Office for National Statistics

Article

Dynamic population model, improvements to data sources and methodology: local authorities in England and Wales, mid-2011 to mid-2023

Update on the data used by the dynamic population model (DPM) to produce admin-based population estimates (ABPEs).

Contact:
Admin-based Population Estimates team
pop.info@ons.gov.uk
+44 1329 444661

Release date:
15 July 2024

Next release:
To be announced

Table of contents

1 . Overview of main changes

We are continuing to develop our research into the dynamic population model (DPM) which produces our mid-year admin-based population estimates (ABPEs) for England and Wales.

We are working towards our ABPEs meeting the standards expected of accredited official statistics by summer 2025.

This article discusses updates and improvements to data sources since provisional mid-2023 ABPEs were published in December 2023. Changes to methods are only planned for implementation when provisional ABPEs are produced around six months after the reference period. We plan to publish a methods guide in summer 2025 which provides a detailed explanation of how our ABPEs are produced.

Updated ABPEs for mid-2023 provide improved estimates which update the timely, provisional mid-2023 ABPEs published in December 2023. They include additional data that have become available. Provisional estimates make use of incomplete data alongside some assumptions about migration. More detail is provided in in our Provisional long-term international migration estimates: technical user guide.

We have also published a fully comparable time series of updated ABPEs for mid-2011 to mid-2022 for all 318 local authorities in England and Wales. These incorporate methodological improvements, detailed in our June 2023 and December 2023 publications available in Section 3: Improvements to data sources, and new and updated data sources including new estimates for migration. This latest series supersedes the time series released in February 2023.

In this methodology, all references to years refer to mid-year (30 June).

These are official statistics in development while we continue to refine our methods. They do not replace official mid-year population and international migration estimates and should not be used for decision making. These outputs should not be used without this warning.

2 . How the dynamic population model estimates the population

Our population and migration statistics are evolving in response to rapid changes in society and technology and to better meet user needs. To do this, we are making greater use of a wide range of administrative data that is acquired in line with our Data acquisition policy and our Data ethics policy. More information on these developments can be found in our Future of population and migration statistics in England and Wales web page.

The dynamic population model (DPM) produces our mid-year admin-based population estimates (ABPEs) and coherent components of population change for all local authorities in England and Wales by single year of age and sex. The DPM supports the use of an increasing range of data sources, allowing us to maximise and optimise the use of administrative data. Our methods and the data sources used to produce our ABPEs are being continually developed and enhanced.

The DPM builds on the cohort component method (as explained in our Population estimates methods guide), which has been used to produce our Population estimates for many years. It extends our existing mid-year population estimate methods using a Bayesian demographic accounts framework to create coherent demographic estimates from multiple imperfect datasets. Statistical modelling takes account of underlying demographic trends and allows for differing levels of coverage and uncertainty associated with input data.

Further information on how our ABPEs are created, their quality assurance, appropriate usage, and strengths and limitations, is available in our Mid-year admin-based population estimates in England and Wales Quality and Methodology Information (QMI) report.

An appendix to a paper, published by the UK Statistics Authority (The Authority) in February 2023, provides technical information on the Design and implementation of the DPM (PDF, 1343KB) including:

- the Bayesian demographic accounting framework used in the DPM
- how system models summarise underlying demographic trends
- how data models summarise the quality of the data sources
- how we estimate the hyper-parameters for the system models for births, deaths, in-migration, and out-migration
- the data models used for population stocks

We plan to publish a methods guide in summer 2025 which provides a detailed explanation of how our ABPEs are produced.

Articles outlining improvements to data sources and methodology used to produce our ABPEs are available on our website. There is also a June 2023 article and a February 2023 article available.

3 . Improvements to data sources

We have updated several data sources used in the dynamic population model (DPM). Unless discussed in this article, assumptions made about individual data sources remain the same as in our previous publications:

- published December 2023: Dynamic population model, improvements to data sources and methodology: local authorities in England and Wales, mid-2021 to mid-2023
- published June 2023: Dynamic population model, improvements to data sources and methodology for local authorities, England and Wales: 2021 to 2022
- published February 2023: Dynamic population model, improvements to data sources and methodology for local authorities, England and Wales: 2011 to 2022
- published November 2022: Dynamic population model for local authority case studies in England and Wales: 2011 to 2022 article

Population stock

The DPM requires unbiased population stock estimates for each year broken down by local authority, single year of age and sex.

The population stocks used for our July 2024 release are:

- for 2011 and 2021, census-based mid-year population estimates
- for 2012 to 2015, the NHS Patient Register (the Personal Demographics Service (PDS) is not available, and the Statistical Population Dataset (SPD) cannot be created for these years)
- for all other years, 2016 to 2020 and 2022 to 2023, the SPDs

Population stocks, including the SPDs, are coverage adjusted to allow for coverage errors.

The SPDs are linked administrative data, which apply a set of inclusion rules to approximate the usually resident population. SPD stocks are produced independently for each year, so any error in one year is less likely to be rolled forward to the next by the method. Our updated admin-based population estimates (ABPEs) use SPD version 5.0 supplemented with a second population stock for persons aged one to four years. SPD version 4.0 provided the second stock for 2016 to 2020 while PDS was used for 2022 and 2023. Our ABPEs published in December 2023 primarily used SPD version 4.1 as the population stock.

The data sources used in SPD versions 5.0 and 4.0 are:

- PDS
- English school census (ESC)
- Welsh school census (WSC)
- Individualised Learner Record (ILR)
- Higher Education Student Record (HESA): data for 2023 were missing a small number of institutions required for SPD version 5.0 but were present for the rest of the time series
- Hospital Episode Statistics (HES)
- Emergency Care Datasets (ECDS)
- birth registrations
- deaths registrations
- HM Revenue and Customs (HMRC) P14
- Lifelong Learning Wales Record (LLWR): version 5.0 only
- Patient Episode Database Wales (PEDW): version 5.0 only
- Emergency Department Data set (EDDS): version 5.0 only
- HMRC Frameworks: version 5.0 only
- Ministry of Justice (MoJ) prisoners data: version 5.0 only for 2021 onwards
- Customer Information System (CIS): version 4.0 only
- Benefits and Income datasets (BIDS): version 4.0 only

Our Data source overviews provide more information on these sources.

CIS and BIDS data are no longer provided by the Department for Work and Pensions (DWP), so are not included in SPD version 5.0; this primarily affects the population at retirement ages.

For this publication, SPD version 5.0 has been subject to data availability challenges, and we have had to make some changes to what was planned.

An undercoverage of 17-year-olds was identified in local authorities in Wales in 2023. As a result, the decision was made to remove this age group from the population stock for these areas in this year. In the absence of a population stock for certain age-groups and areas, the DPM uses just flow rates for births, deaths, and migration to derive estimates of population. This will lead to a slight reduction in the quality of the resultant ABPEs at this age, however, it follows our strategy to draw strength across data sources used in the DPM.

A partial set of HESA data were available for 2023, with student records missing for a small number of institutions. To account for this, we created a replica dataset to SPD version 5.0 for 2022 which reflected the missing institutions. Scaling factors were then calculated by comparing the SPD versions with full and partial HESA data in 2022. The scaling factors were then applied to SPD version 5.0 for 2023 to account for the partial set of HESA data. We plan to include full HESA data for 2023 in the next version of the SPD.

MoJ prisoners' data is only available from 2021 onwards and therefore could not be included for the years 2016 to 2020.

Child benefit data was not available at the time of production resulting in low coverage of persons aged one to four years, which varied over time. The DPM used a second population stock in the estimation process to account for this. We plan to include child benefit data in future SPD versions as we now have a regular supply of these data.

SPDs are quality assured at several stages throughout production and include checks on individual sources, through to processing and output checks. The most detailed checks involve analysts studying output data at local authority level by single year of age and sex.

Coverage adjustment of the population stocks, including SPDs, is required because of coverage errors. Coverage adjustment was not required for 2011 and 2021 since census-based mid-year population estimates represent the best available population stocks for these years. Coverage ratios for 2012 to 2020 were derived in the following way:

- comparing both the patient register and the SPD with mid-2011 population estimates to obtain smoothed coverage ratios for 2011
- comparing both the patient register and the SPD with mid-2021 population estimates to estimate coverage levels for 2021
- linearly interpolating between estimated coverage ratios in 2011 and 2021 to estimate coverage ratios for the intercensal years

For the years ending June 2022 and June 2023, we use one set of smoothed coverage ratios calculated by comparing SPD version 5.0 with mid-2021 population estimates. The same ratios were then used for 2022 and 2023.

We are currently exploring methods using administrative data sources for coverage adjustment. Work to date has focused on applying Dual System Estimation (DSE) to available administrative sources as a possible approach. DSE is a well-recognised and established approach, typically used to ensure that estimates resulting from a census have maximum coverage. It uses a coverage survey after the census to estimate how many people responded. We are currently considering if a similar method could be applied with different sources of administrative data. Further work could cover approaches such as multiple system estimation, use of additional administrative sources and potentially the use of surveys. In the meantime, we will continue to use Census 2021 results to apply coverage adjustment to the ABPEs and will continue consulting with methodological experts as we develop our methods.

Deaths

Our admin-based population estimates (ABPEs) are produced using deaths registration data, by age, sex, year of birth and local authority of usual residence. The registration of deaths is a service carried out by the Local Registration Service in partnership with the General Register Office (GRO) in England and Wales and is a legal requirement.

The regular supply of deaths data which feeds into the official mid-year population estimates does not include year of birth which is required to produce the ABPEs. To account for late registrations and adjust for usual residents who die outside of England and Wales and the small number of deaths where the local authority is unknown, we scaled deaths data used for the ABPEs to align with deaths data used for the official mid-year population estimates. Scaling was performed at age, sex, and local authority level. Methods for adjusting deaths data in the future are in development.

More information on deaths data is available in our Population estimates for England and Wales: methods guide.

Registration delays are explained in more detail in our Impact of registration delays on mortality statistics in England and Wales: 2021 article.

Internal migration and cross-border flows

Internal migration describes moves between local authorities in England and Wales. Cross-border flows describe moves in either direction between England and Wales (combined) and Scotland and Northern Ireland.

Usually, updated ABPEs for the latest year use our internal migration estimates. More information on these is provided in our Mid-year population estimates methods guide. However, for our updated ABPEs, the mid-2023 internal migration estimate had to be produced using an alternative approach after a change to the variables available in Higher Education Statistics Agency (HESA) data, which affected data linkage.

The flexibility of the DPM means that we can use alternative estimates of internal migration and cross-border flows. For mid-2023 ABPEs, the internal migration component was estimated using only Personal Demographics Service (PDS) updates of those changing their address on NHS systems.

The PDS data used to derive the estimates of internal migration has been updated and improved since our December 2023 publication. Some differences between the provisional and updated estimates are expected as a result.

To ensure consistency with the rest of the time series, we scaled the PDS-based migration for mid-2023 to account for internal migration not captured by PDS alone. For example, internal migration of Higher Education students between places of usual residence for study is not always well captured by the PDS. This can particularly affect estimates for local authorities with large student populations.

For provisional mid-2023 ABPEs published in December 2023, a mixed scaling approach was applied to PDS-based migration. This resulted in around half of all local authorities using scaling based on ratio of 2018 and 2019 averaged; the other local authorities were scaled based on 2022.

The availability of improved PDS data for 2022 has enabled a simpler internal migration scaling method for updated mid-2023 ABPEs. This method was applied consistently across all local authorities and used the most recent data available. Scaling involved calculating the ratio between PDS based internal migration estimates and mid-year population estimates of internal migration for 2022. This ratio was then applied to PDS data for 2023 for all local authorities.

Cross-border flows data used for the official mid-year population estimates were also used for our latest ABPEs. More information is available in our Mid-year population estimates methods guide for 2023.

International migration

Our latest ABPEs have been produced using provisional estimates of international migration for the year ending June 2023. The headline international migration estimates are the Long-term international migration published in November 2023 but the methods for disaggregating to local authority, sex and single year of age have been improved and rely more on observed activity in administrative data compared with the estimates used in our December 2023 ABPE release.

For mid-2022, we used updated international migration estimates which address an underestimation of students in a few local authorities.

We continue to use the United Nations' (UN) definition of a long-term migrant: a person who moves to a country other than that of their usual residence for at least a year. A summary of the main concepts behind our international migration estimates can be found in our Understanding international migration statistics methodology.

Combined migration

Our ABPEs are produced using combined migration, summing the internal migration, cross-border flows and international migration by age, sex, year of birth and local authority of usual residence.

In our December 2023, release we assumed an equal split of moves across age for each year of birth. However, for student ages in particular this does not hold true. For example, undergraduates typically move to begin their studies in the autumn aged 18 years. However, at the mid-year reference point most of these students are aged 19 years.

For this release, the relationship between age at time of migration and year of birth observed for internal migration has been used to apportion moves; we have assumed that this relationship also applies to international migration and cross-border flows.

4 . Updated uncertainty measures for combined emigration

The latest set of admin-based population estimates (ABPEs), published for mid-2011 to mid-2023, use updated values for internal migration uncertainty measures for males aged 18 to 23 years in a third of local authorities. These values form part of the input data used by the dynamic population model (DPM). Methods to measure the uncertainty around the estimates remain unchanged.

In our December 2023 release, the standard error of internal emigration estimates was not processed for males aged 18 to 23 years in a third of local authorities. As a result, uncertainty measure values for combined emigration were underestimated. The impact of this is small for all estimates (means) but has sometimes increased the width of the credible interval for combined emigration in the latest estimates.

5 . Future developments

The dynamic population model (DPM) and resulting admin-based population estimates (ABPEs) are showing great potential for producing timely, coherent population statistics.

We have requested an assessment of our ABPEs and are working to ensure that these meet the standards expected of accredited official statistics by summer 2025.

The DPM models migration as combined in-flows and combined out-flows for each local authority in England and Wales. To better understand changes in the size of local authority populations, these combined flows then need to be disaggregated to provide estimates of international, internal, and cross-border flows. Disaggregated migration flows data will be published in the future once we have had more time to develop and quality assure our methods and the resultant estimates.

We plan to publish a methods guide in summer 2025 which provides a detailed explanation of how our ABPEs are produced.

Feedback

User feedback plays an essential part in our ability to improve our statistics. Please email any questions or feedback to pop.info@ons.gov.uk.

Our Local population statistics insight feedback framework enables users of population statistics to provide feedback at local authority level and suggest data sources for us to better understand the quality of our estimates.

You can also sign up to email alerts from the Office for National Statistics Population team for updates on our progress, and to hear about upcoming events and opportunities to share your views.

6 . Glossary

Administrative data

Collections of data maintained for administrative reasons, for example, registrations, transactions, or record-keeping. They are used for operational purposes and their statistical use is secondary. These sources are typically managed by other government bodies.

Coverage errors

A coverage error occurs when a member of the population is not counted (undercoverage), is counted more than once (overcoverage) or is counted in the wrong location.

Credible intervals

The range in which the true value of the quantity being estimated is likely to be contained. We use 95% credible intervals in this article by taking 2.5th and 97.5th percentiles from the distributions of counts produced by our estimation process as the lower and upper bounds of our intervals, respectively. In this case, we can say that the probability that the true value lies in the credible interval is 95%.

Dynamic population model

A statistical modelling approach that uses a range of data to measure the population and population changes in a fully coherent way.

Official statistics in development

Official statistics that are in the testing phase and not yet fully developed. A more detailed explanation is available on the Office for Statistics Regulation website.

7 . Collaboration

The Office for National Statistics (ONS) has been supported in this research by the University of Southampton. Specifically, we would like to thank John Bryant, Peter Smith, Paul Smith, Jakub Bijak, Jason Hilton, Andrew Hind, Erengul Dodd and Joanne Ellison for their guidance and support.

8 . Related links

Admin-based population estimates: local authorities in England and Wales, mid-2011 to mid-2023
Dataset | Released 15 July 2024
Admin-based population estimates for all local authorities in England and Wales from the dynamic population model.

Population estimates for England and Wales: mid-2023
Bulletin | Released 15 July 2024
Accredited official mid-year population estimates for England and Wales by administrative area, age and sex.

Mid-year admin-based population estimates for England and Wales QMI
Methodology | Released 15 July 2024
Quality and Methodology Information for mid-year admin-based population estimates (ABPEs) for England and Wales, detailing the strengths and limitations of the data, methods used, and data uses and users.

Understanding mid-year admin-based population estimates for local authorities in England and Wales
Article | Released 15 July 2024
Important information about our mid-year admin-based population estimates (ABPEs) for England and Wales.

9 . Cite this methodology

Office for National Statistics (ONS), released 15 July 2024, ONS website, methodology, Dynamic population model, improvements to data sources and methodology: local authorities in England and Wales, mid-2011 to mid-2023