



ined
INSTITUT
NATIONAL
D'ÉTUDES
DÉMOGRA
PHIQUES



7ème Conférence Africaine sur la Population

JOHANNESBURG - AFRIQUE DU SUD

30 Novembre - 4 Décembre 2015



Session 145

Data Quality Issues in African Surveys and Censuses

Chair: **Enock Ngome**, University of Botswana

Discussant: **Rolang Majelantle**, University of Botswana

Quality of age data from African censuses.

Evidence from census data matching in Mali.

*Quelle fiabilité accorder aux déclarations sur les âges en Afrique ?
Les enseignements d'un appariement de recensements au Mali.*

Assa Gakou Doumbia (INSTAT, Bamako, Mali)

Véronique Hertrich (INED, Paris, France)

Projet Slam – Suivi longitudinal au Mali

Projet ANR-DyPE – Dynamiques de la parentalité et de l'enfance en milieu rural africain - (ANR-12-BSH1-0005-01)



Estimating ages and dates: a significant issue



- **Difficulties to collect accurate data on ages and dates**
 - ✓ Civil registration is deficient, with limited coverage
 - ✓ Until recently, age and date were not familiar concepts
 - People don't use them in their everyday life
 - They do not memorize dates and ages
- **Census and survey questionnaires are overloaded with questions on age and date (respondent's age, marriages, births...)**
→ How to fill the gap between the need of data and the lack of knowledge?
- **Difficult job for the respondent and for the interviewer**

Various pieces of information may be used to provide answers

 - Administrative papers (certificates of birth, ID card, family card...)
 - Physical appearance
 - Family status (married or not, nb children, being grand-parent...)
 - Norms (age at marriage, age gap between spouses...)
 - Comparison with other family members (younger/older than the husband...), or events (« historical calendar »)

- **Outcomes**

- ✓ age heaping (stating final digit 0 or 5, normative age,...)
- ✓ age transfer:
 - under/overestimation according the respondent's characteristics
 - under/overestimation according the eligibility criteria of the survey
→ *sample distortion*

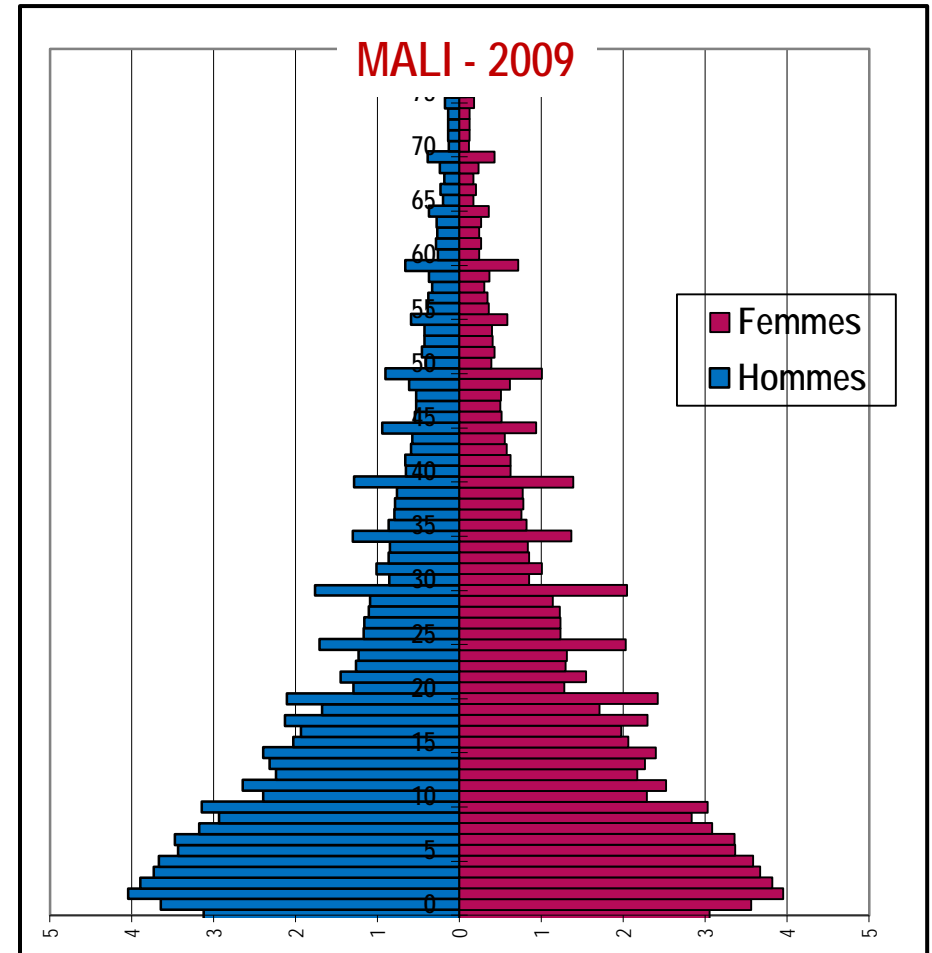
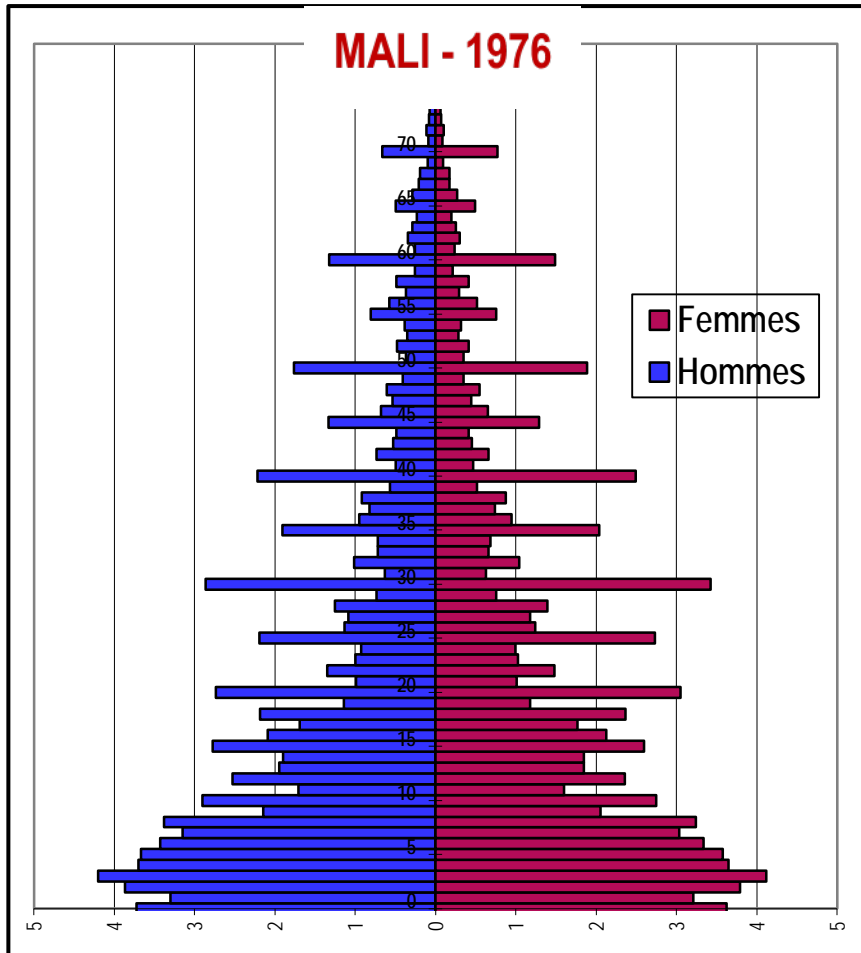
Impact on:

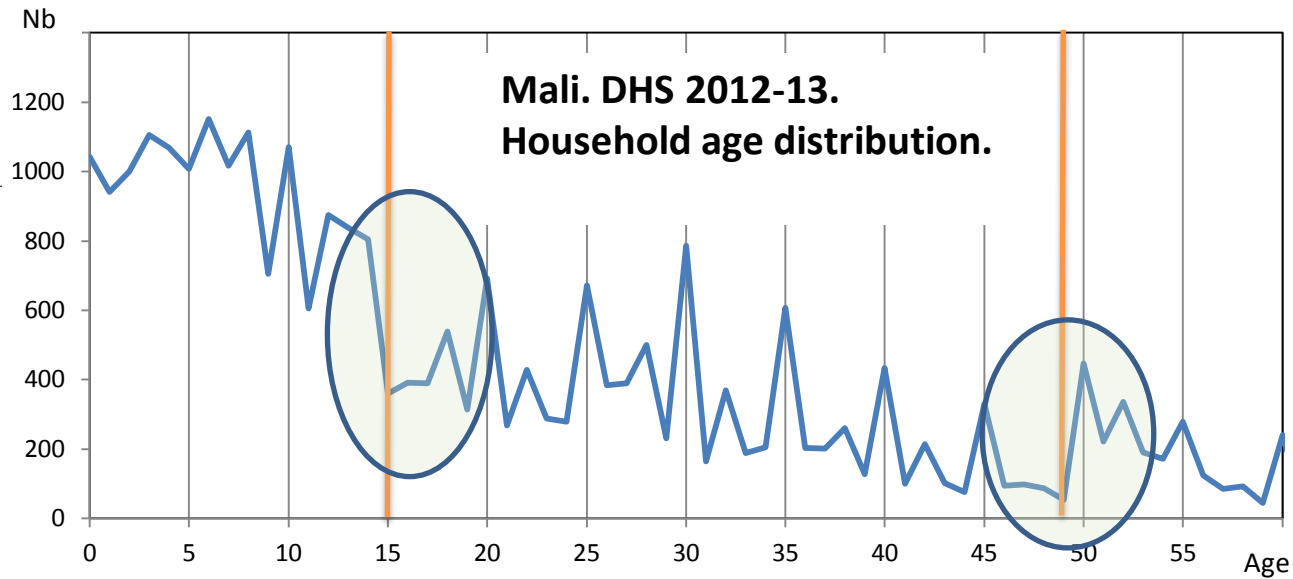
- ✓ age structure
- ✓ demographic levels and trends (fertility, age at marriage, child mortality...)
- Progress have been made but the age accuracy is still an issue for demographic research and a challenge to follow demographic trends in SubSaharan Africa

→ ***Illustration in the case of Mali***

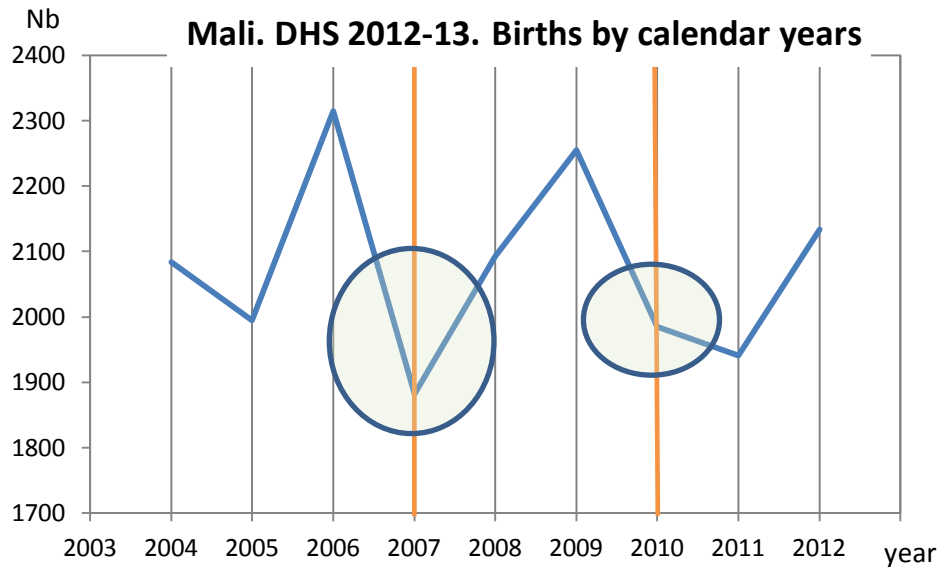
Population Pyramid. Mali.

First national census (1976) and last census (2009)





Eligibility for individual survey: women aged 15 to 49



Eligibility for the sections on:

- Pregnancy and postnatal care
- Child health

- Children born in 2007 or later
- Children born in 2010 or later

- Different publications on the issue of errors on age and their impact on demographic trends

For example: Pison, 1979; Ewbank, 1981; Roger et al, 1981; Lohle-Tart and Clairin 1988; Waltisperger 1988; Blanc and Rutenberg, 1990; Rutstein et al., 1990; Hertrich, 1992; Gage, 1995; Pullum and Stokes, 1997; Lohle-Tart and Francis, 1999; Pison and Ohadike. 2006; Pullum , 2006; Moultrie and Dorrington, 2008; Schoumaker, 2009; Johnson et al., 2009; Hertrich and Lardoux, 2014; Lecomte and al., 2015

- Specific point in every census/survey report

- Usually the issue is adressed through cross-sectionnal indicators: evidence of biases at the level of the outcomes

- *Here*: longitudinal data at the individual level (7 villages in Mali)

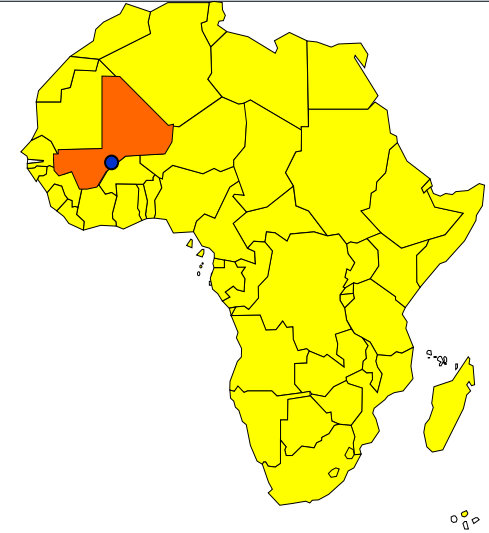
- ✓ Linkages of individual records from 4 national censuses

- ✓ Level of inconsistencies, patterns by age and sex

- ✓ Impact of changes in marital status to the statement on age

Context

- 7 villages (4300 inhab. In 2009)
- Southeast Mali, 450 km from Bamako
- Bwa ethnic group
- Farmers, family-based production
- Low school enrolment until the 1990s (now primary rate about 50%)
- High fertility (TFR=8)



The observation system

Slam Project. Suivi longitudinal au Mali/Longitudinal survey in Mali

- implemented in 1987-89 → *retrospective approach*
- a new round every 5 years → *prospective approach*
- last round: 2009-10.

Follow-up survey (« Enquête renouvelée »)

- local censuses in the 7 villages approximately every 5 years (1988, 1994, 1999, 2004, and 2009).
- questionnaires of the four national censuses (1976, 1987, 1998, 2009)
 - *to match at the individual level*
 - *33 years, 1976-2009*

- **The principle of the survey:**
 - ✓ Input: **cross-sectional**, independant data
 - ✓ Ouput: (semi-)**longitudinal** data, individual itineraries
- **The matching process:**
 - ✓ Data are organized by domestic groups (zû)
 - ✓ First matching is done by hand
 - ✓ Work meetings by families (lineages):
 - To control and complete the matching
 - To collect additionnal data

→ *Objective: know the status and place of residence at every census for each individual registered by at least one census.*
- **The bet of the survey: identify and follow everyone**
- **Few « unknown » individuals :**
3 from the 1976 census (0,1 %), 5 from the 1998 census (0,1%).
- **Database:** N=9200 indiv. recorded as resident at least at 1 census

Potential to evaluate the quality on age report

- 4 national censuses = 4 independent sources
- 1 individual registered by N national censuses → N records of age

Individuals recorded by several national censuses:

	Men	Women	all
At least 2 censuses	1752	1471	3223
At least 3 censuses	750	497	1247
Paired observation: census(t) & census (t+10)	2744	2109	4853

- Approach: Comparing 2 census registrations of the respondent's age
- National censuses at (t) and (t+10) → 4853 linked observations

Indicator of consistency

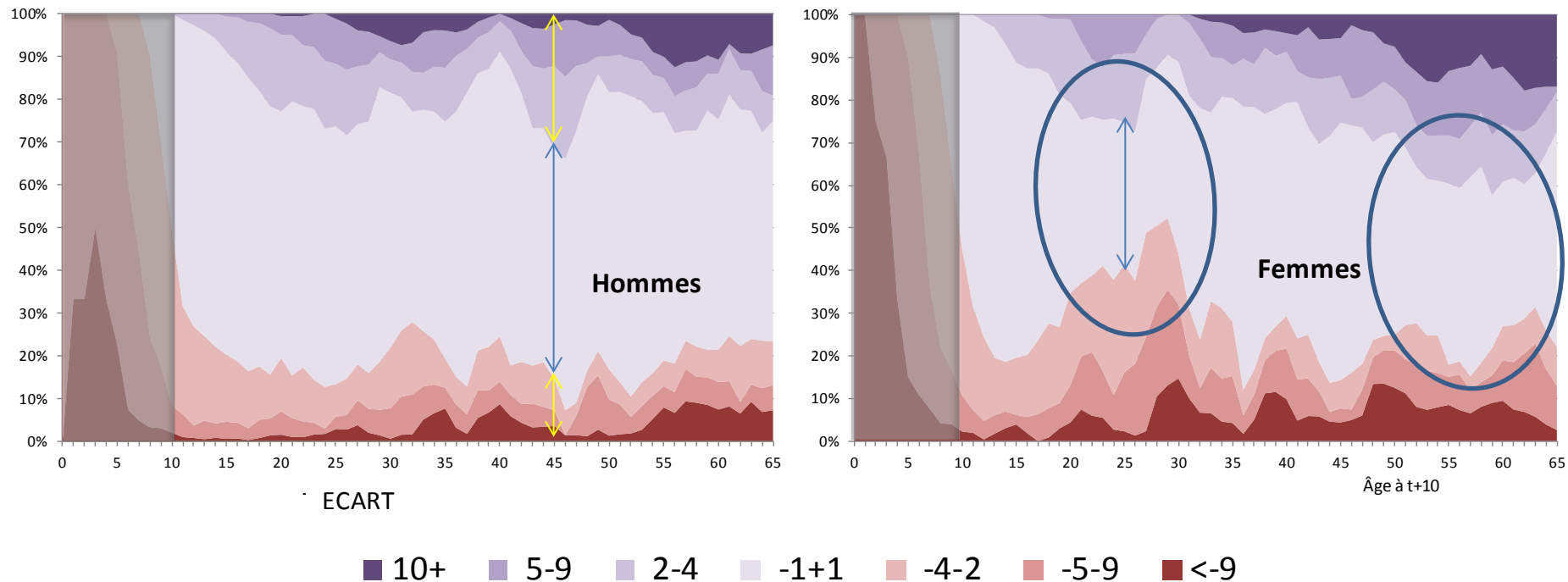
- From each census: an estimation of the year of birth
 $\text{year of birth} = \text{year of census} - \text{age}$
- **Indicator of consistency:**
 $\text{year of birth (census } \underline{t}) - \text{year of birth (census } \underline{t+10})$

	Consistency +- /1 year %	Consistency +- /2 years %	Inconsistency 5 years and more (%)	Inconsistency 10 years and more (%)
C76-C87	46	58	26	10
C87-C98	58	69	20	8
C98-C09	59	70	18	6
Men	59	71	17	6
Women	50	61	26	10
All	55	67	21	8

Inconsistencies on age between censuses (t, t+10) according age at the second census (t+10) (moving average)

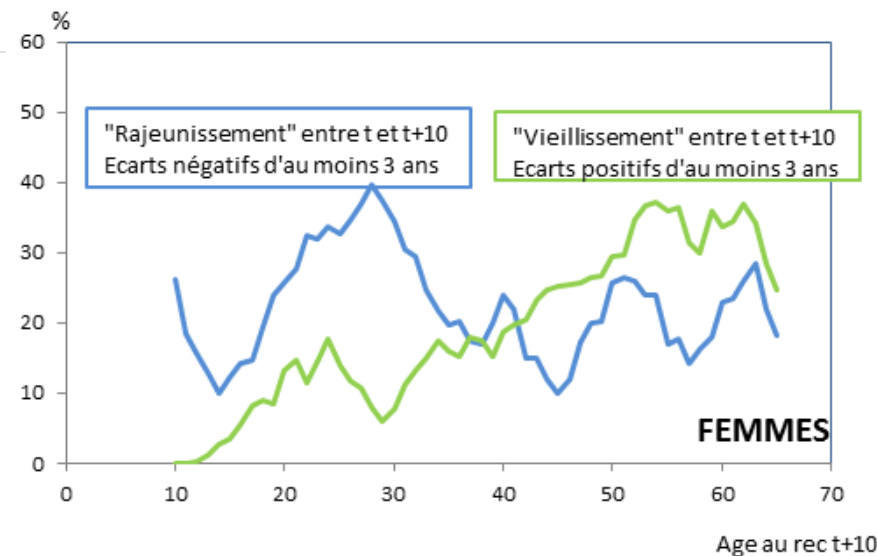
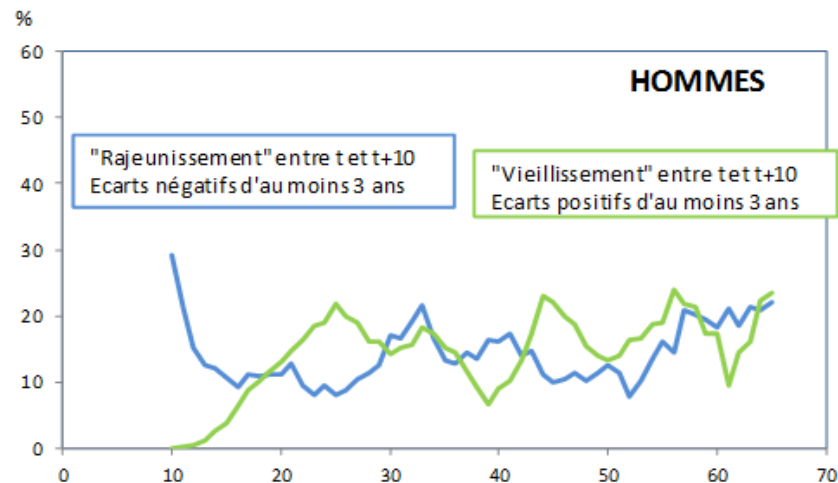
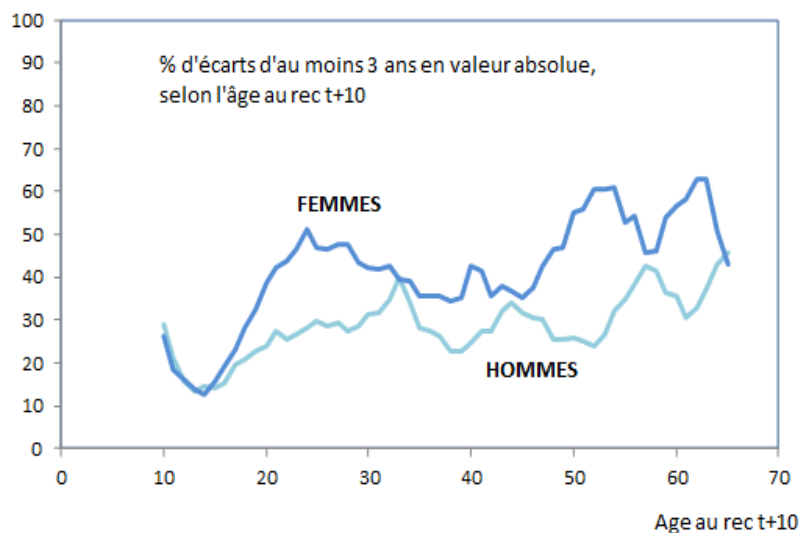
Red-Pink: Negative gap → « rejuvenation », age at t+10 < age predicted by the report at t
Violet: Positive gap → « ageing », age at t+10 > age predicted by the report at t

We do not know which report is the best (neither if one is accurate)



Focus on the inconsistencies of 3 years and more

% discrepancies 3 years and more



Are you getting older more rapidly if you get married?

Marital itineraries between censuses:

First census

Never-married



Second census

Never-married

Never-married



Ever-married

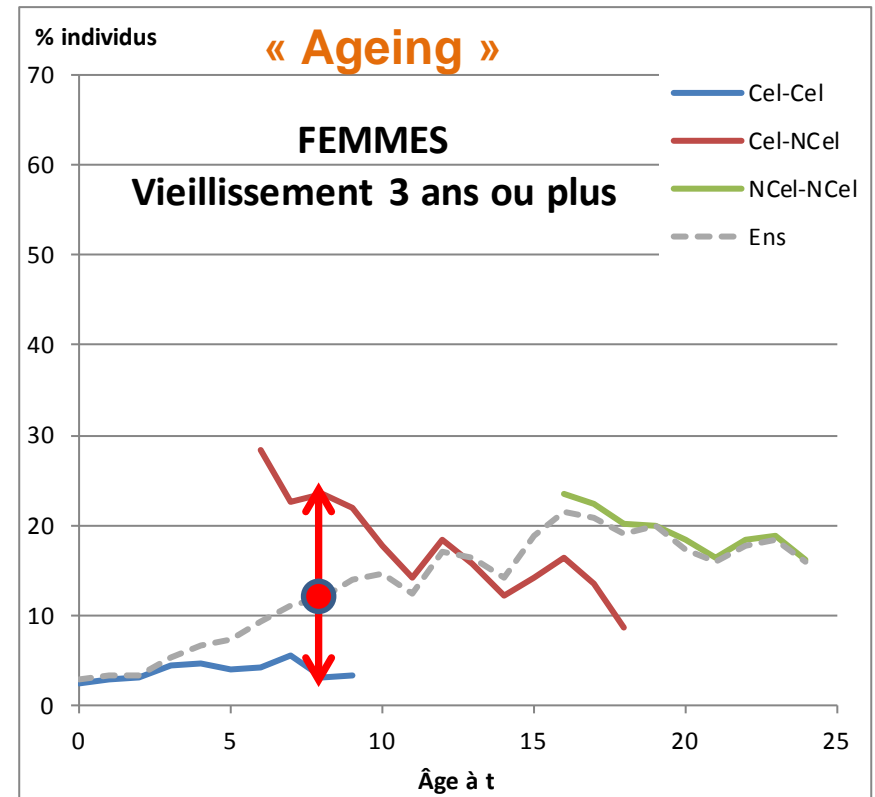
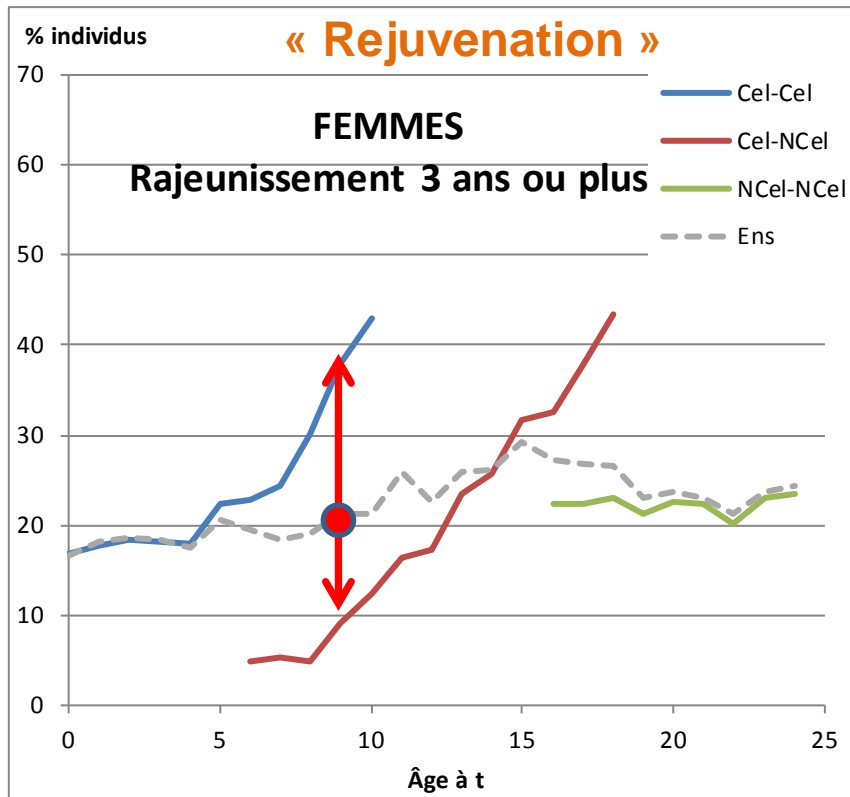
Ever-married

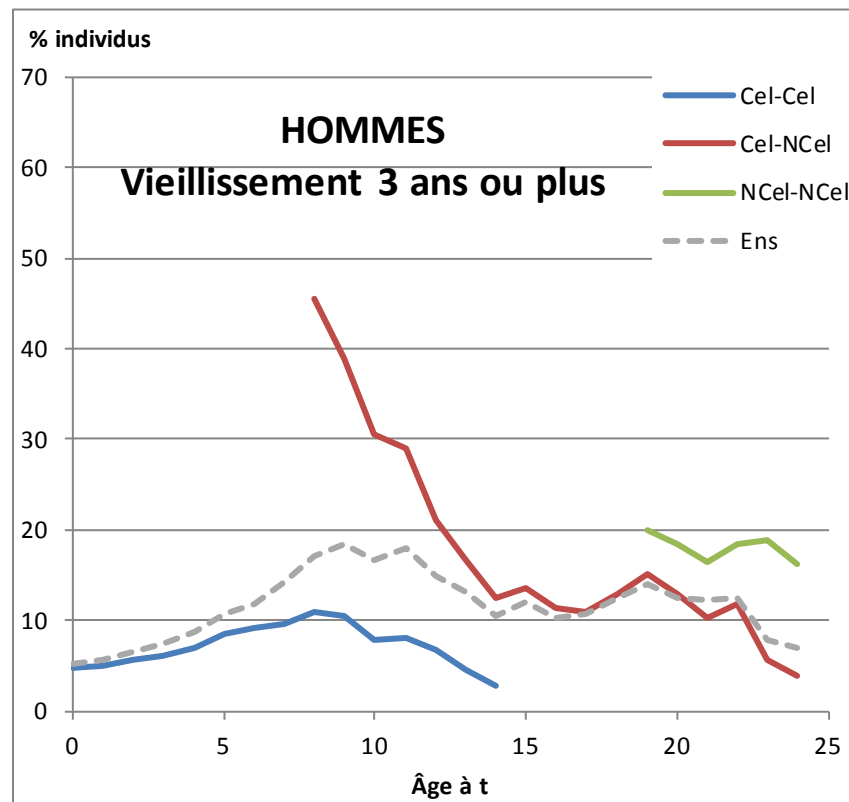
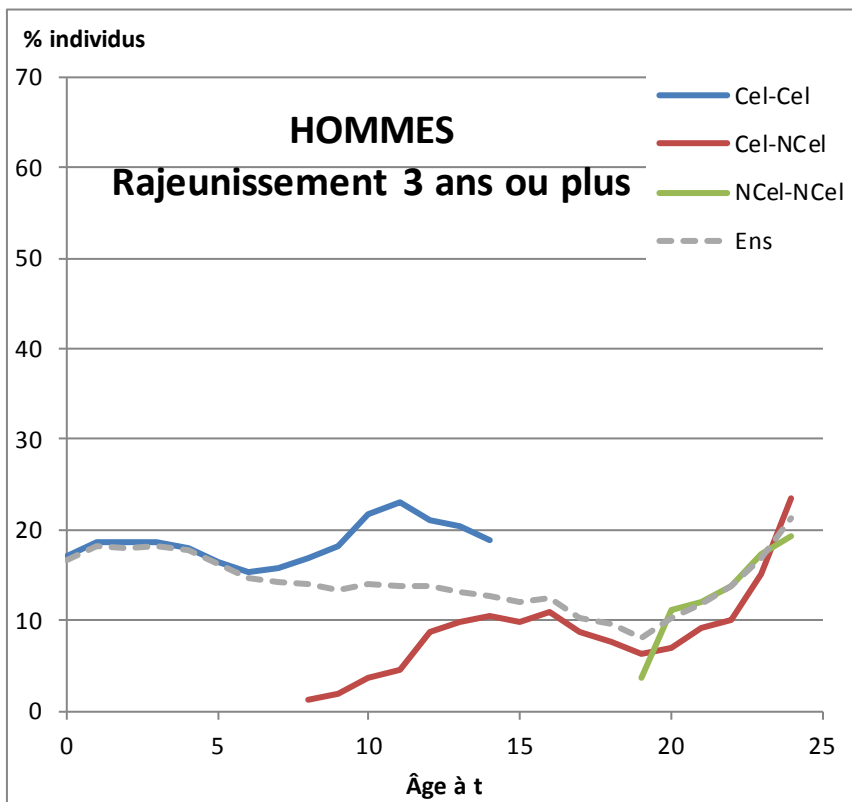


Ever-married

Inconsistencies 3 years and more, according the marital itinerary between censuses

Women





To conclude

- Work in progress. Next steps, taking into account :
 - ✓ Childbearing, nb children
 - ✓ Marital changes at older ages (polygamy, widowhood...)
 - ✓ Status in the household (Head, relationship to the HHH)
 - ✓ Status according relatives: husband getting polygamist, children getting married...
 - ✓ Individual characteristics: schooling, religion
 - ✓ HH characteristics
- Quality of data on age is still a critical issue:
 - ✓ Discrepancies are common
 - ✓ Discrepancies are often large
 - ✓ No random pattern of distortion → biases
- Age is not an objective information: it is significantly reformulated for a same individual when her/his social/family status is changing
- Special issue: adolescent ages and transition to adulthood
→ biases in analysis on adolescent reproductive health
- Faisability of matching national census data with small scale surveys: to replicate?

Thank you Merci !

***Assa GAKOU DOUMBIA
(assdoumbia@yahoo.fr)***

&

***Véronique HERTRICH
(hertrich@ined.fr)***

