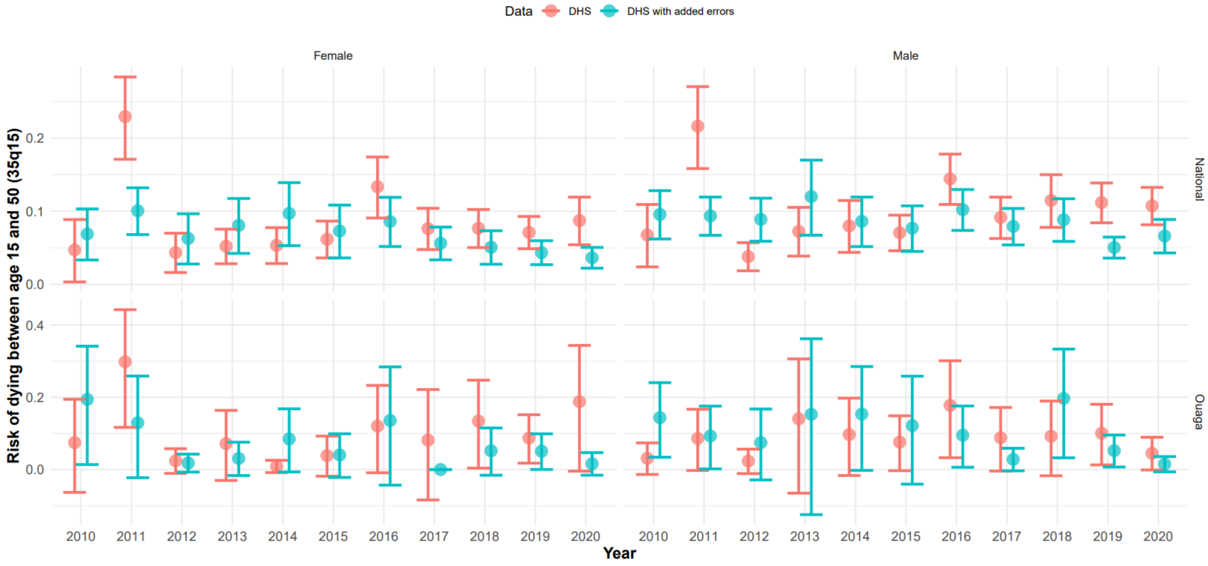


Figure 5: Proportions of parents surviving according to the sex of the parent and the age group of the respondents at the time of data collection contrasting census data with census data with age errors from the RaMMPS validation survey