LIVING to 100

SOCIETY OF ACTUARIES INTERNATIONAL SYMPOSIUM

Jan. 4–6 , 2017
Orlando, FL

Session 2B, Older Age Mortality Trends

Moderator:

W. Ward Kingkade, Ph.D.

Presenters:

Jean-Marc Fix, FSA, MAAA
Samuel Gbari
Robert Bourbeau, Ph. D.
Adrian Gallop, FIA

Improvement in late-life mortality and its impact on the increase in the number of Centenarians in Quebec (Canada)

Robert Bourbeau
Melissa Beaudry-Godin
Bertrand Desjardins

Department of Demography, Université de Montréal

SOA 2017 Living to 100 Symposium
Orlando, USA
January 4-6, 2017

Université de Montréal

Outline of the presentation

- Background: Increase in the number of centenarians
- Data and Methods
- Estimation of centenarians in Canada and Quebec
- Factors responsible for the increase in numbers of centenarians
- Conclusion : Challenges

The 21st century – the century of centenarians

Published on N-IUSSP.ORG November 21, 2016

George W Leeson

"The emergence of large numbers of centenarians has accompanied the ageing of our populations. The number of people aged 100 years and over in England and Wales, for example, increased from less than 200 in 1922 to 570 in 1961. By 1981it had climbed to 2,418 and to 12,318 in 2012…."

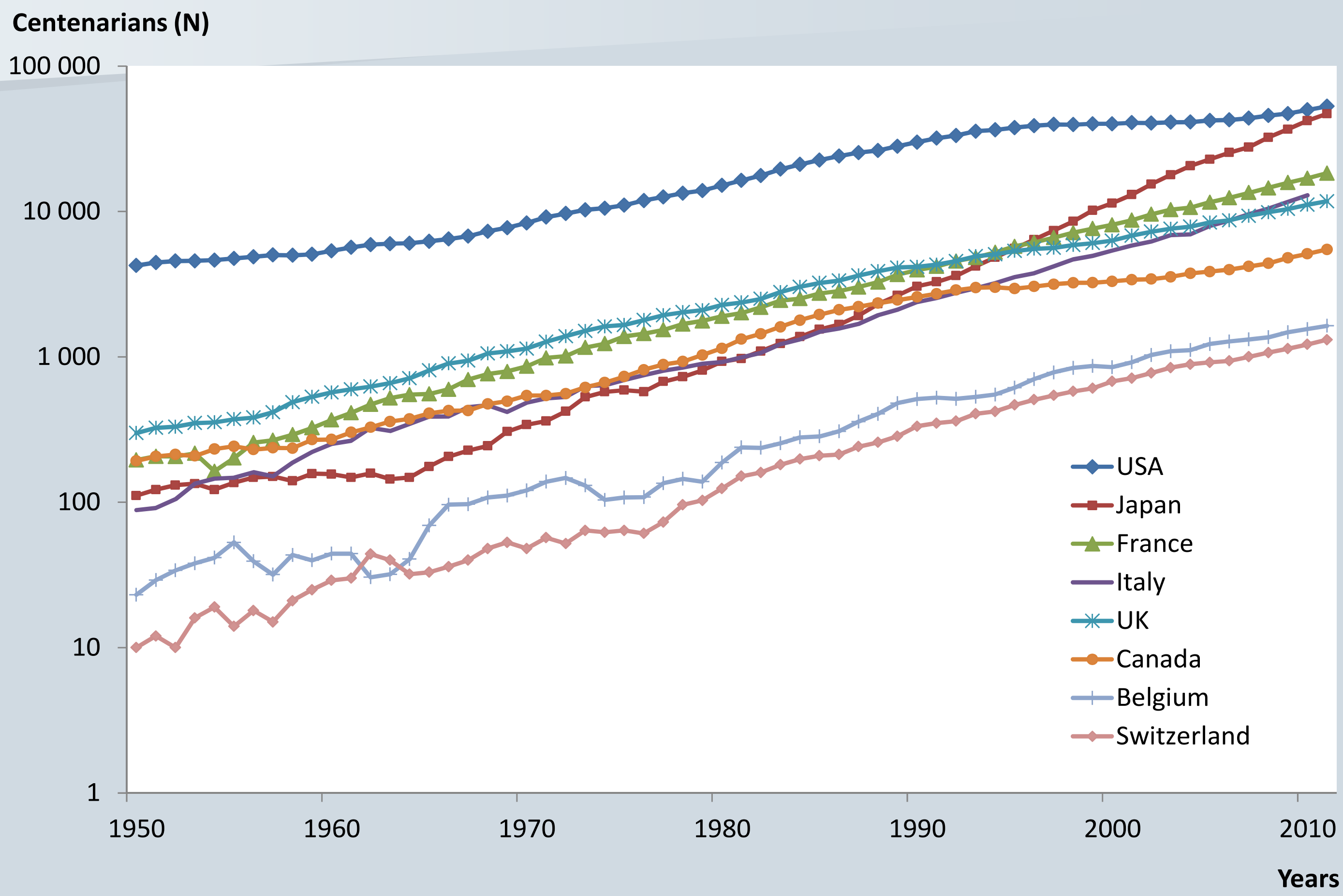
Table I

Number of centenarians and sex ratios of centenarians in selected countries, 1 January 2011

	Men	Women	Total	Centenarians per 10 000 inhabitants	W/M
USA	8 181	44 742	52 923	1.7	5.5
Japan	6 093	40 914	47 007	3.7	6.7
France	2 224	16 112	18 336	2.9	7.2
UK	1 596	10 154	11 750	1.9	6.4
Canada	844	4 638	5 482	1.6	5.5
Belgium	206	1 428	1 634	1.5	6.9
Switzerland	200	1 111	1 312	1.7	5.6

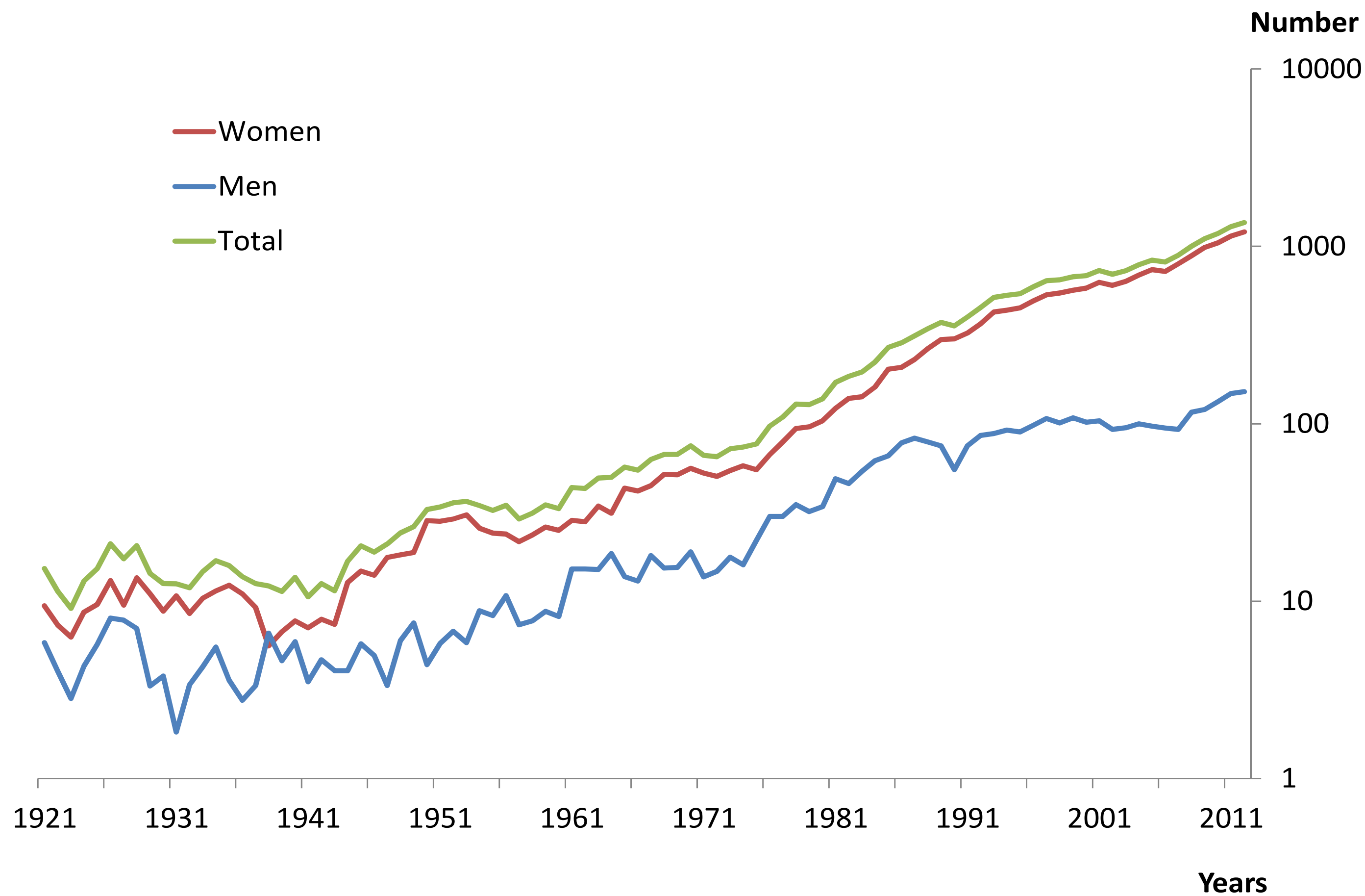
Figure 1

Evolution of numbers of centenarians (those aged 100 and over), selected countries, 1950-2011



Source : HMD

Evolution of numbers of centenarians by sex, Quebec, 1921-2012



Source : BDLC

Ratio of centenarians per 10,000 individuals aged 60 forty years earlier (RC_{60}), Quebec, 1961 to 2011

Year	Men	Women	Total
1961	11.6	19.4	15.5
1971	6.3	29.0	17.4
1981	21.5	55.4	37.9
1991	29.4	104.7	66.6
2001	31.2	147.7	90.3
2011	34.7	196.4	118.3

$$RC_{[60]} = \frac{P(100)^{t}}{P(60)^{t-40}} * 10\,000$$

Comparison of centenarian ratio (RC_{60}), selected countries, 1961 to 2011

	Japan	France	Canada	**Québec**	USA	Italy	Switzerland	UK	Belgium
1961		5.4	24.0	**15.5**		3.9	3.1	8.0*	3.7
1971		11.4	34.0	**17.4**		7.7	7.3	14.7	9.3
1981		22.6	62.8	**37.9**	59.8	14.7	19.9	21.6	13.8
1991	29.2	43.4	99.4	**66.6**	91.3	28.2	35.9	34.6	25.6
2001	81.8	72.0	105.2	**90.3**	103.8	51.1	57.9	48.9	37.7
2011	227.9	135.3	135.0	**118.3**	111.0	91.4	87.4	72.8	63.2

$$RC_{[60]} = \frac{P(100)^{t}}{P(60)^{t-40}} * 10\,000$$

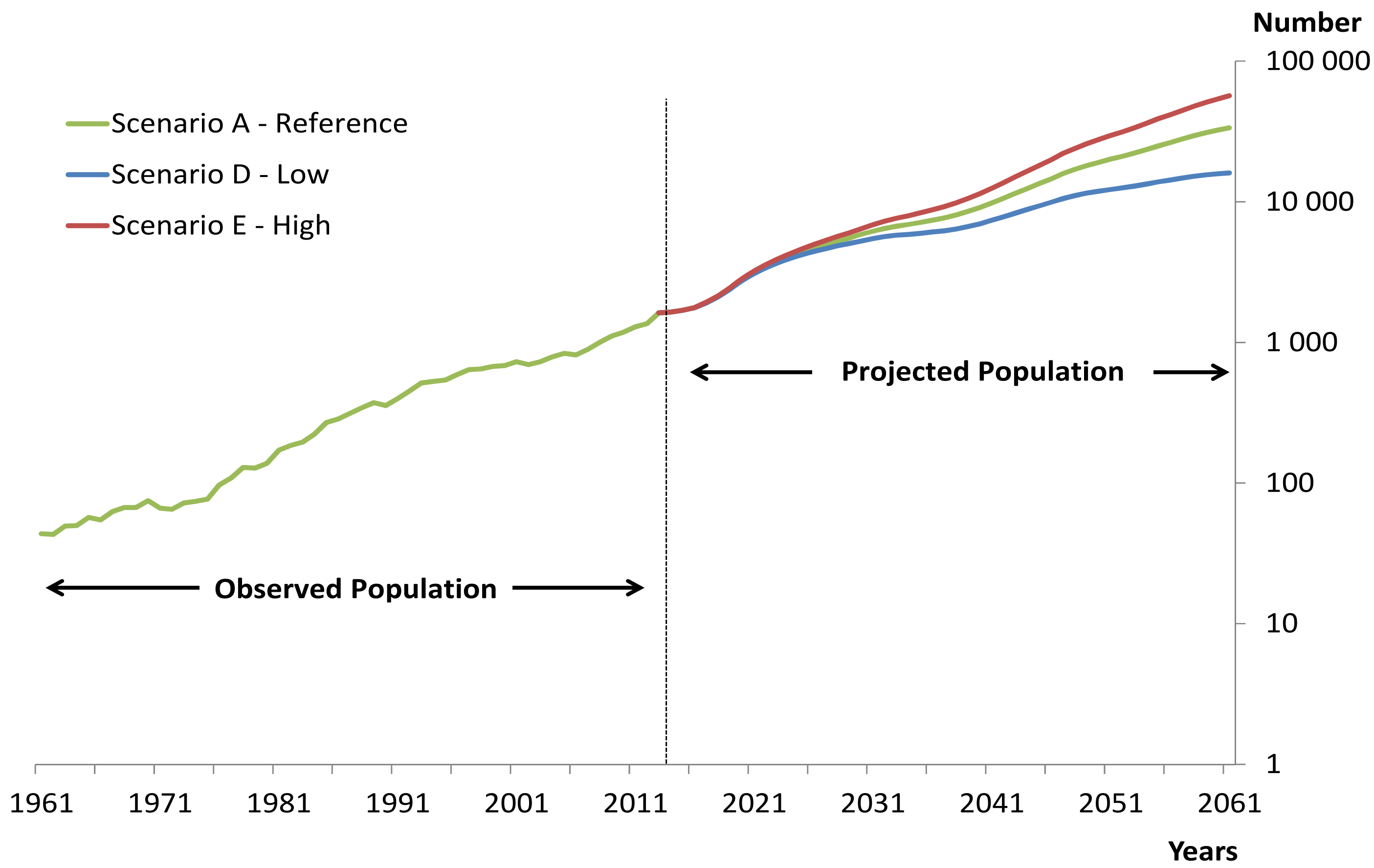
Factors responsible for the increase in numbers of centenarians

- Increased cohort size
- Increase in the probability of survival from birth to age 80
- Increased probability of surviving from age 80 to age 100

Breakdown of factors responsible for the rise in numbers of centenarians, Quebec, generations of 1871 and 1901

Factors involved in increase	Numbers (Sx)			Probability of surviving	
	Age 0	Age 80	Age 100	${}_{80}p_{0}$	${}_{20}p_{80}$
MEN					
1871 cohort	23741	2070	8	0.0872	0.0040
1901 cohort	31882	4372	46	0.1371	0.0105
Growth coefficient	**1.34**	**2.11**	**5.55**	**1.57**	**2.64**
WOMEN					
1871 cohort	22490	2398	21	0.10166	0.0088
1901 cohort	30363	7089	201	0.233504	0.0283
Growth coefficient	**1.35**	**2.96**	**9.46**	**2.19**	**3.20**
TOTAL					
1871 cohort	46231	4469	28	0.0967	0.0066
1901 cohort	62245	11461	255	0.1841	0.0213
Growth coefficient	**1.35**	**2.56**	**8.30**	**1.90**	**3.23**

Observed and projected centenarian population, three forecast scenarios, Quebec, 1961-2061



Source : BDLC and Institut de la statistique du Québec, Perspectives démographiques du Québec et des régions, 2011-2061.

Challenges

- The century of centenarians : a challenge in both the developed and the emerging countries
- Retardation of life cycle (old age and retirement)
- Intergenerational families and caregiving
- Adaptation to the needs of extreme-aged populations

Acknowledgements

- Data on Centenarians:
  Statistics Canada
  CHMD

- Financial support:
  Social Sciences and Humanities Research Council of Canada

EXTREME VALUE ANALYSIS OF MORTALITY AT THE OLDEST AGES : A CASE STUDY BASED ON INDIVIDUAL AGES AT DEATH

Samuel GBARI
samuel.gbari@uclouvain.be

Joint work with M. Poulain, L. Dal and M. Denuit
Universite Catholique de Louvain (Belgium)
Institut de Statistiques, Biostatistiques et de Sciences Actuarielles

SOA Living to 100 Symposium (Orlando)
January 2017

INTRODUCTION

- Exponential increase in mortality force is appropriate for adults and early old ages for actuarial computations (pensions, life insurance).
- There is a need for reliable estimates for mortality rates at oldest ages because of data scarcity in these regions of age.
- Extreme Value Theory (EVT) offers us an alternative to approximate the upper tail of lifetime distribution.
- This theory was applied on a Belgian database recording individual ages at death of the whole population above age 95.
  The model obtained had various applications such as
    - Estimation the ultimate age,
    - Comparison of mortality rates estimated from classical demographical methods,
    - Pointwise estimation of lifespan high quantiles.

AGENDA

1. INTRODUCTION
2. METHODOLOGY
3. DATA
   - Data source and validation process
   - Descriptive statistics
4. ESTIMATION
   - Threshold selection
   - ML estimates
   - Goodness-of-fit
   - Alternative estimators
5. SOME APPLICATIONS
   - Ultimate age
   - Old mortality rates
   - Point estimate of high quantiles
6. CONCLUSION

REFERENCES

METHODOLOGY

- Peaks-Over-Threshold (POT) approach was used.
- It requires to fix a high threshold age $x^\star$.
- 3 automatic selection procedures :
  - Pickands
  - Reiss & Thomas
  - Extreme value mixture model.
- POT states that the residual lifetime distribution at $x^\star$ can be approximated by the Generalised Pareto Distribution (GPD).
- GPD has 2 parameters
  - ξ (Tail index)
  - β (Scale parameter).
- Parameters are fitted by maximum likelihood estimation.
- When $\xi < 0$, there exits an theoretical ultimate age defined as

$$w = x^\star - \frac{\beta}{\xi}.$$

DATA SOURCE AND VALIDATION PROCESS

- Data are from the Belgian national population register.
- A total of 22 extinct cohorts from 1886 to 1904.
- It consists in a list of 46,666 persons who died after 95 years old since 1981 till 2015.
- The dataset contains
  - 10,050 males (22%)
  - 36,616 females (78%).
- Data have been validated in a number of ways:
  - Individuals not born in Belgium were excluded.
  - Date of birth recorded from official certificates.

DESCRIPTIVE STATISTICS (Overall)

	Males	Females
Mean	97.35	97.75
Std. Dev.	2.05	2.37
Q1	95.77	95.91
Q2	96.79	97.12
Q3	98.36	98.99
Max	111.47	112.58

Table 1 : Descriptive statistics for the observed ages at death above 95.

- All female age statistics (mean, quartiles, maximum) are greater than male ones for the overall population.
- The oldest ages recorded in the dataset are respectively **111.47** for males and **112.58** for females.

DESCRIPTIVE STATISTICS (by cohort)

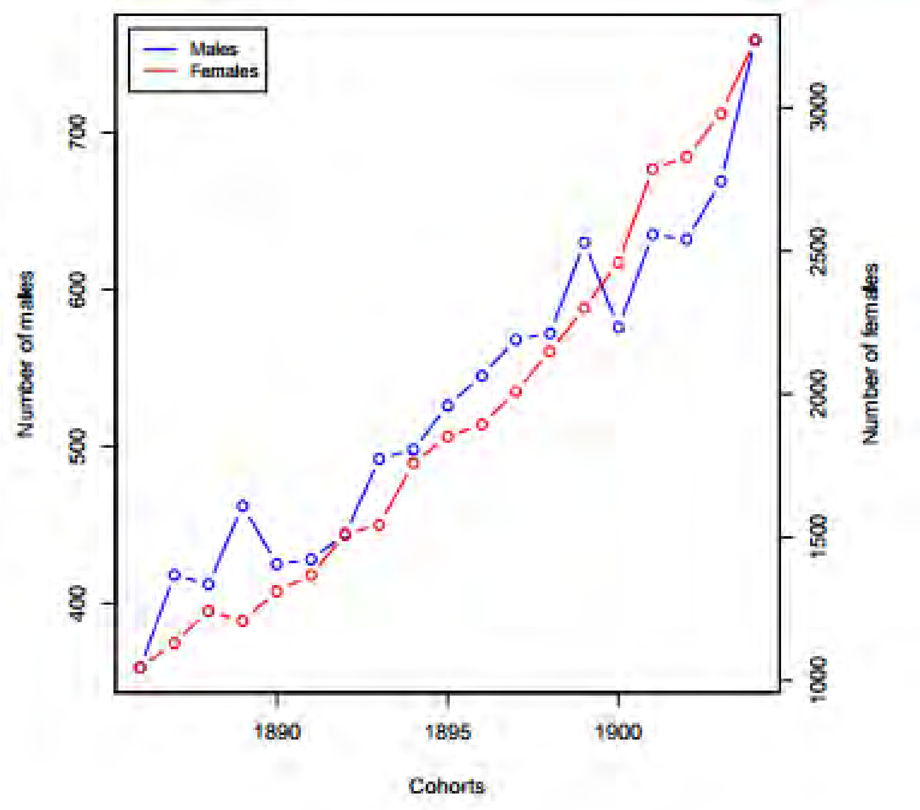


Figure 1 : Cohort numbers of survivors at 95 by sex.

DESCRIPTIVE STATISTICS (by cohort)

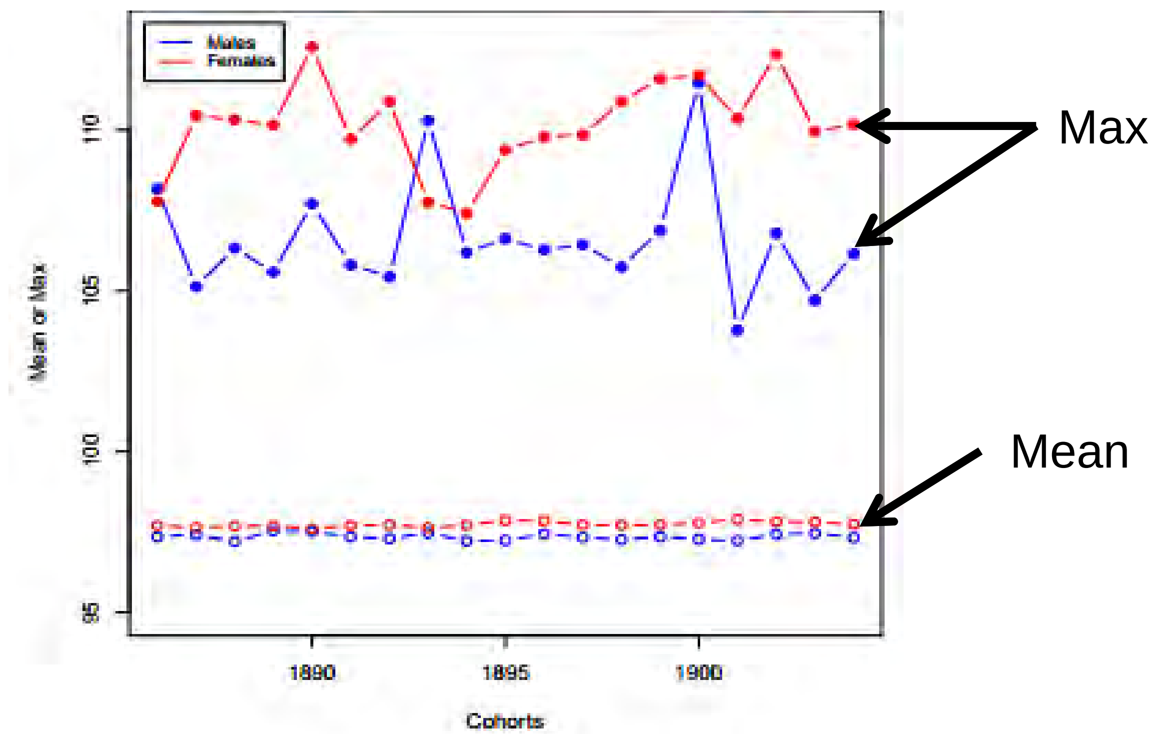


Figure 2 : Cohort mean and maximum ages at death by sex.

- It seems that e_{95} did not improve across the 19 birth cohorts.
- It suggests no potential gain of extreme longevity through generations at these advanced ages.

THRESHOLD SELECTION

	Males		Females	
	≥ 95	≥ 98	≥ 95	≥ 98
Pickands	95.24	98.40	95.41	98.60
Reiss&Thomas	95.03	98.07	100.89	**100.89**
Mixture model	97.82	**98.89**	100.01	99.99

Table 2 : Thresholds selected by automatic procedures.

- Threshold age for males is lower than the female one.
- Threshold age becomes more stable for $x \geq 98$.
- Optimal threshold was set at maximum of selected thresholds thanks to the GPD threshold stability property :
    - $x^{\star}$= 98.89 (Male).
    - $x^{\star}$= 100.89 (Female).

ML ESTIMATES

	Males	Females
$x^\star$	98.89	100.89
$L_{x^\star}$	1,940	4,104
$\hat{\xi}^{ML}$	-0.132	-0.092
$s.e\ (\hat{\xi}^{ML})$	0.015	0.014
$\hat{\xi}_+$	-0.102	-0.065
$\hat{\beta}^{ML}$	2.098	2.019
$s.e\ (\hat{\beta}^{ML})$	0.057	0.042

Table 3 : ML estimates for the GP parameters β and ξ.

- There is a substantial number of observations for estimation.
- $\hat{\xi}_F < 0$ and $\hat{\xi}_M < 0$.
- As $\hat{\beta}_F > \hat{\beta}_M$ suggests $\hat{w}_F > \hat{w}_M$ for a fixed value of $\hat{\beta}$.
- Confidence interval upper bounds $\hat{\xi}_+ < 0$ confirms that $\xi < 0$.

GOODNESS-OF-FIT

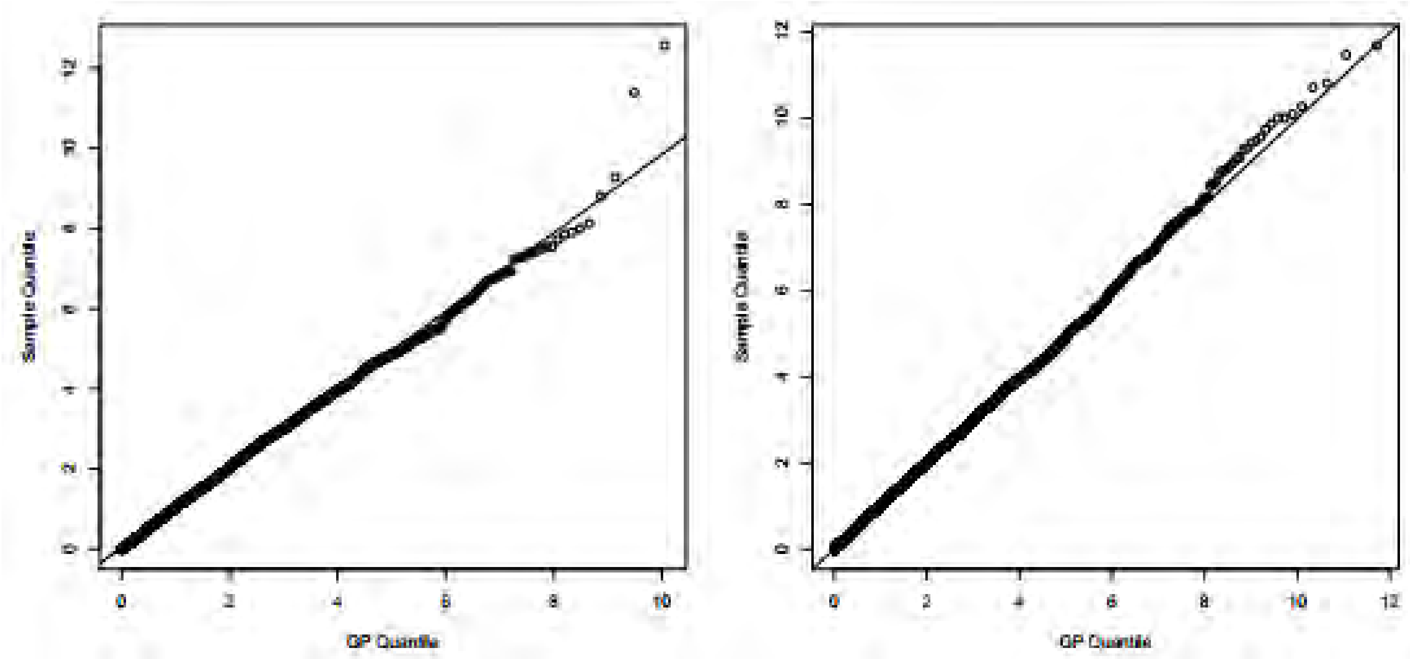


Figure 3 : GP QQ-plots for the remaining lifetimes at age $x^{\star}$ for both males (left panel) and females (right panel).

- KS test pvalue : 0.828 (Male) and 0.951 (Female).
- This supports the good fitting of the GPD.

ALTERNATIVE ESTIMATORS

		ML	MOM	WMOM	M	H
Males	$\hat{\xi}$	-0.132	-0.163	-0.171	-0.163	-0.155
	$\hat{\beta}$	2.098	2.160	2.175	2.128	2.113
Females	$\hat{\xi}$	-0.092	-0.097	-0.105	-0.097	-0.094
	$\hat{\beta}$	2.019	2.027	2.043	1.997	1.991

Table 4 : Alternative estimators of GP parameters

- Same differences previously observed between gender.
- Estimation method clearly impacts on $\hat{\xi}$.
- This will lead to different conclusions about the value of w.
- Despite these differences, all estimators support that $\xi < 0$.
- It suggests existence of a theoretical finite ultimate age estimated by

$$\hat{w} = x^{\star} - \frac{\hat{\beta}}{\hat{\xi}} .$$

ULTIMATE AGE

		ML	MOM	WMOM	M	H
Males	$\hat{w}$	**114.82**	112.14	111.64	111.91	112.53
	$\hat{w}^-$	112.31	109.81	109.05	109.07	109.29
	$\hat{w}^+$	118.87	116.20	116.65	114.75	115.76
Females	$\hat{w}$	**122.73**	121.86	120.30	121.41	112.14
	$\hat{w}^-$	118.19	117.43	115.80	115.80	116.09
	$\hat{w}^+$	131.21	130.22	129.53	127.03	128.20

Table 5 : Estimated ultimate ages with 95% simulated confidence intervals.

- Results are consistent with Belgian maximum observed age.
- They are also worlwide interesting (122.42 Calment (female), 116 Kimura (male)).

MORTALITY RATES

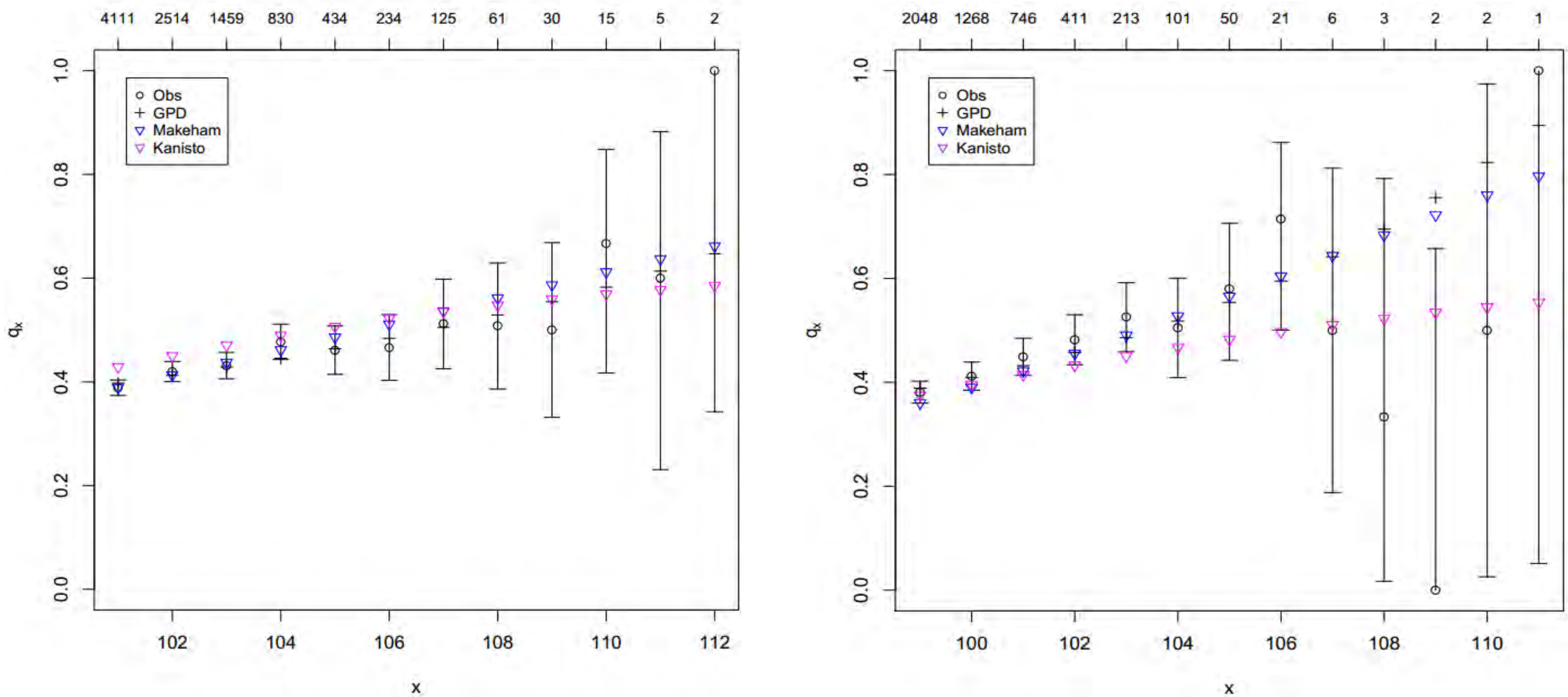


Figure 4 : Crude, GPD, Makeham and Kannisto one year mortality rate estimates for females (left panel) and males (right panel) above $x^{\star}$.

POINT-ESTIMATE OF HIGH QUANTILES

- As one can decompose death probability,

$$
\begin{aligned}
{}_{x}q_{0} &= 1 - {}_{x}p_{0} \\
&= 1 - {}_{x^\star}p_{0}\ {}_{x-x^\star}p_{x^\star} \\
&\approx 1 - {}_{x^\star}p_{0}\left(1 - G_{\hat{\xi},\hat{\beta}}(x - x^\star)\right)
\end{aligned}
$$

- It is easy to approximate $t(\epsilon)$ the $1-\epsilon$ quantile of the lifespan,

$$
\hat{t}(\epsilon) = x^\star + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{L_{95}}{L_{x^\star}}(1-\epsilon)\right)^{-\hat{\xi}} - 1\right).
$$

CONCLUSION

- EVT was applied on Belgian old individual data.
- ML estimation gives a negative tail index.
- This supports the existence of an ultimate age
    - 114.82 for males (very close to 115 from Dong et al. (2016)).
    - 122.73 for females (very close to 122.42 of Jeanne Calment).
- Those ages must not be interpreted as biological maximum lifespan.
- They can serve as upper bounds for actuaries to close mortality tables for actuarial computations.
- Mortality rates from different models agree for early old ages.
- But diverge as the size of sample decreases.
- Results show sensitivity to the chosen estimation procedure.
- As further research, the dynamic behavior of the different cohorts will be investigated.

REFERENCES

- Dong, X., Milholland, B., Vijg, J. (2016). Evidence for a limit to human lifespan. Nature, in press.
- Gampe, J. (2010). Human mortality beyond age 110. In *Supercentenarians*, pp. 219-230. Springer, Berlin.
- Gavrilov, L.A., Gavrilova N.S. (2011). Mortality measurement at advanced ages : A study of the social security administration death master file. *North American Actuarial Journal*, 15, 432-447.
- Gbari S., Poulain M., Dal L. and Denuit M. (2016). Extreme value analysis of mortality at the oldest ages : a case study based on individual ages at death. *ISBA Discussion paper submitted for publication*,1-25.
- Hanayama, N., Sibuya, M. (2015). Estimating the upper limit of lifetime probability distribution, based on data of Japanese centenarians. *The Journal of Gerontology : Series A*.
- Ouellette, N., Bourbeau, R. (2014). The trajectory of mortality at ages 100 and beyond : An analysis of individual level data in Canada. *Paper presented at the Annual Meeting of the Population Association of America*, Boston, May 1 - 2, 2014.
- Robine, J. M., Cournil, A., Gampe, J., Vaupel, J. W. (2005). IDL, the international database on longevity. *Living to 100 and beyond symposium monograph*, Society of Actuaries.

THANKS FOR YOUR ATTENTION !

EXTREME VALUE ANALYSIS OF MORTALITY AT THE OLDEST AGES : A CASE STUDY BASED ON INDIVIDUAL AGES AT DEATH

Samuel GBARI
samuel.gbari@uclouvain.be

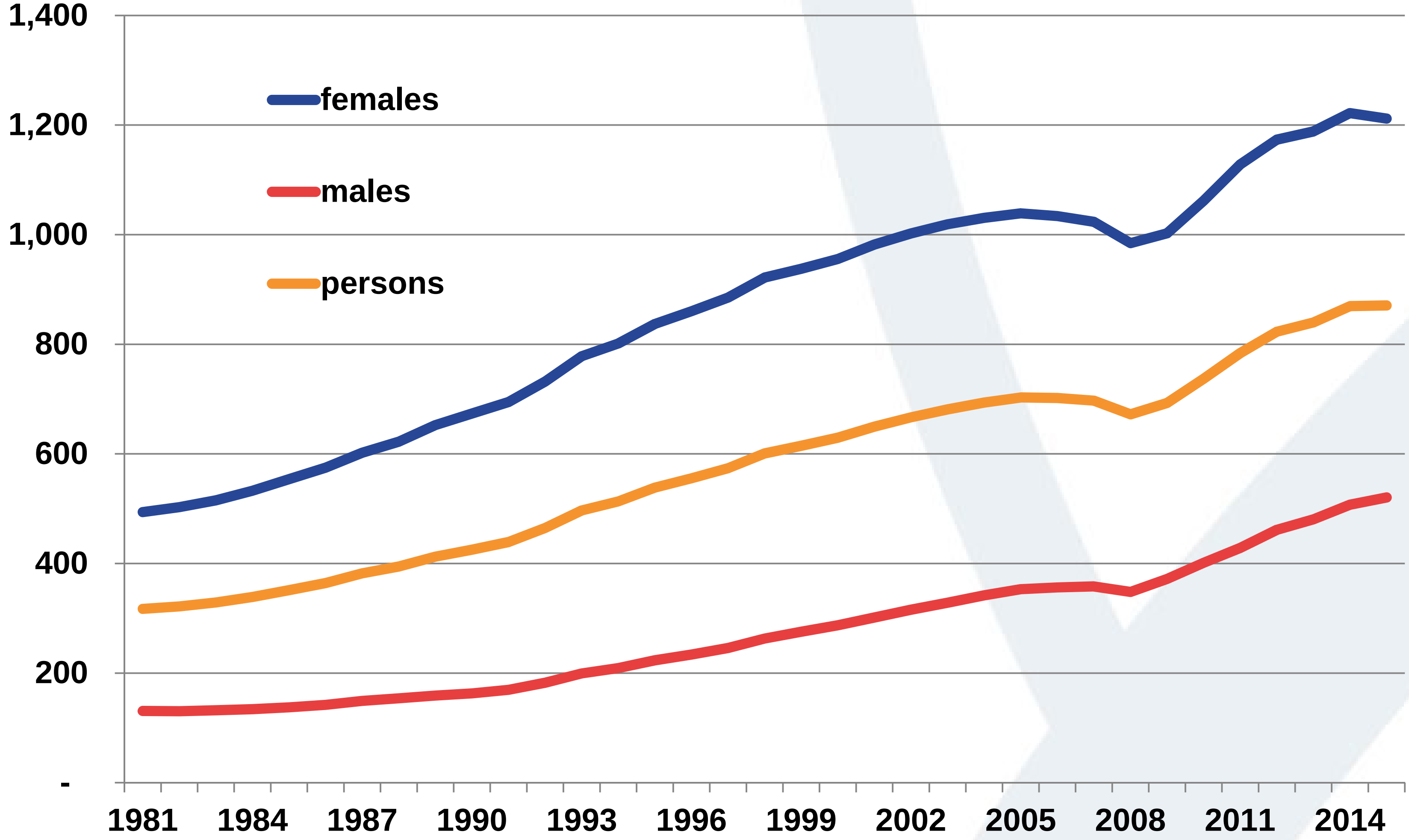
Joint work with M. Poulain, L. Dal and M. Denuit
Universite Catholique de Louvain (Belgium)
Institut de Statistiques, Biostatistiques et de Sciences Actuarielles

SOA Living to 100 Symposium (Orlando)
January 2017

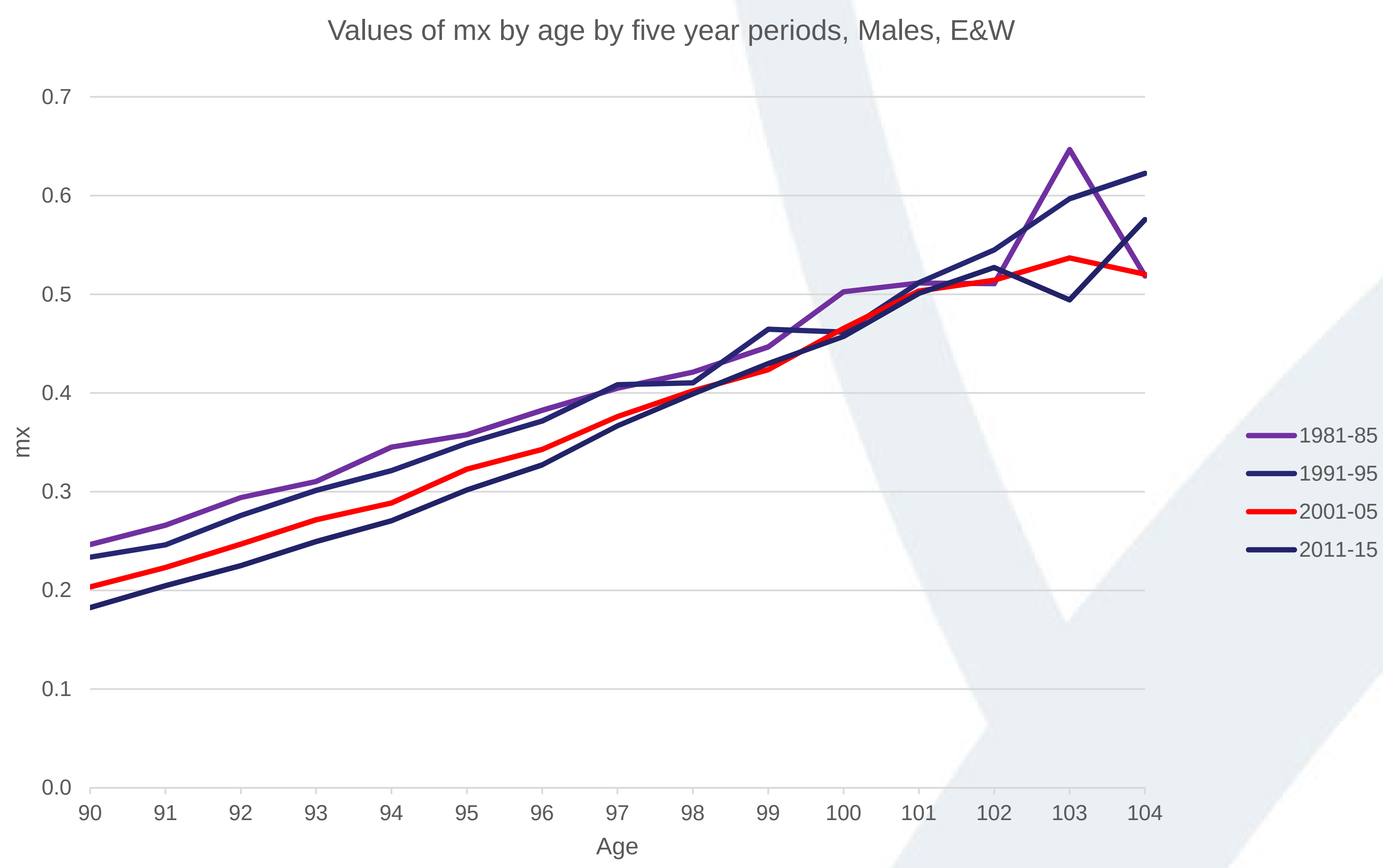
Population mortality at the oldest ages in England and Wales

Adrian Gallop and Angele Storey
Office for National Statistics
United Kingdom

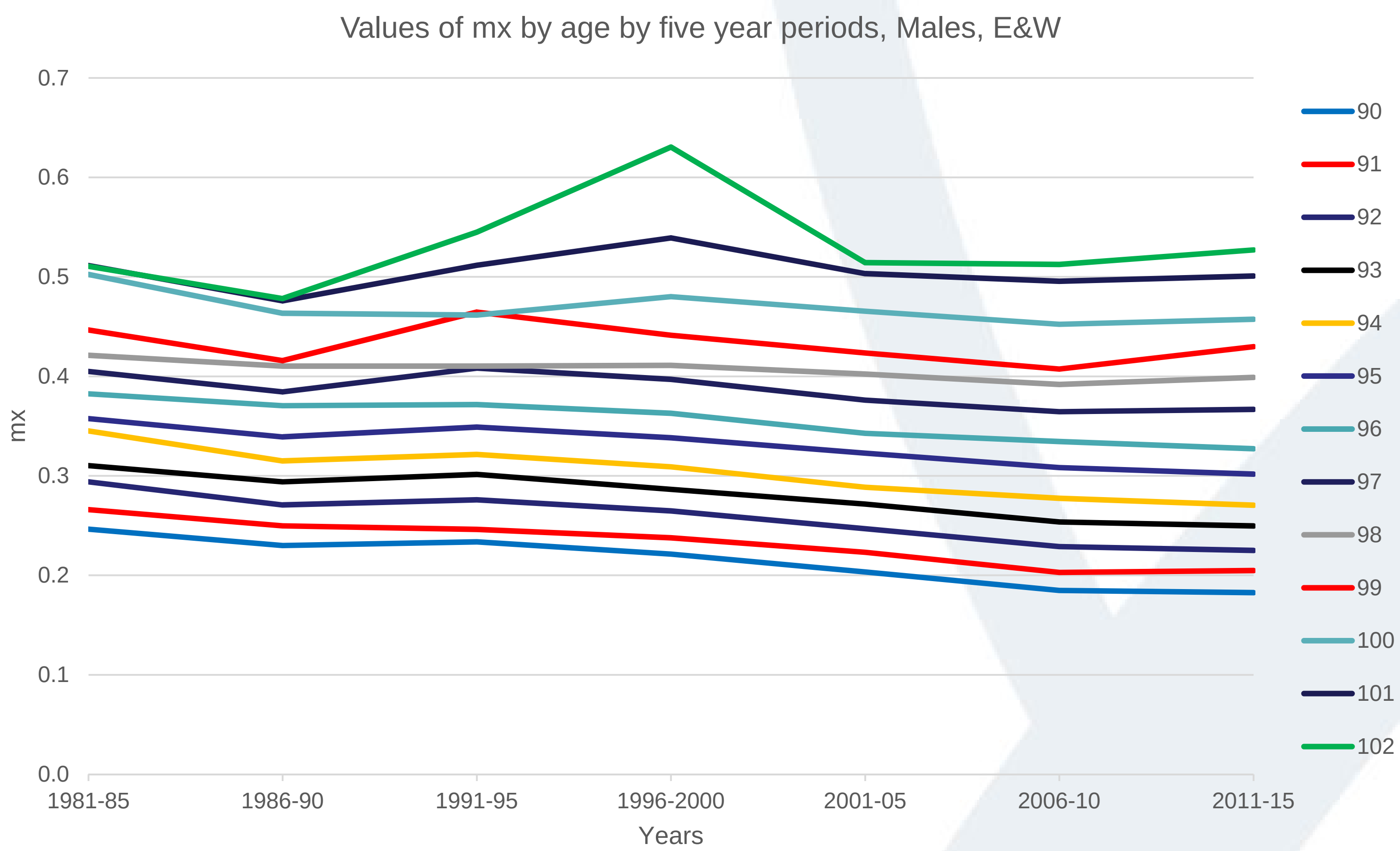
90 and over population per 100,000 population, England and Wales, 1981 to 2015



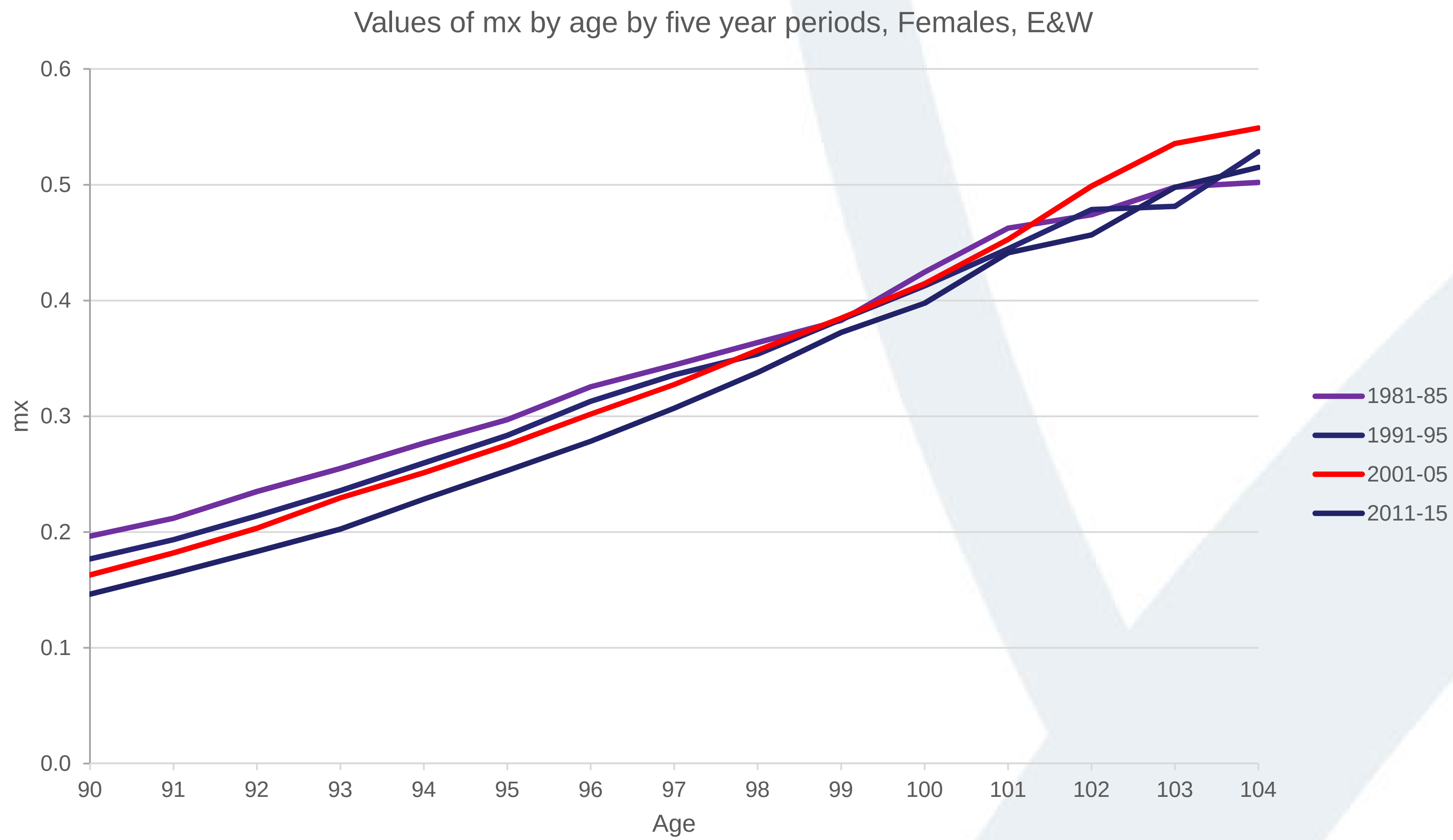
Male mx, E&W, 5-year groups



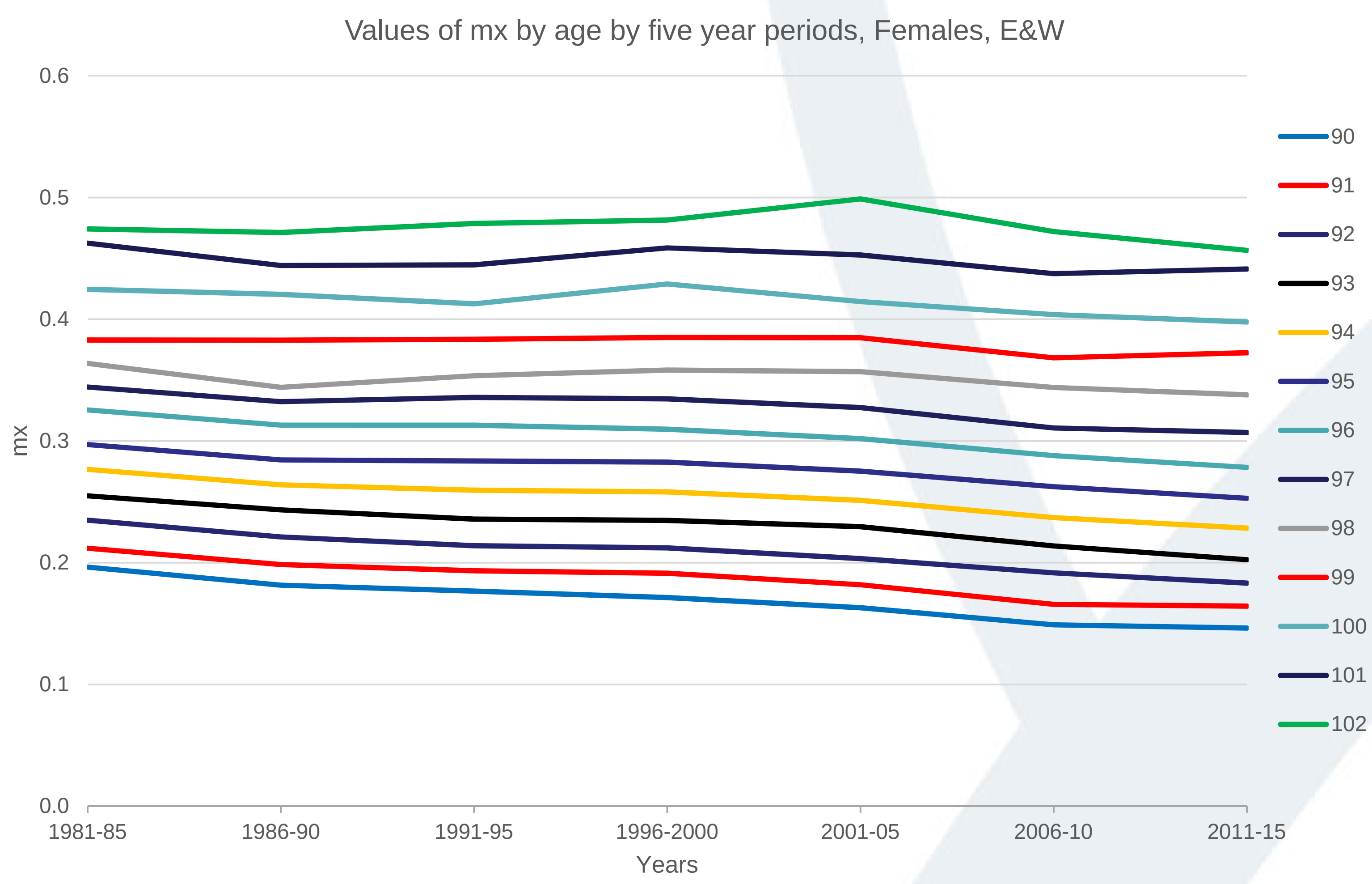
Male mx, E&W, 5-year groups



Female mx, E&W, 5-year groups



Female mx, E&W, 5-year groups



Drivers for the research

- Increasing numbers of people living to 90+ →accurate population estimates and mortality rates at high ages gaining importance
- 30,000 fewer people aged 90 and over in 2011 Census (E&W) than had been estimated by rolling forward the 2001 Census data to 2011 →Q:2011 Census?
- Assessment of quality of population and deaths data + methods used to construct population estimates at ages 90 and above

Input data to high age estimates

- Official MYE published by syoa & sex to age 89 and 90+
- MYE produced by cohort component method : census estimate aged on - deaths + births + net migration
- 90+ estimates by syoa and sex to age 104 and 105+; produced from deaths data using the KT method
- KTs estimates constrained to 90+ total in MYE before publication
- Quality of deaths data, census data and migration data for the oldest ages all impact quality of high age estimates

Age at death validation

- Informants asked to supply date of birth of the deceased when registering a death
- Age at death is derived from subtracting date of birth supplied by informant from date of death
- ONS do not validate age at death except deaths of supercentenarians (those aged 110 & over)
- University of Southampton funded age at death validation of a sample of semi-centenarians (those aged 105-109) (Validation being extended to those aged 90-104)

Age at death validation - sample

- 375 male deaths and 3,652 female deaths registered in E&W aged 105 to 110+ between 2000 to 2014

- 55 of the deceased males and 285 of the deceased females were born outside E&W or no place of birth recorded. These cases were excluded.

- All remaining males cases (320) plus a stratified sample of the remaining female cases (821) were included in the analysis

- Validated by matching dob on death certificate to dob on birth certificate of the deceased

Age validation results

- High validation rate where person born in E&W;
  - males 96% full validation; 98% incl. partial
  - females 96% full validation; 97% incl. partial

- 15% of male and 8% of female deaths registered during 2000-2014 in E&W born outside E&W
  - validation not possible
  - can't say if high validation rate would also apply

- Extrapolating results suggests up to 2,000 of 103,400 deaths aged 90 and over in 2014 could have incorrect age at death on death record

Completeness of deaths data

- Legal requirement for all deaths occurring in E&W to be registered (except Foreign Armed Forces)

- Deaths of residents of E&W that occur and registered outside of E&W not included (except HM Forces)

- ONS publish deaths data on a registration basis

- Annual extract of death occurrences
  - used for seasonal analysis
  - calculation of population estimates

Census data

- Main component of the 90+ official mid-year estimate is the rolled forward census estimate of the 90+ population
- Census estimates = census counts + imputation (to account for non-response)
- Research investigated sources of error other than non-response – misreporting/recording of dob by respondent/proxy respondent + possible processing/scanning errors

- Research conducted using the ONS-Longitudinal Study (ONS-LS)

ONS-LS

- ONS-LS is a 1.1% sample of the population of England and Wales
- Original sample selected from 1971 Census using 4 (undisclosed) birthdays
- Updated at each successive census
- Births, deaths, cancer registrations and immigration emigration events are linked
- Excludes imputed cases

Analysis

Study Sample

- All those in ONS-LS sample aged 80 and over at 2011 Census (27,307 cases)

Analysis

- Year of birth recorded at 2011 Census compared with
  - year of birth recorded in previous censuses (1971, 1981, 1991, 2001)
  - NHS Medical Research Information Service (MIDAS) record
  - cancer records (if applicable)
  - embarkation records (if applicable)

Analysis

Number of records of year of birth for sample

	Number	Percent
2	851	3.1
3	364	1.3
4	644	2.4
5	2,469	9
6	15,834	58
7	7,110	26
8	35	0.1
Total	27,307	100

Source: ONS Longitudinal Study

Results 1

- 93.5% of sample cases had no year of birth discrepancies across all their available records
- 1,781 cases had at least 1 discrepancy

2011 Census age	Number of cases with at least one discrepant record	Number in sample	Percent (of cases with at least one discrepant record)
80-84	827	14,239	5.8
85-89	586	8,646	6.8
90-94	241	3,466	6.9
95-99	106	842	12.6
100+	21	114	18.4
Total	1,781	27,307	6.5

Source: ONS Longitudinal Study

Analysis

Probable 'true' year of birth

- For cases with no year of birth discrepancies, recorded yob assumed to be true value
- For 1,781 cases with at least one discrepancy, modal year of birth was designated the 'true' year of birth
- For cases with same number of unique year of birth values as records, mode was designated as not applicable (61 cases)
- For cases with multi-modes (36 cases), 2011 census was designated as mode

Results 2

Probability that 2011 Census year of birth record is 'true' where 2011 Census year of birth is the mode

Probability	Number	Percent
Certain	25,526	94.7
0.7-0.9	1,337	5.0
0.5-0.6	73	0.3
less than 0.4	9	-
Total	26,945	100 .0

Source: ONS Longitudinal Study

Only 300 cases where 2011 Census year of birth was not the mode - 1.1% of sample

Type of discrepancies (where 2011 Census year of birth is not mode; 300 cases)

- Young children recorded on census as born in 20th Century rather than 21st
- Reversal of digits e.g.1902 in 2011 Census: Modal = 1920
- Problems capturing some written responses e.g. 1907 in Census: Modal =1967; 1912 in Census: Modal = 1972
- Cluster of cases found where year of birth was 1918 in the 2011 Census and modal year of birth was 1919 and vice versa

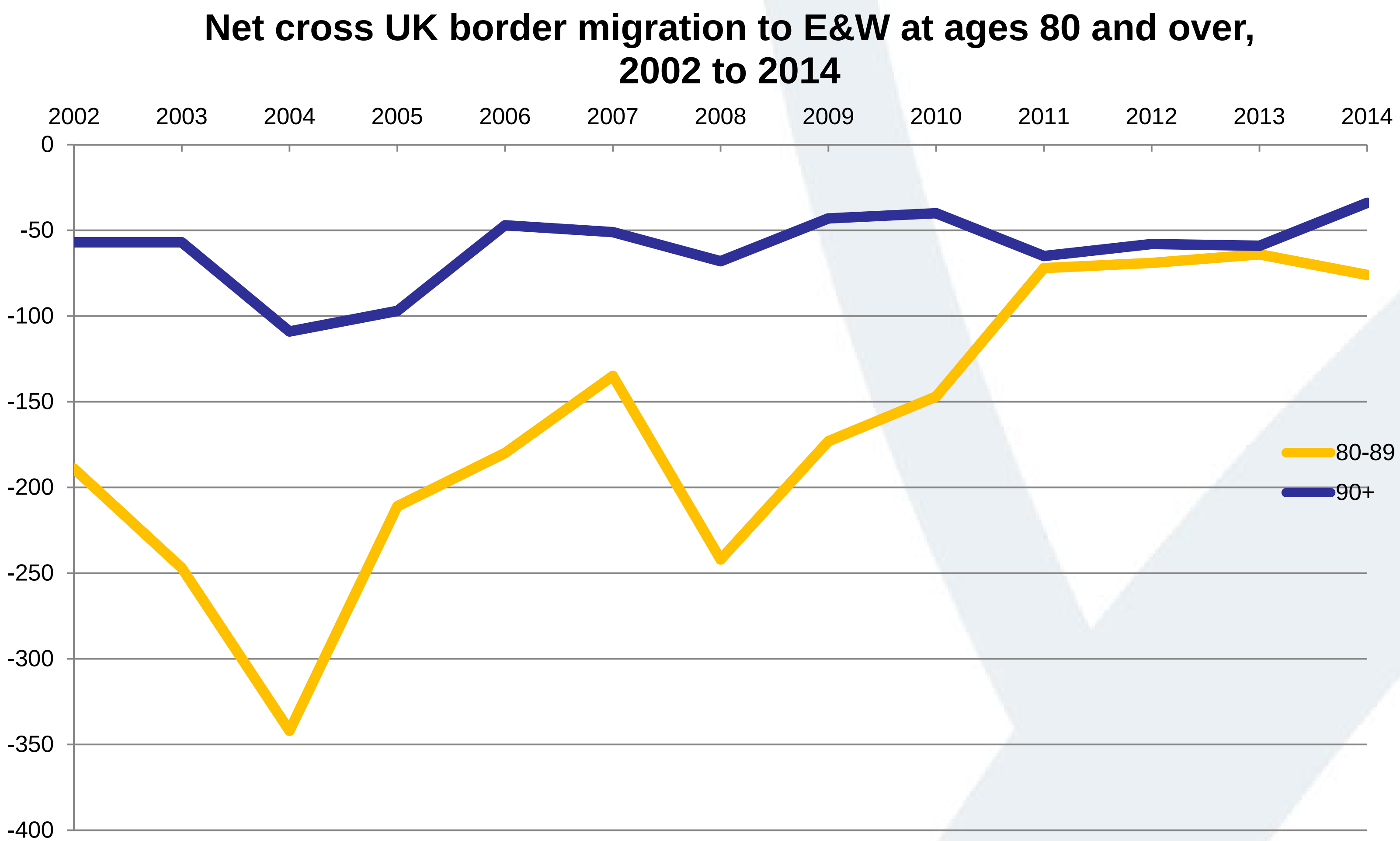
2011 Census

- Overall high degree of accuracy in birth data in 2011 census data for those aged 80 and over in ONS Longitudinal Study implying a high degree of accuracy in overall 2011 Census
- For 1.1% of cases, the 2011 Census year of birth not the modal year of birth
- Reliability of birth data decreased with age (for the 80 and over age group)

Migration data

- KT method makes assumption migration at the oldest ages is minimal
- KT estimate is constrained to the 90 and over total in the official mid year estimates (MYE)
- 2 migration components in MYE
  - net internal migration
  - net international migration

What is the level of UK cross-border migration at the oldest ages?



International migration component of MYE

- Based on International Passenger Survey (IPS)
- sample survey; migrant respondents scaled to produce national migration estimates

a) International immigration component of MYE
- 2011 Census question on address one year ago used to obtain age/sex distribution of immigrants
- distribution is applied and constrained to the IPS national long-term international in-flow total

b) International emigration component of MYE
- from IPS national long-term international migration out-flow total
- age/sex distribution based on current and previous 2 years of data

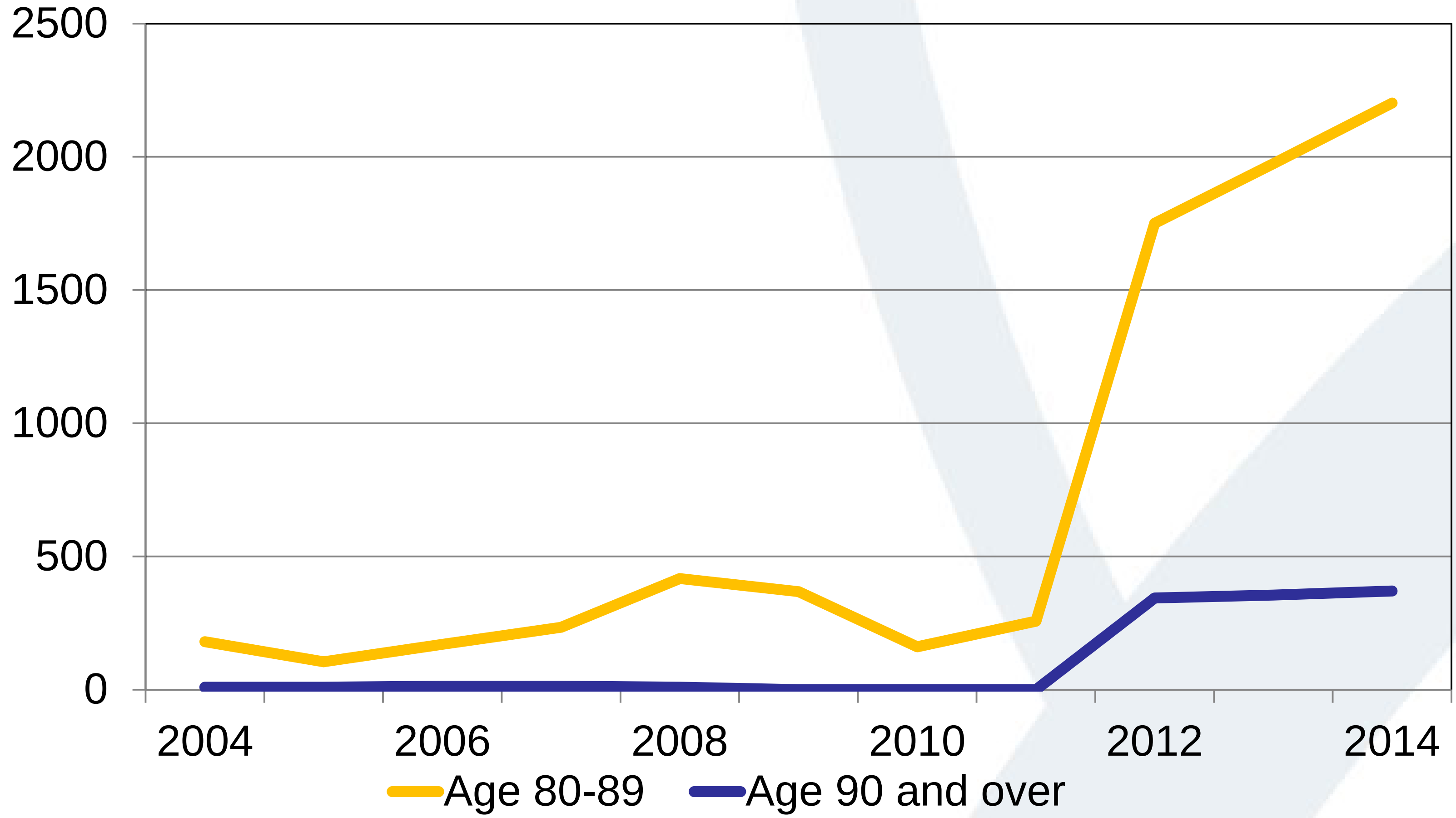
Quality of international migration estimates at the oldest ages

- Flows of international migrants to and from E&W too small to be reliably measured in IPS

- Less of a problem for immigration estimate because 2011 Census age/sex immigration distribution is applied to the total IPS figure

- Emigration estimates arguably the weakest component of the MYE of people aged 80 and over

International immigration at the oldest ages, 2004 - 2014



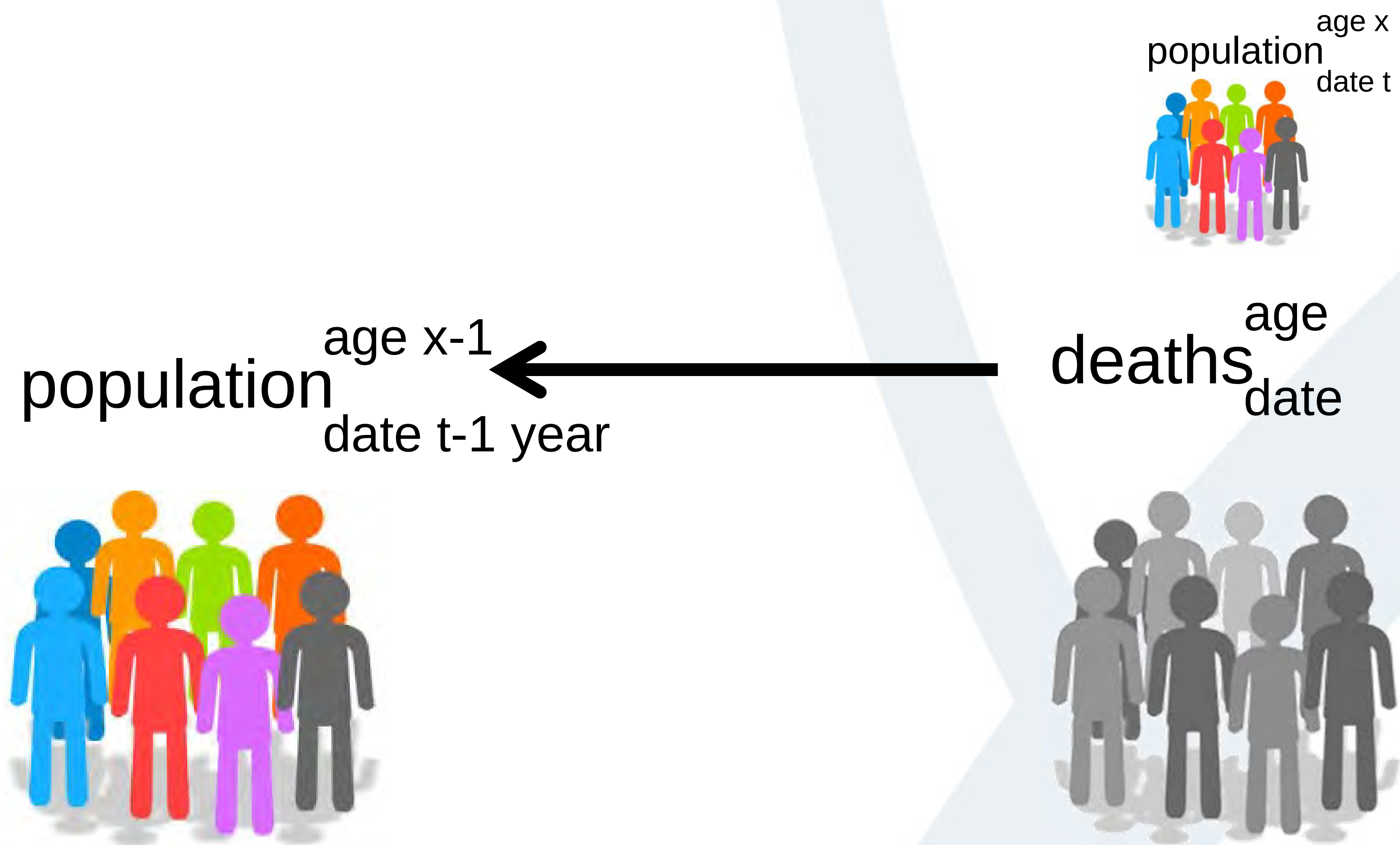
Net migration component of population change in MYE, E&W, 2013 to 2014

Age group	Change in population 2013/2014	Net UK cross border migration	Net international migration	Total net migration
80-89	33,687	-76	2,142	2066 (6.1%)
90 and over	21,013	-34	364	330 (1.6%)

Migration data component of MYE: Summary

- Quality of UK cross-border migration data at the oldest ages appears reasonably good
- Quality of international immigration data at oldest ages appears reasonable; however emigration data is likely to be of poorer quality
- Levels of both UK cross border migration and international migration are very low at the oldest ages i.e. migration has minimal impact on annual MYE of oldest old
- KT method assumption of minimal migration holds

Kannisto-Thatcher (KT) method



Kannisto-Thatcher (KT) method (as used by ONS)

- At high ages and for dates sufficiently far in the past age specific population estimates can be obtained directly from deaths data

- Once all the members of a given birth cohort have died numbers alive at earlier dates reconstructed from their dates of birth and death

Kannisto-Thatcher (KT) method

- For cohorts which are almost extinct the ratio of the number of survivors who are still alive to the numbers in the cohort who died in the previous *k* years can be estimated from the experience of previous cohorts
- This estimated survivor ratio can then be applied to the known number of deaths in the given cohort which occurred over the last *k* years

Kannisto-Thatcher (KT) method

- Assumes survivor ratio is the same as for the immediately preceding cohort
- However, may be an atypical cohort
- To dampen fluctuations in the ratios calculate the average survivor ratio over the preceding *m* cohorts, rather than just a single cohort

Kannisto-Thatcher (KT) method

- Kannisto and Thatcher proposed various modifications to allow for mortality rates changing over time
- Reducing mortality may increase the size of the survivor ratio over time; apply a correction factor to the survivor ratios calculated
- This can be set to constrain the estimates to sum to the official population estimate for a given age group (say, 90 and over, as in the ONS method)
- or so that the estimates join to the official estimates in a specific way.

Kannisto-Thatcher (KT) method

- For all years prior to the year

$$P_x^t = P_{x+1}^{t+1} + D_x^t$$

where

P is population ag ex at beginning of the year,

D is deaths age x at beginning of the year,

x is age, and

t is year

Kannisto-Thatcher (KT) method

- For the year under consideration

$$P_x^T = \left(D_{x-1}^{T-1} + D_{x-2}^{T-2} + D_{x-3}^{T-3} + D_{x-4}^{T-4} + D_{x-5}^{T-5}\right) \times S_x^T \times c$$

where S is survivor ratio for age x in year t
and c is a correction factor

Kannisto-Thatcher (KT) method

- ONS use data for 5 cohorts (*m* = 5) and going back for 5 years (*k* = 5) so the survivor ratio is calculated as:

$$S_x^T = \sum_{T-1}^{T-5} P_x \Bigg/ \left(\sum_{T-2}^{T-6} D_{x-1} + \sum_{T-3}^{T-7} D_{x-2} + \sum_{T-4}^{T-8} D_{x-3} + \sum_{T-5}^{T-9} D_{x-4} + \sum_{T-6}^{T-10} D_{x-5} \right)$$

Kannisto-Thatcher (KT) method

Assumptions:

- Deaths data accurate
- Death data available for 12 month periods by age at start of 12 month period
- No migration

ONS Kannisto-Thatcher (KT) method

- ONS population estimates are as at 30 June
- Need deaths data in mid-year to mid-year periods by age at start of period
- Published data is deaths in a calendar year by age at death
- ONS approach is to calculate pop estimates at 1 Jan using KT method and interpolate to get mid-year pop estimates

ONS Kannisto-Thatcher (KT) method

- Working with 1 Jan population estimates need deaths data in a calendar year by age at the start of the calendar year
- We have deaths in calendar year by age at death
- It is assumed half the people died before their birthday in the year, and half after their birthday
- If number of deaths aged x in calendar year Y = D(x,Y) then number of deaths in calendar year Y aged x at 1 Jan Y is assumed to be

$$0.5 \times [D(x,Y)+D(x+1,Y)]$$

- This may not be accurate, especially for years where deaths are not evenly spread over a year
- Or birth cohorts not evenly spread

Analysis idea

- Q1 – How well does the KT method work when data is in the required format?
- Q2 – How well does the KT method work when data has to be manipulated into the required format?

Method: compare KT against a benchmark

What data needed

- Reliable comparator estimates
- Death by age at reference day

- We obtained population register data from 2 countries, Finland and Sweden

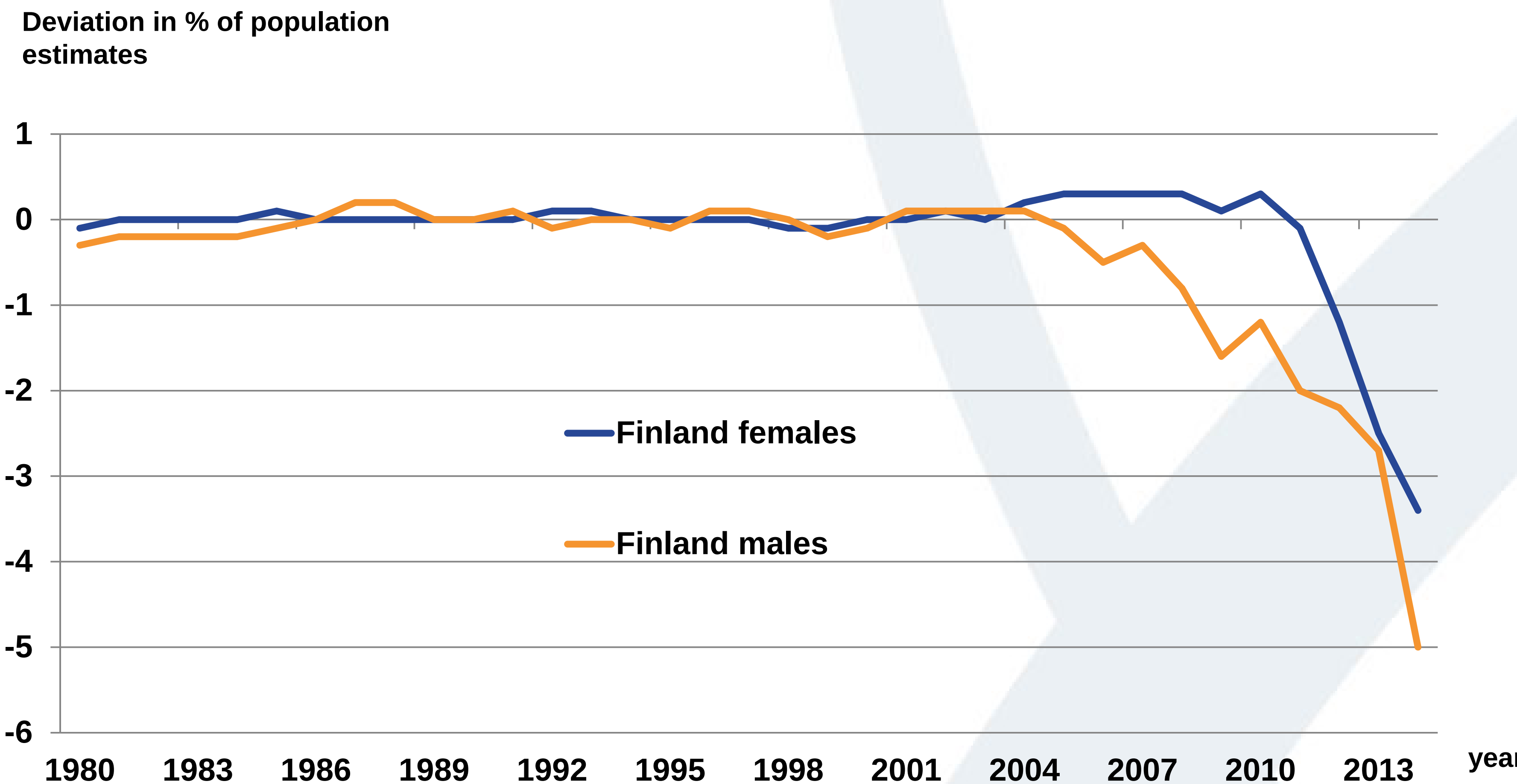
- **Finland only reported in the presentation**

Metadata

DEATHS	FINLAND
Reference date	1st January
Period	1980 - 2014
Top age	112

POPULATION	FINLAND
Reference date	1st January
Period	1980 - 2014
Top age	100+

Fit with deaths data in required format



Fit by age & year

Finland

Deviation in % of population estimates

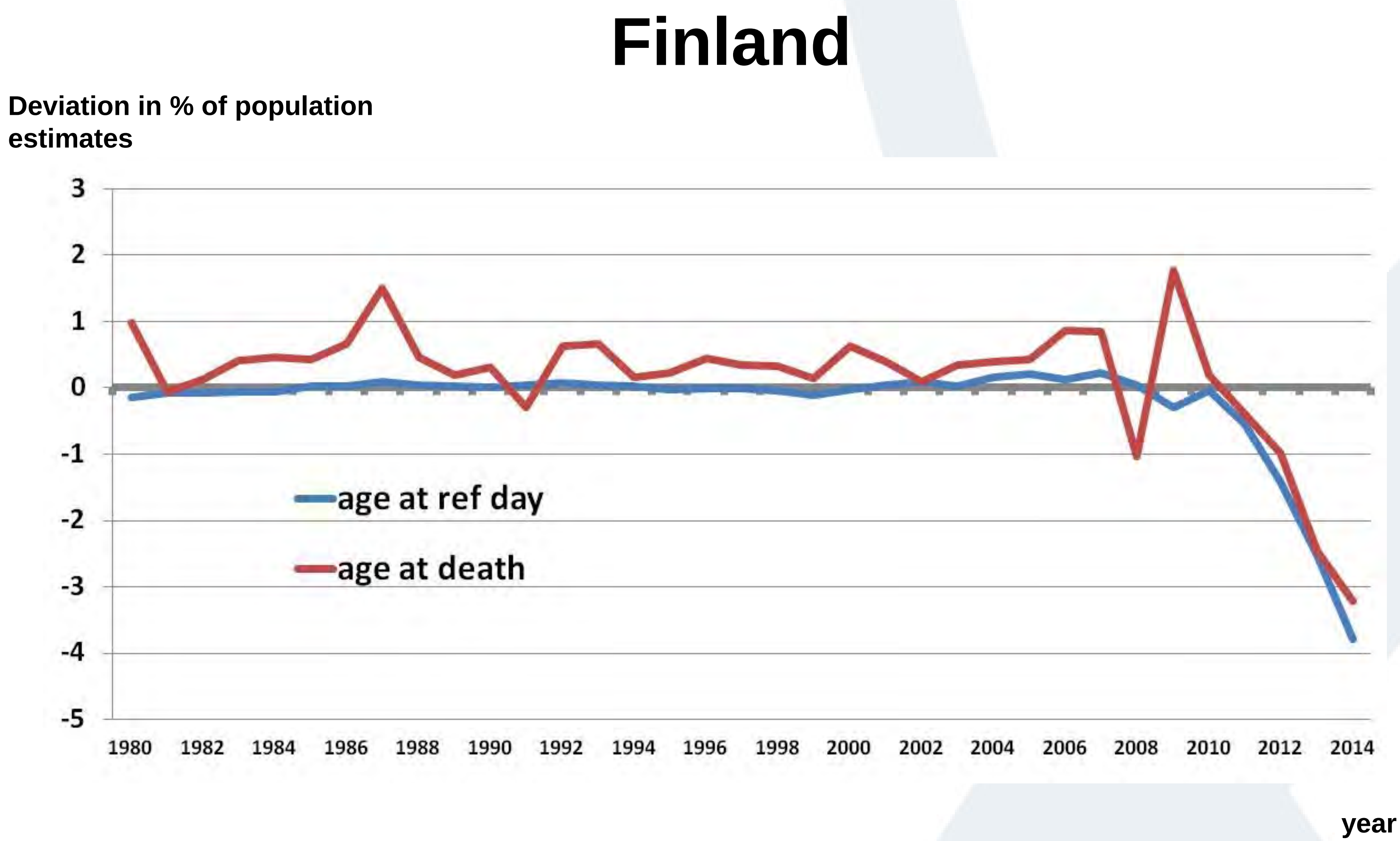
year

age\year	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
85	0.0	0.2	0.2	0.2	0.0	0.3	0.2	-0.1	0.4	-0.3	-0.7	0.6	-1.1	-1.9	-2.4	-3.0	-5.3	-5.2	-6.7	-11.2	-11.0
86	0.1	-0.1	0.2	0.3	0.2	0.0	0.3	0.3	-0.2	0.4	-0.3	-0.8	0.6	-1.2	-2.1	-2.7	-3.4	-5.9	-5.8	-7.4	-12.2
87	0.1	0.1	-0.1	0.2	0.2	0.2	0.0	0.3	0.3	-0.2	0.4	-0.4	-1.0	0.6	-1.4	-2.3	-3.1	-3.8	-6.6	-6.4	-8.2
88	0.1	0.0	0.1	-0.1	0.2	0.2	0.1	0.0	0.4	0.3	-0.2	0.4	-0.5	-1.1	0.7	-1.6	-2.6	-3.4	-4.3	-7.4	-7.3
89	0.1	0.1	0.0	0.0	-0.2	0.2	0.2	0.1	-0.1	0.4	0.3	-0.3	0.4	-0.6	-1.3	0.8	-1.8	-3.0	-4.0	-4.9	-8.5
90	-0.1	0.0	0.1	-0.1	0.0	-0.2	0.3	0.3	0.1	-0.1	0.4	0.2	-0.3	0.5	-0.7	-1.5	0.9	-2.1	-3.5	-4.6	-5.7
91	0.2	-0.1	0.0	0.1	-0.1	0.0	-0.2	0.3	0.4	0.1	-0.1	0.5	0.3	-0.4	0.5	-0.9	-1.9	1.1	-2.5	-4.2	-5.5
92	0.1	0.1	-0.1	0.0	0.0	-0.1	-0.1	-0.2	0.3	0.5	0.2	-0.2	0.5	0.3	-0.5	0.6	-1.1	-2.3	1.4	-3.1	-5.1
93	0.1	0.0	-0.1	0.0	0.0	0.0	-0.2	-0.1	-0.3	0.4	0.6	0.2	-0.2	0.7	0.4	-0.6	0.8	-1.3	-2.9	1.7	-3.9
94	0.0	0.0	-0.2	-0.1	-0.1	0.1	0.0	0.0	-0.2	-0.4	0.4	0.8	0.1	-0.3	0.9	0.6	-0.8	1.2	-1.6	-3.8	2.2
95	0.0	-0.1	0.0	0.0	-0.3	-0.1	0.0	-0.1	-0.1	-0.3	-0.4	0.6	1.1	0.2	-0.4	1.3	0.7	-0.9	1.5	-2.2	-4.8
96	-0.2	-0.2	0.0	0.0	-0.1	-0.3	0.0	0.1	-0.2	-0.2	-0.3	-0.4	0.8	1.5	0.3	-0.6	1.6	1.1	-1.2	2.0	-2.8
97	-0.3	-0.3	-0.2	0.0	0.0	-0.6	-0.4	-0.2	0.0	-0.3	-0.1	-0.3	-0.6	1.1	2.0	0.3	-0.9	2.5	1.4	-1.7	3.2
98	0.0	0.0	0.0	0.0	0.0	0.0	-0.6	-0.6	0.0	0.0	-0.7	-0.2	-0.4	-0.9	1.5	3.0	0.4	-1.0	3.5	2.1	-2.3
99	0.6	0.0	0.0	0.0	0.0	0.0	0.0	-1.3	-0.9	0.0	0.0	-0.9	-0.3	-0.6	-1.3	2.3	4.2	1.0	-1.8	4.9	3.2
100 +	0.0	0.5	0.0	0.0	0.0	-0.4	-0.7	-0.4	-0.6	-1.2	-1.1	-1.2	-1.5	-0.7	-0.8	-1.6	0.2	3.5	4.4	3.5	7.8

Analysis idea

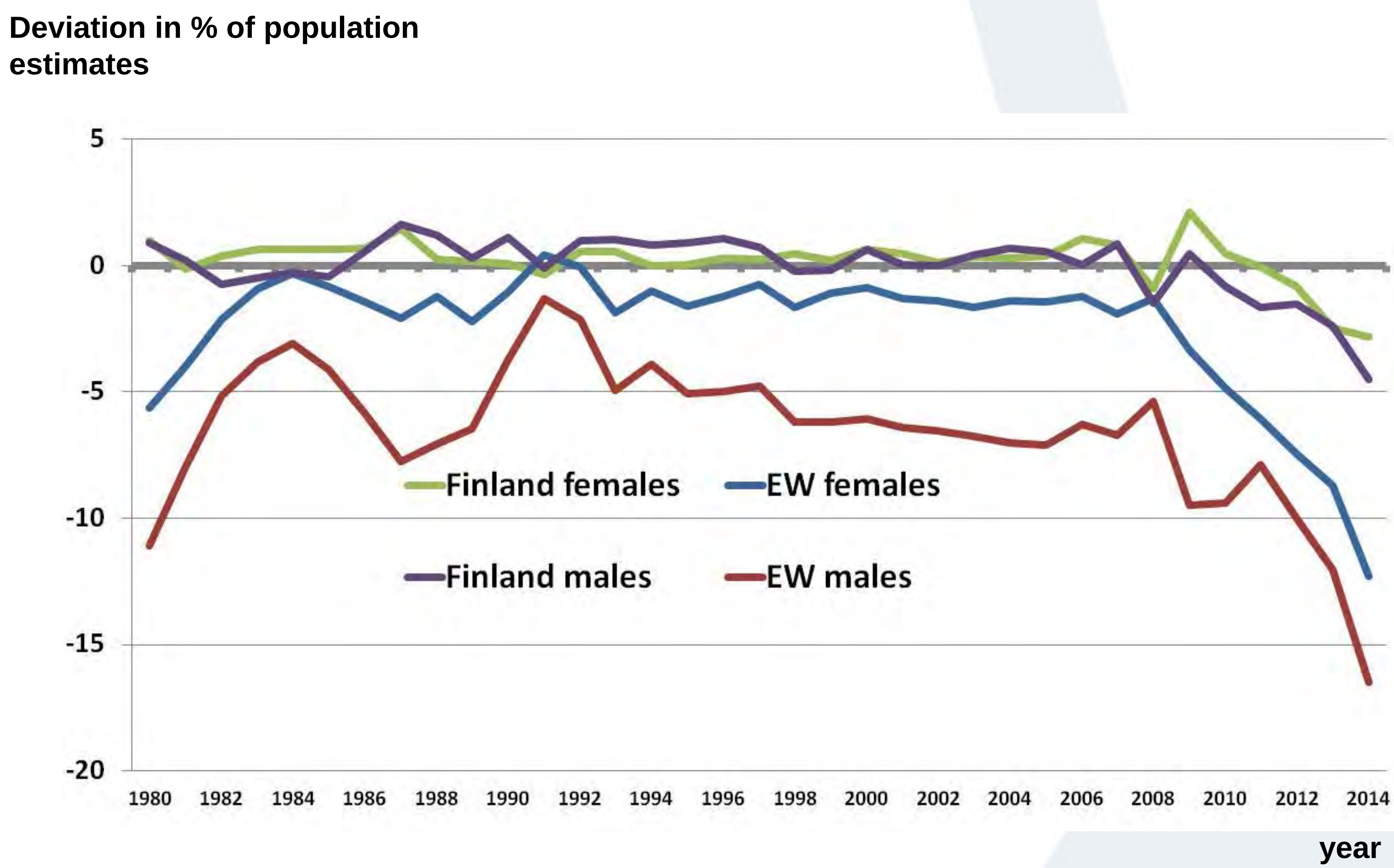
- Q2 – How well does the KT method work with when deaths data have to be manipulated into the required format?

Differences between KT estimates derived from deaths data on different definitions and population register estimates: Finland



Percentage differences between KT estimates and official estimates: E&W and Finland, 1980 to 2014

Finland vs ONS



Accuracy

Conclusions

- The KT method works well where
  - data are in the correct format
  - mortality rates at the oldest ages are not changing over time
- The greater the changes in mortality rates at the oldest ages the greater the correction factor required
- Where mortality rates at oldest ages are changing over time a different method for allowing for mortality improvement may produce better estimates

Comparison of official estimates of the 90+ population and estimates derived from administrative data sources

	MALES		FEMALES	
	SPD V1.0	CENSUS/ MYE	SPD V1.0	CENSUS/ MYE
2011	109,286	114,506	304,082	314,511
2013	127,398	134,797	330,165	343,420
2014	134,507	143,536	339,860	355,694
2015	138,627	148,630	338,499	355,400

SPD V1.0 are population estimates derived from administrative data; they are not official statistics

Quality of input data – Key findings

- Legal requirement for all deaths occurring in E&W to be registered; deaths data taken as complete
- Age at death data very high quality in terms of accuracy for those dying at ages 105 and over in E&W who were born in E&W – validation work being extended
- High degree of accuracy in year of birth data for the sample of 2011 Census records of those aged 80 and over; however decreases with age
- Migration at the oldest ages is very low and has minimal impact on the annual MYE of those aged 90 and over

Further research identified

- Extend death validation to those dying at ages 90 - 104
- Model potential impact of differing levels of migration on high age estimates
- Investigate ways to improve the emigration estimation at the oldest ages as a component of change in the mid-year estimates
- Consider impact of changing mortality rates over time on KT estimates and investigate different methods for allowing for these (e.g. incorporating trends in survivor ratios)
- Run the model populations using different averaging periods (k and m) for calculating the survivor ratio

Any questions?

More details of the research are available in the report :

‘Accuracy of official high age population estimates: An evaluation’

available at https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/methodologies/accuracyofofficialhighagepopulationestimatesinenglandandwalesanevaluation

Additional slides not used in the presentation

Age at death validation
- Males aged 105-110+

All deaths registered in E&W who were born in E&W							
Age at death	Number registered	Number fully validated*	Number partially validated#	Number incorrect	Number not validated	Percentage fully validated*	Percentage fully or partially validated
105	190	186	1	1	2	97.9%	98.4%
106	63	58	2	1	2	92.1%	95.2%
107	36	35	1	0	0	97.2%	100.0%
108	23	22	0	0	1	95.7%	95.7%
109	3	2	1	0	0	66.7%	100.0%
110+	5	5	0	0	0	100.0%	100.0%
Total	320	308	5	2	5	96.3%	97.8%

* Fully validated cases are those where a birth record was found which matched the same date and place of birth and the same full name (and aliases).

Partially validated cases are those which appear to have a transcription error which does not invalidate the age at death.

Age at death validation
- Females aged 105-110+

	Validation of sample of all deaths registered in E&W for those born in E&W							
Age at death	Number registered	Number fully validated*	Number partially validated#	Number possibly correct	Incorrect age	Number not validated	Percentage fully validated*	Percentage partially validated
105	162	154	1	3	0	4	95.1%	95.7%
106	124	121	1	2	0	0	97.6%	98.4%
107	155	149	0	0	0	6	96.1%	96.1%
108	170	159	2	1	2	6	93.5%	94.7%
109	128	124	3	0	0	1	96.9%	99.2%
110+	82	80	1	0	0	1	97.6%	98.8%
Total	821	787	8	6	2	18	95.9%	96.8%

* Fully validated cases are those where a birth record was found which matched the same date and place of birth and the same full name (and aliases).

Partially validated cases are those which appear to have a transcription error which does not invalidate the age at death.

Registrations v. Occurrences

- Death occurrences:
  - Advantage: deaths assigned to year of occurrence (Accuracy)
  - Disadvantage: Late death registrations may not be included in data (Completeness)

- Death registrations:
  - Advantage: (Historically) all deaths captured (Completeness)
  - Disadvantage: Some deaths registered in a particular year will have occurred in previous years (Accuracy)

Internal migration component of MYE

- At E&W level, internal migration = UK cross-border flows
- Based on Patient Register (PR) and National Health Service Central Register (NHSCR) data
- PR and NHSCR identify people who have moved across internal UK borders since previous year
- Total flows to and from constituent countries are agreed between ONS, NRS and NISRA based on records of in-migration to the relevant country

Quality of cross-UK border migration estimates at the oldest ages

- Some people don't register promptly with a GP after they move or may never register

- Time lag in processing death registrations

Cross UK border migration, NHS/PR data, mid-2011			
	Age group	In-migration; to E&W	Out-migration; from E&W
Northern Ireland	80-89	53	48
	90 & over	6	15
Scotland	80-89	442	529
	90 & over	89	145

Cross UK border migration, 2011 Census data, March 2011			
	Age group	In-migration; to E&W	Out-migration; from E&W
Northern Ireland	80-89	47	59
	90 & over	10	16
Scotland	80-89	423	419
	90 & over	100	119