

Estimates of the very old, including centenarians, QMI

Quality and Methodology Information report for annual mid-year population estimates for the very old, detailing the strengths and limitations of the data, methods used, and data uses and users.

Contact:
Ageing Analysis team
pop.info@ons.gov.uk
+44 1329 444661

Release date:
17 May 2024

Next release:
To be announced

Table of contents

1. [Output information](#)
2. [About this QMI report](#)
3. [Important points](#)
4. [Quality summary](#)
5. [Quality characteristics of the data](#)
6. [Methods used to produce the data](#)
7. [Related links](#)
8. [Cite this methodology](#)

1 . Output information

- National Statistic: yes
- Frequency: annual
- How compiled: based on third party data
- Geographic coverage: UK, England and Wales

2 . About this QMI report

This Quality and Methodology Information report contains information on the quality characteristics of the data (including the European Statistical System's five dimensions of quality) as well as the methods used to create it.

The information in this report will help you to:

- understand the strengths and limitations of the data
- learn about existing uses and users of the data
- understand the methods used to create the data
- decide suitable uses for the data
- reduce the risk of misusing data

3 . Important points

- Estimates of the very old are mid-year population estimates (MYEs) by single year of age and sex for ages 90 to 104 years and for the 105 years and over age group; estimates are rounded to the nearest 10.
- We produce these estimates annually for England, Wales and the UK; comparable statistics for Scotland and Northern Ireland are produced by [National Records of Scotland \(NRS\)](#) and the [Northern Ireland Statistics and Research Agency \(NISRA\)](#), respectively, and are used in our UK estimates.
- The 2002 to 2022 estimates are for England and Wales only; figures for the UK, including a full data back series to 2002, will be published in autumn 2024.
- For the most accurate data, use the most recent version of estimates of the very old, even if you wish to use estimates for a previous year, as the accuracy of the estimates in previous years improves with each successive version.
- Each version of estimates of the very old includes estimates going back to 2002.
- Estimates of the very old are calculated from death registration data using the Kannisto–Thatcher (KT) method; they are constrained to the age 90 years and over totals in the official published MYEs.
- As the KT method relies on death registrations from previous years, it underestimates in conditions of decreasing mortality. Therefore, estimates of the very old are constrained to the published MYEs for those aged 90 years and over; this means that while the MYEs determine the total population size, the KT method determines the distribution of the population at different ages.

4 . Quality summary

Overview

Data sources for [estimates of the very old, including centenarians release](#) are death registration data and official mid-year population estimates (MYEs) of those aged 90 years and over.

The estimates were designated as [National Statistics](#) in 2011.

For a particular mid-year (30 June), estimates are published approximately 15 months after the reference date.

The [estimates of the very old, including centenarians release](#) provide annual mid-year estimates by sex and single year of age for people aged 90 to 104 years and for the 105 years-and-over age group.

Interest in population estimates at the oldest ages by single year of age has increased as life expectancy has increased and the number of centenarians grows. In recognition of this, we began publishing these estimates in 2007.

Estimates are based on [death registrations](#) statistics and are constrained to be consistent with the 90 years and over totals in the official annual MYEs. A guide to [Calculating population estimates of the very old](#) is available.

Information on the suitability of the death registration data used in producing these statistics is provided in the [Mortality statistics in England and Wales Quality and Methodology Information \(QMI\) report](#). Information on the quality of the MYEs is provided in the [mid-year estimates QMI report](#).

Statistics are produced for the UK, England and Wales separately. Comparable estimates for Scotland and Northern Ireland are produced and published by [National Records of Scotland \(NRS\)](#) and the [Northern Ireland Statistics and Research Agency \(NISRA\)](#) respectively. These feed into our UK estimates.

These data are produced primarily for our internal customers who require them for the production of [national life tables](#) and [national population projections](#). However, interest in these statistics has increased, and external users of the data include demographers, actuaries, medical researchers and others interested in longevity, population numbers and projected mortality rates.

Uses and users

The estimates of the very old are produced primarily for our internal customers who require them for the production of [national life tables](#) and [national population projections](#).

There is also a growing external user demand for the publication of population estimates at the oldest ages by single year of age as life expectancy increases and the very old population grows.

In addition to national and local government users, the estimates of the very old are also used by demographers, actuaries, medical researchers and others interested in longevity, population numbers, and/or past and projected age-specific mortality rates at the oldest ages.

We have not conducted a formal user consultation, but we have regular contact with our users. User need is evident from requests for these data, for example, from the Department of Health and Social Care, for use in outputs such as dementia prevalence rates at national level for health and social care planning. The estimates also feed into work by the Department for Work and Pensions and HM Treasury in formulating or assessing future policy on pensions.

We identified a user need to have Kannisto–Thatcher (KT) estimates for England and Wales separately and implemented this for the 2016 release. Previously, KT estimates were published for England and Wales combined.

Strengths and limitations

The main strengths of these statistics include:

- they are calculated using internationally recognised methods
- the calculation uses high-quality administrative data
- they use existing data sources
- they are produced annually, providing timely statistics to users
- they are consistent with the official mid-year estimates of the population aged 90 years and over
- equivalent comparable estimates are produced for the UK, England and Wales (separately) by the ONS and for Scotland and Northern Ireland by NRS and NISRA respectively

The main limitations of these statistics include:

- they are estimates and therefore some error is inevitable
- they are published rounded to the nearest 10 and grouped at age 105 years and over
- the method would not produce reliable results at local area geographies because of internal migration and the small numbers involved for the population aged 90 years and over
- there is a 15-month lag between the reference period and publication date

Recent improvements

We contribute anonymised information relating to people who have died aged 110 years and over (supercentenarians) in England and Wales to the [International Database on Longevity](#). This work involves validating the reported age at death, with the person's birth certificate from the General Register Office. In recent years we have seen increasing numbers of potential supercentenarian deaths that cannot be validated or where evidence suggests the deceased was younger than recorded on the death certificate.

The Kannisto–Thatcher (KT) method is based on the likelihood of survival from one age to the next, seen in deaths at each age. As the number of people in a cohort surviving to each successive age decreases, the calculations become more sensitive to fluctuations in the number of deaths occurring at the very oldest ages. In 2020, we incorporated the validation of deaths at aged 110 years and over into the production of the estimates.

Where there is evidence that the age at death is incorrect and the correct age is available, records are reassigned in the input data. Where there is evidence that the individual did not die at aged 110 or over but we do not know their correct age at death, the records are removed. Additionally, we exclude unvalidated deaths above aged 113 for males and 115 for females (the oldest ages at which deaths have been validated in the UK).

This change has improved the quality of the input data and has reduced the likelihood of overestimating the proportion of the aged 90 years and over population surviving to the very oldest ages. As the estimates are constrained to the 90 years and over total in the mid-year population estimate (MYE), there is no effect on the overall numbers.

5 . Quality characteristics of the data

Relevance

The [estimates of the very old, including centenarians release](#) provides estimates by sex and single year of age for persons aged 90 to 104 years and for the 105 years and over age group.

We produce these population estimates for England, for Wales and for the UK. Estimates for Scotland and Northern Ireland are produced and published by [National Records of Scotland \(NRS\)](#) and the [Northern Ireland Statistics and Research Agency \(NISRA\)](#) respectively. These feed into our UK estimates.

The primary use of these statistics is in the production of [national life tables](#) and [national population projections](#). Prior to 2007, these estimates were made available for research purposes but were not officially published. We began to publish these estimates as [Experimental Statistics](#) in 2007, in recognition of increased interest in population estimates at the oldest ages. At that time, they were named Estimates of the Very Elderly. Since 2010, we have also published these estimates at UK level.

In 2011, Estimates of the Very Elderly were assessed by the UK Statistics Authority and given [National Statistics](#) status. They were published as National Statistics for the first time in September 2011. In 2013, the estimates were renamed as [Estimates of the Very Old \(including Centenarians\)](#).

In 2015, the [Estimates of the very old, including centenarians](#) were reassessed by the UK Statistics Authority. The UK Statistics Authority required the Office for National Statistics (ONS) to assess the feasibility of producing estimates of the very old for England and Wales separately, rather than combined, as previously published. Separate estimates for England and Wales have been published since 2016.

The estimates are produced annually and published in [datasets](#) on the ONS website together with a [statistical bulletin](#) providing descriptions of and commentary on the estimates.

In addition to their primary use in the production of life tables and population projections, estimates of the very old are used for resource allocation and planning for older people. They are of policy interest because of implications for pensions, health and social care. They are used by:

- the Department for Work and Pensions, the Department of Health and Social Care, and HM Treasury for formulating or assessing future policy on pensions and health care
- demographers, health, and medical professionals and other researchers interested in longevity
- life insurance companies and the actuarial profession for the calculation of mortality rates at the oldest ages
- answering Parliamentary Questions and responding to media and public interest

Estimates of the very old are constructed using the Kannisto–Thatcher (KT) model of population at advanced age. They are based on [death registration statistics](#) and are constrained to be consistent with the 90 years and over totals in the official annual mid-year population estimates (MYEs). A guide to [Calculating population estimates of the very old](#) is available.

Accuracy and reliability

Estimates of the very old are constrained to the ONS MYEs of those aged 90 years and over by sex. MYEs are produced using a well-established demographic approach called the [cohort component method](#). This involves combining information from a number of data sources including the previous census, survey data and administrative registers. The data sources used are the best that are available on a nationally consistent basis; however, the estimates are subject to the coverage and error associated with these sources.

Any error in the 90 years and over census estimate is carried forward to the inter-censal MYEs and will be reflected in the estimates of the very old. In addition to non-response, other possible sources of error in the census estimate for people aged 90 years and over include inaccuracies in reporting of dates of birth (for example, proxy reporting by carers). Further information on the quality of the MYEs, including a brief explanation of the cohort component method, can be found in the [MYEs Quality and Methodology Information \(QMI\) report](#).

Survivor ratio methods such as the KT method provide age-specific estimates of the population for those aged 90 years and over using data from death registrations. The main assumption in these methods is that all deaths are recorded and that the recording of information on age at date of death is sufficiently accurate and reliable.

Statistics on death registrations are collected through administrative sources maintained by the General Register Office (GRO). These data are considered very reliable for two reasons. First, there is a legal requirement to register a death and the certificate issued at registration is needed and used by the recipient. Second, administrative data are not subject to sampling error in the way that survey data are.

We attempt to validate the age at death of all records where the date of birth recorded on the death registration makes the deceased person aged 110 years or older. A small number of records have been identified where the date of birth on the death registration has been recorded incorrectly; this can be expected at the very oldest ages because of proxy reporting during death registrations, which makes the deceased appear older than they were. The attained age at death is corrected by locating the individual's birth certificate where possible or excluded (see [Section 6: Methods used to produce the data](#) for more information) before calculating the estimates, to prevent skewing the distribution to the very oldest ages.

In the KT method, it is also assumed that international migration at the oldest ages is minimal so this component of population change can be ignored.

The survivorship ratio used is weighted over five years; this takes into account variations in the cohort size at each specific age and any cohort-specific fluctuations in mortality. Age-specific survivorship ratios are calculated using age-specific deaths data in both the denominator and, in a more complex way, the numerator. Registered deaths are a component of population change and are included in the MYEs. When creating mortality rates for the population aged 90 years and over, users should be aware that deaths data have been used to generate the estimates of the very old (the denominator in the calculation).

The estimates produced by the method are constrained to the official published mid-year population estimate of those aged 90 years and over. In effect this means that while the KT method determines the estimated distribution of the population aged 90 years and over, the accuracy of the overall KT estimates is dependent on the accuracy of the 90 years and over total in the official mid-year estimates. Estimates of the very old are published rounded to the nearest 10 people.

Output quality

The KT method overestimates the population if their mortality is rising over time and underestimates the population if their mortality is falling. This is because the estimation process takes into account the mortality in previous years, which may not be the same as in the most recent year. For this reason, estimates produced by the method for the current year and the back years are constrained to sum of the 90 years and over total in the MYEs.

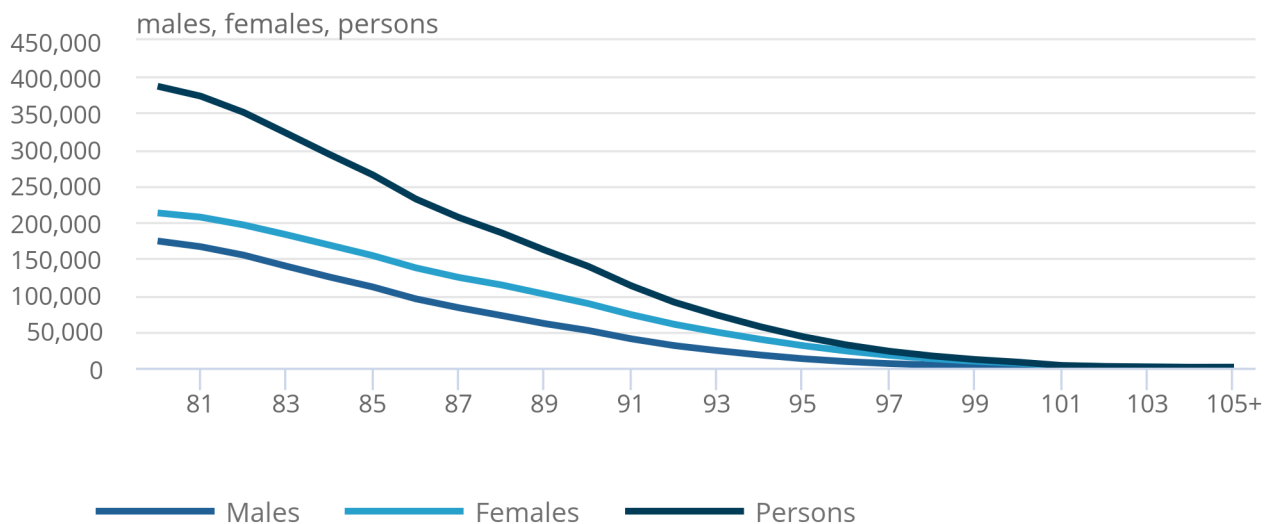
Constraining the KT estimates to sum to the 90 years and over totals in the MYE also provides users with a consistent set of single-year estimates beyond age 89 years and produces a smooth join between the population estimates produced by the two methods, cohort component and KT, at the age 89 and 90 years boundary.

Figure 1: There is a smooth join at aged 89 to 90 years between population estimates produced using the cohort component and Kannisto–Thatcher (KT) methods

UK population estimates aged 80 to 105 years and over, 2020

Figure 1: There is a smooth join at aged 89 to 90 years between population estimates produced using the cohort component and Kannisto–Thatcher (KT) methods

UK population estimates aged 80 to 105 years and over, 2020



Source: Office for National Statistics, National Records of Scotland, Northern Ireland Statistics and Research Agency

Notes:

1. Population estimates for those aged 80 to 89 years have been rounded to the nearest 10.
2. Pre-rebased mid-2020 population estimates for the UK are shown, despite more recent figures now being available. This is because the 2020 estimates of the very old were constrained to the aged 90 years and over total from the 2020 mid-year population estimates, which have now been superseded following the rebasing of Census 2021.

There is a trade-off between timeliness and accuracy in the production of estimates of the very old.

Previously, estimates of the very old were produced using calendar year death registrations for the reporting year and calendar year death occurrences for all preceding years. This is because there is a time lag between the occurrence of a death and registration of that death (particularly for deaths referred to the coroner for investigation), with deaths data by calendar year of occurrence not being available until several months after the availability of death registration data for that year. The number of deaths registered was replaced by the number of deaths occurring in that year when the following year's estimates were calculated.

From the [2002 to 2018 estimates of the very old release](#) onwards, deaths that are used in the calculations have been taken directly from the vital statistics life events system and are based on age at the beginning of the mid-year reference period, with age at this time point being calculated directly from the date of birth data on the death record. This change has enabled death occurrences data to be used for the reference year as well as the back years.

Coherence and comparability

We use an internationally recognised methodology in the construction of the estimates of the very old.

We calculate the estimates of the very old for England and Wales and, until 2010, also did so for Northern Ireland. Scotland and Northern Ireland use similar methodology to produce their age 90 years and over estimates by single year of age. The estimates are therefore comparable across UK countries, and this allows the estimates to be aggregated to produce estimates for the UK. A report has been published on the [comparability of estimates of the very old](#) produced by the ONS, NRS and NISRA.

Comparable time series are published back to the year 2002. Each annual set of estimates of the very old is derived using the same methodological approach. A feature of the methodology used is that previous years' estimates may change when a new year of data is added. Estimates are constrained to the published MYEs for the 90 years and over age group for the reporting year and re-constrained for previous years as the series is updated. The estimates of the very old are therefore consistent with the MYEs for England, Wales and the UK.

The estimates are always published as "provisional" because updates are made to the back series of data every time a new year's figures are added.

MYEs for the UK and constituent countries are produced for those aged up to 89 years by single year of age and for age 90 years and over using the cohort component method (see the [MYEs Quality and Methodology Information \(QMI\) report](#)). The census provides the mid-year resident population base for the cohort component method. Mainly owing to levels of uncertainty in the reporting of age in the census at older ages, this method is not used to produce single year of age data for those aged 90 years and over. Estimates of the very old are constrained to the 90 years and over totals in the MYEs and are therefore consistent with them. However, owing to the differences in the methods used to compile these sets of estimates, there may be some discontinuities between the oldest age in the MYEs (89 years) and the next age (90 years), where the figure is derived by the KT method.

Following the 2011 Census, the MYEs were revised back to mid-2002. Estimates of the very old were also revised accordingly.

Concepts and definitions (including list of changes to definitions)

These statistics provide estimates of those aged 90 years and over by sex and single year of age up to age 104 years and for the 105 years and over age group for England, Wales and the UK.

The National Statistics country classification's standard classification is used. The [National Statistics country classification](#) is based on ISO 3166-1, adapted to meet the data needs of users and producers of UK National Statistics. For the purposes of the National Statistics country classification, a country is the name, either short or official, of a current country, dependency or other geographic area of interest. This includes administrative subdivisions, particularly the nations of the UK: England, Wales, Scotland and Northern Ireland.

NRS and NISRA publish equivalent estimates for Scotland and Northern Ireland respectively, using the same definitions, classifications and methodology. These feed into the UK estimates.

There is no legislation governing the output. There is no deviation from agreed standards.

Geography (including list of changes to boundaries)

Estimates of the very old are produced by the ONS for England and Wales separately, and for the UK.

Accessibility and clarity

Our recommended format for accessible content is a combination of HTML web pages for narrative, charts and graphs, with data being provided in usable formats such as CSV and Excel. Our website also offers users the option to download the narrative in PDF format. In some instances, other software may be used or may be available on request. Available formats for content published on our website but not produced by us, or referenced on our website but stored elsewhere, may vary.

For information regarding conditions of access to data, please refer to the following links:

- [terms and conditions \(for data on the website\)](#)
- [copyright and reuse of published data](#)
- [accessibility](#)

In addition to this QMI report, basic quality information relevant to each release is available in the quality and methodology section of the relevant statistical bulletin: [Estimates of the very old, including centenarians, UK, 2002 to 2018](#).

Any enquiries regarding the estimates of the very old can be sent to pop.info@ons.gov.uk.

Timeliness and punctuality

Estimates of the very old, including centenarians, are usually published annually at the end of September. For a particular mid-year (30 June), they become available about 15 months after the reference date. This time lag reflects the availability of the data sources and the time required to process the data and calculate the estimates.

The publication of the estimates of the very old would be later than the planned date only if the input data used to calculate the estimates were not available, for example, if deaths data were unavailable or if substantial problems were encountered with the processing systems used to calculate the estimates. In previous years, the pre-announced publication date has always been met. In the year following the release of census estimates, the estimates of the very old are published later than the usual September release date.

For more details on related releases, the [GOV.UK Research and statistics release calendar](#) provides 12 months' advance notice of release dates. In the unlikely event of a change to the pre-announced release schedule, public attention will be drawn to the change and the reasons for the change will be explained fully at the same time, as set out in the [Code of Practice for Statistics](#).

6 . Methods used to produce the data

How we collect the data, main data sources and accuracy of data sources

Estimates of the very old are calculated from death registration data. They are constrained to age 90 years and over totals in the official published mid-year population estimates (MYEs).

Statistics on death registrations are collected through administrative sources maintained by the General Register Office (GRO). These data are considered very reliable for two reasons. First, there is a legal requirement to register a death and the certificate issued at registration is needed and used by the recipient. Second, administrative data are not subject to sampling error in the way that survey data are.

MYEs for the UK and constituent countries are produced for those aged up to 89 years by single year of age and for age 90 years and over using the cohort component method (see [population estimates methodology](#)).

Recently we have seen increasing numbers of deaths at aged 110 years and over in the input data that cannot be validated or where evidence has been found to suggest the deceased was younger than recorded. Where there is evidence that age at death is incorrect and the correct age is available, records are reassigned in the input data. Where there is evidence that the individual did not die aged 110 or over but we do not know their correct age at death, the records are removed. Additionally, we exclude unvalidated deaths occurring above aged 113 for males and 115 for females (the oldest ages at which deaths have been validated in the UK). This change has improved the quality of the input. Given that the estimates are constrained to the aged 90 years and over total in the mid-year population estimate (MYE), there is no effect impact on the overall numbers.

How we process the data

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, historical 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 that 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, we can calculate the average survivor ratio over the preceding m cohorts, 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 (or increased) mortality at higher ages may change 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 deaths data becomes available to inform the age profiles.

Methodology

Our methodology follows the KT method described earlier, 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 90 years and over population estimate for that year.

To carry out the calculations, deaths data need 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 England and for Wales were obtained by calendar year by age at death. 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.

From 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 Office for National Statistics (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 on the basis of date of birth and date of death.

Estimates of the very old, including centenarians, for the UK are produced by aggregating the age 90 years and over single-year estimates by sex for England, Wales, Scotland and Northern Ireland.

A guide to [Calculating population estimates of the very old](#) is available.

Further detail on the demography of centenarians and the KT method can be found in [Thatcher, R. \(1999\) \(PDF, 144KB\)](#) and [Thatcher, R., Kannisto, V., and Andreev, K.F. \(2002\)](#).

A summary of the cohort component method used to produce MYEs for ages up to 90 years can be found on the [population estimates methodology](#) page.

How we quality assure the data

All calculations are replicated using different software packages to check the accuracy of the figures.

Visual quality assurance is also carried out; this includes checking that:

- age and cohort distributions are plausible
- there are no large year-on-year changes
- totals sum to the annual MYEs

How we disseminate the data

[The estimates of the very old](#) are published annually. Links from the [release calendar](#) make the release date and location of each new set of estimates clear. A statistical bulletin, which describes the main patterns and trends in the data, accompanies the estimates. The estimates and the underlying data for the charts and tables in the bulletin can be downloaded free of charge in Excel format.

7 . Related links

Useful links

[Estimates of the Population aged 85 years and over, Northern Ireland](#)

[Dini, E. and Goldring, S. \(2008\) Estimating the changing population of the 'oldest old', Population Trends 132, pages 8 to 16 \(PDF, 157KB\)](#)

8 . Cite this methodology

Office for National Statistics (ONS), released 17 May 2024, ONS website, quality and methodology information report, [Estimates of the very old, including centenarians, QMI](#)