

Calculating population estimates of the very old

Methods used to produce the population estimates of the very old (aged 90 years and over) by single year of age and sex, UK, England, and Wales.

Contact:
Vasita Patel
pop.info@ons.gov.uk
+44 (0)1329 444661

Release date:
25 September 2019

Next release:
To be announced

Table of contents

1. [Introduction](#)
2. [Methodology](#)
3. [Calculations](#)
4. [Useful links](#)

1 . Introduction

The mid-year population estimates published by the Office for National Statistics (ONS) include estimates by single year of age up to 89 years, with a final category for ages 90 years and over.

Until 2007, population estimates for single years of age beyond 90 years were calculated for England and Wales by the ONS (and previously the Government Actuary's Department) solely for use in compiling national life tables and in producing the national population projections. These estimates were made available for research purposes but were not officially published.

Interest in population estimates at the oldest ages by single year of age has grown in line with increases in life expectancy and the growth in the number of centenarians. In recognition of this, the ONS began to publish these estimates in 2007 as Experimental Statistics. Since 2011, these estimates have been published as National Statistics.

Since 2016 (2002 to 2015 estimates) the statistics provide estimates of the population aged 90 years and over by single year of age and sex in the UK, England and in Wales. Prior to this the estimates were for England and Wales combined. This change was made in response to user demand. These separate country estimates have gone through both internal and external quality assurance processes.

Population estimates of the very old are constructed using a version of the [Kannisto-Thatcher \(KT\) model](#) of population at advanced ages. They are consistent with the official mid-year population estimates and death registrations statistics.

Estimates of the very old are produced by the Centre for Ageing and Demography within the ONS. The estimates are published annually; the tables provide mid-year population estimates by sex and single year of age for 90 to 104 years and for the 105 years and over age-group.

Quality and methodology information reports for the [mid-year population estimates](#) and [estimates of the very old](#) are available.

Equivalent estimates of the 90 years and over population by single year of age and sex for Scotland and Northern Ireland, are produced and published by [National Records of Scotland \(NRS\)](#) and the [Northern Ireland Statistics and Research Agency \(NISRA\)](#). They are used in the production of the UK estimates of the very old, which are an aggregation of the individual UK country estimates.

2 . Methodology

Population estimates of those aged 90 years and over by single year of age are constructed using the [Kannisto-Thatcher \(KT\) model](#), which is a version of survivor ratio methodology. The KT method produces age-specific estimates of population at older ages using deaths data.

At high ages and for dates sufficiently far in the past, historic age-specific population estimates can be obtained directly from deaths data. Once all the members of a given birth cohort have died it is possible to reconstruct the numbers who were alive at earlier dates from their dates of birth and death.

For cohorts that 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. The past population for this cohort can then be recreated by adding back the deaths. If the highest age x , at which there is expected to be a survivor, is known, the whole process can be repeated to obtain survivor ratios to estimate the numbers aged $x-1$, then $x-2$ and so on, in an iterative process.

Applying this method directly assumes that the survivor ratio is the same as that in the immediately preceding cohort. However, this may be an atypical cohort for various reasons. In order to dampen fluctuations in the ratios, the average survivor ratio over the preceding m cohorts can be calculated, rather than just a single cohort.

In circumstances where mortality rates are changing over time, or where estimates are required down to ages as low as 90 years, Kannisto and Thatcher proposed various modifications to the survivor ratio method. To compensate for the fact that reduced mortality at higher ages may increase the size of the survivor ratio over time, a correction factor is applied 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 (for example, 90 years and over) or so that the estimates join to the official estimates in a specific way.

One consequence of this method is that each year the estimates for earlier years become more accurate as more death data becomes available to inform the age profiles.

Current Office for National Statistics (ONS) methodology

The current ONS methodology follows the KT method described previously, with values $k=5$ and $m=5$ and a constraint that the total estimates derived for the most recent year being estimated sum to the official population estimate for those aged 90 years and over for that year.

To carry out the calculations, deaths data needs to be in the format of deaths during the annual mid-year to mid-year period by age at the start of the period.

Up until the reference year 2018, input deaths data for both England and Wales were obtained by calendar year by age at death. However, official population estimates are published at a mid-year date. This required two adjustments: adjusting the deaths data to age at the start of the calendar year, rather than age at death, and adjusting the resulting KT estimates from 1 January to mid-year date.

Since the reference year 2018 it has been possible to obtain historical deaths data for England and for Wales in the required format, that is, by age at the start of the mid-year to mid-year period. Deaths data are obtained directly from the ONS Vital Statistics Life Event system. This enables deaths by date of occurrence to be assigned to the relevant mid-year to mid-year period and age at the beginning of that period to be assigned based on date of birth and date of death.

3 . Calculations

The method as used by the ONS can be expressed as follows.

For all years prior to the year that is being calculated (where P_x^t is the population aged x at the beginning of the mid-year to mid-year period starting in calendar year t , D_x^t is the number of deaths during the mid-year to mid-year period age x at the beginning of the year, where x is age and t is the calendar year in which the start of the mid-year to mid-year period falls):

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

For the year under consideration (where c is the correction factor, S is the survival ratio and T is the 'current' mid-year):

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

where S is calculated as:

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

The calculations are performed sequentially for single years of age, starting with the oldest age beyond which none is assumed to survive, taken to be 120 years.

The value of the correction factor c is derived such that:

$$\sum_{x=90}^{x=120} P_x^T = E^T$$

where E^T is the "official" population estimate of the population aged 90 years and over at 1 July in year T .

The methodology produces single year of age estimates at 1 July in year T and earlier years. Rating factors are then applied to the derived mid-year estimates for earlier years so that they also total to the official mid-year estimates for those years (for a given year the same rating factor is applied to the estimates derived for all ages 90 years and over).

For more information on the Kannisto-Thatcher methodology, see [The survivor ratio method for estimating numbers at high ages](#).

4 . Useful links

[Research on accuracy of high age estimates – update \(Word, 29.2Kb\)](#)

[Quality and Methodology Information for Mid-year National and Local Authority Population Estimates](#)

[Estimates of the very old, including centenarians, QMI](#)

[Comparison paper: 90 and over single year of age and sex population estimates produced by ONS, NRS and NISRA](#)