

Quality overview of data sources used in mid-2024 admin-based population estimates for England and Wales

This quality overview provides descriptions of the strengths and limitations of all the individual administrative (admin) data sources used directly in producing the mid-2024 admin-based population estimates (ABPEs) for local authorities (LAs) in England and Wales.

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

Release date:
30 July 2025

Next release:
To be announced

Table of contents

1. [Overview](#)
2. [Use of data sources](#)
3. [How we assess quality](#)
4. [Personal Demographics Service](#)
5. [English School Census](#)
6. [Welsh School Census](#)
7. [Higher Education Statistics Agency](#)
8. [Individualised Learner Record](#)
9. [Lifelong Learning Wales Record](#)
10. [Emergency Care Dataset \(England\)](#)
11. [Hospital Episode Statistics](#)
12. [Emergency Department Data Set \(Wales\)](#)
13. [Admitted Patient Care, Outpatient and Critical Care Datasets for Wales](#)
14. [HMRC Corporate Data Frameworks](#)
15. [HMRC P14](#)

16. [HMRC Child Benefit](#)
17. [Ministry of Justice Prisoners](#)
18. [Births and deaths registrations](#)
19. [Related links](#)
20. [Cite this methodology](#)

1 . Overview

This quality overview report provides descriptions of the strengths and limitations of all the individual administrative (admin) data sources used directly in producing the mid-2024 admin-based population estimates (ABPEs) for local authorities (LAs) in England and Wales.

The ABPEs are [official statistics in development](#), while we refine methods and data sources used to estimate the usual resident population. They do not currently replace the existing accredited official mid-year population estimates (MYEs) and international migration estimates, and should not be used for policy or decision making while their development continues. Detailed information on how the ABPEs are derived is provided in the [Methodology Information](#) accompanying the release, while this report focuses on the admin sources used as input data for the estimates.

[Section 2: Use of data sources](#) contains a very brief description of how the different data sources feed into the ABPEs.

[Section 3: How we assess quality](#) sets out our standard approach to assessing quality, which is applied to each data source covered in this report.

The main body of this report contains summary information on the quality of each admin data source contributing directly into the ABPEs. For this, we have adapted the [European Statistical System \(ESS\) Quality Dimensions](#) of:

- relevance
- accuracy and reliability
- timeliness and punctuality
- accessibility and clarity
- coherence and comparability

Information on the relevance of each data source, and thus why it was selected, is summarised in the introductory paragraphs for each data source. A section on coverage describes how well the data source covers the population of interest and any differences to standard definitions. Summary information on accuracy and completeness, and timeliness is provided in subsequent sections. Any lack of comparability with previous versions of the data is highlighted where relevant. We welcome [users' feedback](#) on this approach to setting out the strengths and limitations of each source, and whether there are any areas where more detailed information would be useful.

Other published reports contain similar information for data sources, which indirectly feed into the ABPEs. Links to these reports are provided in [Section 2: Use of data sources](#). All the admin data described here have been acquired in line with the Office for National Statistics's (ONS's) [data acquisition policy](#) and [data ethics policy](#) to assist with our research into transforming census and data collection practices. Data supply agreements with the supplier are in place for each source and the data are subject to robust controls to ensure that individuals cannot be identified. Further information, including ONS's privacy statement and data protection policy, can be found on our [Data protection](#) page.

2 . Use of data sources

Detailed descriptions of how the admin-based population estimates (ABPEs) are created, and on their quality assurance, appropriate usage, and strengths and limitations, are available in our [Mid-year admin-based population estimates for England and Wales quality and methods guide](#).

The use of admin data sources can be summarised as the following steps of the method.

1. A wide range of admin data sources are combined to produce a dataset covering individuals who might appear in the usual resident population. In this report we will call these sources "coverage sources". To ensure individuals are only counted once, we apply a consistent deduplication methodology across all data sources, using unique identifiers (such as NHS numbers or learner IDs) and retaining the most recent or most relevant record where multiple entries exist.
2. This dataset is filtered to remove records which are likely to relate to people who are not part of the usually resident population (for example, having left England and Wales) at the reference date of the estimates. We do this by looking for evidence of the person having some interaction with admin data. For this, we use the "coverage sources" described previously and a set of other admin data sources, which we will call in this report, "signs of activity" sources. We also use information from these latter sources to improve the accuracy of the information on records - for example, in providing more up-to-date information on local area of residence.
3. This filtered dataset (known as the Statistical Population Dataset) gives us initial estimates of the population at the current and previous mid-year points (30 June). These initial population estimates, called "stocks" undergo coverage adjustment to reflect expected differences between the combination of data sources and the true usually resident population.
4. Population changes over the year because of births and deaths are derived from civil registration data for those events.
5. Population changes over the year because of migration are estimated using the ONS's [Long-Term International Migration \(LTIM\) estimates](#) (in conjunction with data sources used to disaggregate those estimates to local authority level) and estimates of migration within the UK used in the mid-year estimates. Both these sets of estimates themselves rely on administrative data sources, which thus feed indirectly into the ABPEs.

Following these steps (described within Step 1: Create initial estimates in our [Mid-year admin-based population estimates for England and Wales quality and methods guide](#)), statistical methods are then used to reconcile the change in the population derived from Step 3, with the change estimated from Steps 4 and 5 to produce the best estimate of the population given the evidence from the different data sources.

List of data sources

The following data sources are used to produce the Statistical Population Dataset. As described in the previous section, we divide these sources into "coverage sources" and "signs of activity sources"

Coverage sources:

- Personal Demographic Service (Annual Stocks)
- Higher Education Statistics Agency (Student Record)
- English School Census
- Welsh School Census
- Individualised Learner Record
- Lifelong Learning Wales Record
- HM Revenue and Customs (HMRC) Frameworks
- Death Registrations

Signs of activity sources:

- Hospital Episode Statistics
- Admitted Patient Care, Outpatient and Critical Care Datasets for Wales
- HMRC Child Benefits
- HMRC P14
- Emergency Care Dataset
- Emergency Department Dataset (Wales)
- Ministry of Justice (MoJ) Prisoners (Record Level)

These sources, together with the births and deaths registrations data used in Step 4, are covered in this report, since they feed directly into the mid-2024 ABPEs. Admin sources indirectly feeding into the ABPEs through their use in Step 5 are described in more detail in separate reports available for the [mid-year estimates](#) and [Long-Term International Migration](#).

The ABPEs are subject to future development, including the use of new data sources. Future releases of the ABPEs will be accompanied by updated versions of this report, reflecting any changes in the data sources used.

3 . How we assess quality

Every data source used in the ABPEs is subject to continued evaluation using a Data Maturity Framework. This Framework considers how well the data source meets our requirements in producing estimates, the governance around the data source and the systems used in processing it, the quality and sustainability of the data, and the presence of contingency plans if the data are not provided to the expected time or quality.

In addition to this general assessment of quality when using data sources in the ABPEs, three further steps of evaluation and assurance are conducted for each supply of a selected data source.

1. We hold discussions with the supplier to identify any changes in definitions, collection procedures, quality assurance or similar aspects that might affect the data.
2. We conduct a wide range of standard checks on the structure and validity of the data received. These checks include: that data are of the expected format; the degree of missingness for each variable; numbers of invalid or unexpected values (including dates outside an expected range) for each variable; the quality of variables used in linking the data source to other sources; numbers of duplicate records; and evidence of any large clusters of records assigned to single addresses or postcodes.
3. We examine demographic and geographical patterns in the data and compare these with patterns in previous supplies and other sources.

These quality assurance (QA) checks are generally run using a QA [Reproducible Analytical Pipeline](#), which applies a suite of standard checks to each of the data sources described in this report. This system produces a QA Report for each source, which is interpreted by a researcher specialising in that and related sources. Any anomalies identified in the data are investigated and raised with the data suppliers where needed.

When reading this report, it is important to remember that the ABPEs are not solely dependent on any single administrative data source. For example, each coverage source is likely to have some elements of both undercoverage and overcoverage. Similarly, no single "signs of activity" source will reflect all usual residents at a particular time.

The ABPEs' methods are designed to draw on the strengths of the data sources when combined, while also accounting for differing levels of coverage and uncertainty associated with each of them. This means that "biases", in the sense of systematic undercoverage, overcoverage, or sources of inaccuracy identified for an individual source in this report, will not necessarily point to any corresponding bias in the final estimate.

4 . Personal Demographics Service

The NHS [Personal Demographics Service](#) (PDS) is the national electronic database that contains demographic data as well as a national unique identifier - the [NHS number](#) - for those who have interacted with an NHS service in England, Wales, the Isle of Man, or UK defence medical services, including interactions through GP practices and hospital visits.

We use two types of extract from the PDS in producing the admin-based population estimates (ABPEs).

The first type is the "annual stocks" extract, which is a snapshot of the PDS as of the 31 July of the year, and which is one of the "coverage sources" used as the basis of the initial population stocks. This extract is also used in the production of [Long-Term International Migration](#) estimates to distribute international migrants, both EU and non-EU, under the age of 16 years to local authority areas.

The second type is the "weekly update" files, which record changes of address. These files are used in estimating internal migration moves, which feed into the ABPEs, but are not used directly in the ABPEs.

Coverage

The PDS has a very high coverage of the usually resident population at all ages at the national level (England and Wales). Records are created for newborns or when a patient contacts an NHS service, primarily by registering with a General Practitioner (GP), but also through accessing Accident and Emergency (A&E) or attending hospital. Further PDS records are created from information collected and provided by the Home Office for visitors and migrants who have registered under the European Union Settlement Scheme, who have paid the Immigration Health Surcharge, or are part of a vulnerable cohort (such as asylum seekers and victims of trafficking).

People having private health care and irregular migrants are less likely to be included, as seen in our [Administrative data used in Census 2021, England and Wales methodology](#). It is believed that only a small minority of the total population is not known to the NHS and therefore not included in the PDS data. The data delivered contain approximately 85 million patient records of which over 60 million are for people currently registered with a GP.

We are also aware of a level of overcoverage in the PDS data. The data contain information on the reasons individuals can be removed from the database such as if they relocated within the UK, embarked abroad or died. We use this information to flag around 15 million records, unlikely to relate to a usual resident. However, not everyone is immediately removed from the system, for example, when usual residents emigrate without informing the NHS.

We are working with NHS England, and investigating further methods, to better understand and address overcoverage in the data.

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the ABPEs.

Interactions with PDS do not generally occur on an annual basis for most of the usually resident population. Records are only updated when an individual updates their details at a GP or through other interactions with health services via forms completed online or on paper under the direction of departmental staff. Different collection methods, particularly paper forms, can lead to an increase in entry error, therefore reducing the accuracy of the data.

However, NHS England has developed a new [GP registration service](#) that provides all GP practices with an integrated online option for patients. This allows online updates to contact details to improve the accuracy of the PDS and reduce third-party input errors.

Overall, the accuracy is usually high for age and sex, but address information might be affected by delays in people updating their records. Comparisons of the PDS with [Census 2021](#) found that the PDS tends to be more accurate for younger ages (schoolchildren and particularly those aged under 1 year old) and older ages (those aged 55 years and over), whereas student ages and young men aged in their 20s and 30s are less likely to interact with public services on a regular basis, as was reported in previous findings from [ONS's methodological work for producing census-based 2021 mid-year estimates \(MYE\)](#).

Similarly, findings from our [linkage between Census 2021 and the PDS](#) carried out in 2023, suggest that the accuracy of the PDS data on address varies across LAs between England and Wales. In areas which have higher population churn, for example urban areas and those with more higher education institutions, the PDS tends to be less accurate. This lower level of accuracy can lead PDS to show some overcoverage in more urban LAs where there is a higher rate of people moving in and out of the LA. For example, those visiting an LA may temporarily interact with an NHS service and then move out of the area but remain on the PDS (for example, short-term migrants).

Timeliness

The annual stocks file relates to the PDS data as they were on 31 July of that year. This reference date accounts for the typical one-month delay in GP registration following a move. This means that the ABPEs' reference date of 30 June will be less affected by a lag in updating records than if the same reference date was used for the PDS extract. The 2024 PDS stocks extract was delivered to the Office for National Statistics (ONS) in August 2024 and thus provides very timely information that aligns well with the ABPEs mid-year point.

Weekly update files are also received in a timely manner, supporting the ABPEs production schedule.

5 . English School Census

The English School Census (ESC), conducted by the [Department for Education](#) (DfE), collects demographic information on all pupils attending state-funded schools in England. This includes local authority-maintained schools (where funding goes to schools through the local authority) and those where funding goes directly to schools (for example, academies and free schools). The data cover nursery, primary, secondary, and special educational needs schools.

The ESC data are collected three times a year (January, May, October), to support school funding and to evaluate, or assess education policy. The Office for National Statistics (ONS) uses the January data as they fall mid-academic year and is the DfE's primary source for official statistics. This data source has good coverage for a subset of the population and is used as a coverage source in the admin-based population estimates (ABPEs).

Coverage

The ESC has high coverage of the usually resident population aged 5 to 15 years in England, when education is compulsory. In addition to the 5 to 15 years age group, the ESC also includes a proportion of children aged under 5 years, in early years education, and over 15 years, in state school sixth forms, or because they receive special educational needs provision. In the academic year ending 2024, the ESC covered 22,032 state-funded schools and 8,498,587 children, including all pupils enrolled in an English state school from the start of the nursery to the end of secondary. For more information see [Schools, pupils and their characteristics, academic year 2023/24](#).

There are several elements of undercoverage. Children not educated in a state school, for example, those attending private (independent) schools and home-schooled children, are not included. There are also a small number of pupils permanently excluded from school prior to the ESC Census Day who, depending on the circumstances, may not appear on the dataset. Further undercoverage may arise from recent migrants' children who delay school enrolment or from pupils joining between the January Census Day and the end-June reference date of the ABPEs. Children who are usually resident in England but attend a state school in Wales will not appear on the ESC but will appear in the Welsh School Census (see [Section 6: Welsh School Census](#)).

Overcoverage may occur where children who are not usually resident, such as short-term migrants to the UK, are enrolled in school and where pupils enrolled at the Census Day leave the country before the reference date of the ABPEs. Pupils attending Ministry of Defence schools overseas are filtered out from our supply because of non-residency in England and Wales. Pupils enrolled in state schools in England but living outside England will appear on the ESC, but these records are filtered out during the derivation of the ABPEs, which use the ESC data on home address, rather than school address, as the indicator of place of usual residence.

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the ABPEs.

ESC data collection is a statutory requirement from all schools under [Section 537A of the Education Act 1996](#) to finalise funding allocations. Therefore, there is a vested interest for schools to ensure data accuracy. The statistics on [Schools, pupils and their characteristics](#) produced by the DfE using ESC data are designated as [accredited official statistics](#), which means they are compliant with the [Code of Practice for Statistics](#).

Given that each school manages their collection of data independently, there may be some geographical variation in terms of accuracy and reliability. For example, the frequency that updates are requested from parents or guardians will vary and so information for some schools will be more up to date, and therefore more accurate than others.

Data are submitted to the DfE via the [COLLECT system](#) for England. The system runs automatic validation checks on the data. Schools are required to amend or provide suitable explanations for all errors. Guidance for completing this process within the validation period is provided, helping to further improve the accuracy of the data.

Timeliness

The data used in the 2024 ABPEs are from the annual spring collection supply for the academic year ending 2024, reflecting enrolments as of 18 January 2024. The data thus have a reference date around five months earlier than the reference date of the ABPEs. The ONS received these data on 1 July 2024, around five months after the reference date of the data.

6 . Welsh School Census

The Welsh School Census (WSC), also known as Pupil Level Annual School Census (PLASC), is an electronic collection of pupil-level demographic information from all local authority maintained, primary, secondary, middle, nursery and special schools (state funded) in Wales. The data are collected each January.

Completion of the WSC is a statutory requirement. The data support resource allocation including the Welsh Local Government Finance Settlement and the Pupil Development Grant, and inform education policy and research. This data source has good coverage for a subset of the population and is used as a coverage source in the admin-based population estimates (ABPEs).

Coverage

The WSC has high coverage of the usually resident population aged 5 to 15 years in Wales, when education is compulsory. In addition to the 5 to 15 years age group, the WSC also includes a proportion of children aged under 5 years, in early years education, and aged over 15 years, in state school sixth forms, or because they receive special educational needs provision. In the academic year ending 2024, the WSC covered 1,460 local authority maintained schools and 465,840 pupils (see [Schools' census results: January 2024](#)) including all pupils enrolled in a Welsh state school from the start of nursery to end of secondary.

As with the ESC, pupils not educated in a state school are excluded from the WSC. This includes those attending [independent \(private\) schools](#) and home-schooled children. Similarly, there are also a small number of pupils permanently excluded from school prior to the WSC Census Day who, depending on the circumstances, may not appear on the dataset. Further undercoverage may arise from recent migrants' children who delay school enrolment or from pupils joining between the January Census Day and the end-June reference date of the ABPEs. Children who are usually resident in Wales but attend a state school in England will not appear on the WSC but will appear in the English School Census (see [Section 5: English School Census](#)).

Overcoverage may occur where children who are not usually resident, such as short-term migrants to the UK, are enrolled in school and where pupils enrolled at the Census Day leave the country before the reference date of the ABPEs. Pupils enrolled in state schools in Wales but living outside Wales will appear on the WSC data, but these records are filtered out during the derivation of the ABPEs, which use the WSC data on home address, rather than school address, as the indicator of place of usual residence.

Accuracy and completeness

Like the ESC, the WSC is a statutory data collection and therefore there is a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the ABPEs.

Data are collected throughout the year via [the school's Management Information System \(MIS\)](#) and submitted on Census Day, with a one- to two-month window for corrections. This flags errors and inconsistencies that schools must resolve, before submission. Further checks are conducted by the Welsh Government's statistical service.

WSC data accuracy is judged to be high. Published outputs using the data, such as [Schools' census results: January 2024](#) are designated as [accredited official statistics](#).

Timeliness

The annual WSC extract for the academic year ending 2024 reflects enrolments as of 16 January 2024. The data thus have a reference date around five months earlier than the reference date of the ABPEs. The Office for National Statistics (ONS) received the data in August 2024, around seven months after their reference date.

7 . Higher Education Statistics Agency

The [Higher Education Statistics Agency \(HESA\)](#) is part of [Jisc](#), the UK's digital, data and technology agency for tertiary education, that collects and disseminates data on higher education (HE) across the UK. Jisc is the designated data body for England under the Higher Education and Research Act 2017. The HESA Student record data are submitted by [higher education providers](#) for all students attending a course that leads to the award of a higher education qualification or higher education-level credit. This includes those awarded qualifications by the provider or another awarding body, including international students.

As students in higher education often move to study, affecting local populations, we use the annual Student record extract (covering 1 August to 31 July) as one of the "coverage sources" used as the basis of the initial population stocks.

The data can also supplement Personal Demographics Service (PDS) data when producing internal migration estimates, since they provide better quality data on residential location of a large group of [young healthy adults, particularly males, that tend to be slow in updating their details](#) in other data sources (PDF, 256KB).

Our [Long-term international migration: quality assuring administrative data](#) report, provides more information and assurance about the quality of the HESA Student record data used to produce migration estimates of EU nationals.

Coverage

The [HESA Student record](#) includes data from HE providers with a statutory duty to report to funding and regulatory bodies across the UK. Although some privately funded institutions are not required to provide these data, the HESA data provide strong coverage of students in HE, including international students residing in England and Wales.

Those studying overseas or in the UK for fewer than eight consecutive weeks are not included in the HESA Student record. An exception applies to a small number of distance learners, such as Crown servants overseas or members of the British armed forces, whose tuition is funded by a UK government department, or who are eligible for a tuition fee loan. We filter out these students to align with the admin-based population estimates (ABPEs) definition of usual residence.

The HESA Student record includes anyone in active study at least two weeks after their start date (or anniversary of it) during the reporting period (1 August to 31 July). This may result in some overcoverage, as students who drop out after the start of the term remain in the data, even if they no longer reside at their term-time address.

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the ABPEs.

HESA collects data from a large number of HE providers, who are required to report to their funding and regulatory bodies. To support this, HESA provides guidance and resources to ensure consistency. All submissions undergo automated [validation checks](#), and any that fail are returned to providers for correction. Therefore, HESA Student record data are considered a high-quality source.

However, there are some limitations for ABPEs use. Student address information may become less accurate over time. Permanent home postcodes are collected at the start of study, while [term-time postcodes](#) are updated annually with each [course session](#). Practices vary across and even within HE providers.

Students can move to a new term-time address shortly after registering their details with their HE provider, but these updates may not be reflected in the HESA data. This means that the address recorded within the HESA Student record may not reflect the student's address on 30 June, the ABPEs reference point.

The impact of this is not thought to be significant since most in-term student moves are usually short-distance and remain within the same local authority. Since ABPEs are produced at local authority level, such moves are unlikely to affect the statistics. Assuming these moves are not systematically patterned, any random errors in these tend to cancel out. Additionally, when HESA Student record data are used for internal migration estimates, we supplement them with more recent address information from other data sources like patient registrations.

Timeliness

The extract for the academic year ending 2024, used in the 2024 ABPEs, was delivered to the Office for National Statistics (ONS) in February 2025. The bulk of data collection is done at the start of the academic year by the HE providers with a lag of around 18 months from this point to the data being delivered to the ONS. However, the HE providers update records throughout the year ahead of the final submission to HESA, including where students alert them of a change (for example, an address change). The period of collection covers 1 August 2023 to 31 July 2024, with the reference date for the ABPEs being 30 June 2024. We are working with Jisc to improve the timeliness of the data.

8 . Individualised Learner Record

The Individualised Learner Record (ILR), supplied by the Department for Education (DfE), contains information about individuals enrolled in further education or training through providers in England's Further Education (FE) and Skills sector. For more details, see the [ILR data submission guidance](#).

ILR data are used alongside the Welsh Government's Lifelong Learning Wales Record (LLWR) as coverage sources, capturing those who are in further education and may be missing from other admin data sources.

Coverage

The Further Education and Skills sector includes FE colleges, sixth form colleges, training organisations, local authorities, academies, and voluntary and community organisations. The ILR provides strong coverage of the population in further education in England.

Learners are removed from ILR if they withdraw before completing one episode of learning (a period of continuous enrolment at a single provider).

Accuracy and completeness

Completion of the ILR is mandatory for publicly funded providers. The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the admin-based population estimates (ABPEs).

The accuracy of the data submitted to the DfE is high because of the application of rigorous [validation rules](#). These are enforced through systems like the Funding Information System (FIS), which providers use to check data quality before submission. Providers do not receive funding for learners having invalid or incomplete entries. Data outside permitted parameters are flagged as a validation error and must be corrected.

Providers may amend and resubmit ILR data throughout the academic year, supporting continuous data quality improvement and enabling error correction before final submission to the Office for National Statistics (ONS).

Timeliness

Data are collected from learning providers throughout the academic year, which runs from 1 August to 31 July, and are submitted monthly to the DfE. The 2024 ILR extract was delivered to the ONS in September 2024. The ILR provides very timely information for the production of the mid-2024 ABPEs, with the reference period aligning well with the ABPEs' mid-year point.

9 . Lifelong Learning Wales Record

The [Lifelong Learning Wales Record](#) (LLWR) is a collection of data used by the Welsh Government on learners and the learning undertaken by them from learning providers funded by [Medr](#). The data include information on learners in post-16 years education and training, excluding those at schools but including those at further education institutions, other work-based learning providers and community learning provision within Wales. The data also provide the official source of statistics on post-16 years non-higher education (HE) learners in Wales.

The data collected via the LLWR underpin many aspects of Medr's work, including the planning, funding, monitoring and quality assurance of post-16 years provision.

As with the ILR data described in [Section 8: Individualised Learner Record](#), the LLWR data are used as a "coverage source" in the derivation of the admin-based population estimates (ABPEs).

Coverage

The LLWR provides high coverage of learners in publicly funded post-16 years further education in Wales (excluding higher education learners captured by HESA).

The LLWR excludes school-based provision, so sixth form learners typically aged 16 to 18 years, and studying for A-levels or equivalent, are not included. These are captured by the Welsh School Census.

As the data cover learners enrolled in Medr-funded programmes delivered within Wales, it includes a small number of non-usual residents of England and Wales.

Accuracy and completeness

The Lifelong Learning Wales Record (LLWR) is updated annually, and its completion is compulsory for providers seeking funding. As such, the data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the ABPEs.

The LLWR operates as a rolling data collection, allowing providers to submit and amend data continuously, throughout the academic year. Monthly "freezes" of the live data are taken, with the December freeze following the end of an academic year, typically used for statistical purposes.

As a business as usual (BAU) collection, every year we replace the previous LLWR extract with the latest version received from the Welsh Government, meaning that data accuracy is constantly improved.

Timeliness

The 2024 extract of data was supplied in December 2024 and relates to the academic year ending July 2024. Hence, the data are timely and align well with the ABPEs mid-year point.

10 . Emergency Care Dataset (England)

The [Emergency Care Dataset](#) (ECDS) is the national dataset for urgent and emergency care in England, replacing [Hospital Episode Statistics Accident and Emergency \(A&E\) from 2020](#). It captures patient attendances at NHS hospitals in England, including minor injury units, walk-in centres, and 24-hour and consultant-led emergency departments. Data are submitted by NHS hospitals and urgent care providers, with NHS England overseeing data collection, processing and dissemination.

Each record represents a single episode of care, and includes:

- clinical outcomes (for example, diagnoses)
- patient demographics (for example, age, sex, ethnicity)
- administrative information (for example, attendance dates)
- geographical details (for example, treatment location, patient's area)

An extract of the data (excluding medical details) is used by the Office for National Statistics (ONS) alongside Personal Demographics Service (PDS) and Hospital Episode Statistics (HES) data to infer presence in England's resident population through the population's interactions with health services. ECDS data are used as a "signs of activity" source, providing evidence that records on other sources relate to people who are usually resident in England and Wales, and improving the accuracy of information for those records.

Coverage

ECDS only includes NHS-funded emergency care from hospitals and independent providers across England. It excludes private providers and emergency care in Scotland, Wales and Northern Ireland. However, it may include non-residents in England, such as individuals from other UK countries, foreign visitors, or short-term migrants, if they receive care in England. We mitigate this overcoverage by excluding patients with a recorded home address outside England during ECDS pre-processing.

ECDS may also include records for individuals who died in emergency care or emigrated after treatment. The latter population groups might introduce some overcoverage, which is accounted for in the ABPEs' estimation process.

Undercoverage may occur where data are incomplete: for example, patients leaving before treatment.

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the ABPEs. NHS England publishes an interactive [ECDS Data Quality Dashboard](#), which shows the levels of data completeness across the data items.

Data are recorded by healthcare professionals during care episodes. While NHS England applies quality checks, inconsistencies may arise because of varying recording practices by NHS trusts or emergency conditions. We work closely with NHS England to understand any inconsistencies with the data.

Timeliness

NHS England supplies us with ECDS data monthly and annually. The reference period for the annual supply of ECDS data is the financial year (1 April until 31 March). The annual dataset is delivered to the ONS in October, with the monthly data available shortly after the reference date. The reference period for the data aligns well with the reference date of the ABPEs.

11 . Hospital Episode Statistics

The [Hospital Episode Statistics](#) (HES) dataset captures all outpatient appointments, and admissions to NHS hospitals in England. It covers all Sub-Integrated Care Boards and Integrated Care Boards (formerly Clinical Commissioning Groups) and is managed by NHS England. HES captures healthcare activity, not individuals, with data submitted by NHS hospitals and processed centrally.

Each record reflects a care or appointment episode, and includes:

- care dates
- patient demographics (for example, age, sex, ethnicity)
- administrative information (for example, provider)
- geographical details (for example, treatment site, patient's postcode)

Between April 2007 and March 2020, HES consisted of three components: Accident and Emergency (A&E), Outpatient (OP), and Admitted Patient Care (APC). From April 2020, only OP and APC remain active, as the A&E component was no longer updated and replaced by the Emergency Care Dataset (ECDS).

An extract of the data (excluding medical details) is used by the Office for National Statistics (ONS) as a "signs of activity" source in the admin-based population estimates (ABPEs), improving the recognition of which records relate to usual residents.

Coverage

HES includes NHS-funded inpatient and outpatient care in hospitals and independent providers across England. It excludes private care not funded by the NHS and care outside England. HES captures a portion of the usual resident population through hospital interactions, but does not cover all residents, particularly those who rarely use the NHS services.

However, HES may include non-usual residents, such as people from other UK nations or overseas, if treated in England. The ONS mitigates this overcoverage by excluding records with addresses outside England.

HES may also include individuals who died or emigrated after treatment. The latter population groups might introduce some overcoverage, which is accounted for in the ABPEs' estimation process.

Coverage will be lower for groups less likely to use hospital services (for example, young adults, migrants, or those using primary or private care).

Accuracy and completeness

HES data have a high level of completeness for the population covered, with low levels of missingness for the main variables used in the derivation of ABPEs. Healthcare professionals record this information during patient interactions. Each record includes a patient reference, allowing the ONS to link multiple episodes for the same individual.

Under [Section 45c of the Statistical and Registration Service Act 2007](#), all NHS hospitals are legally required to submit HES data. NHS England cleans, standardises and validates the data using [validation checks](#) to improve coherence and comparability. These checks focus on completeness, though verifying accuracy is more challenging.

HES data are collected monthly throughout the financial year, with each update refreshing records back to the start of the year. A final annual refresh follows year-end, improving overall accuracy.

Timeliness

NHS England supplies the ONS with HES [data monthly and annually](#). The reference period for the annual supply of HES data is the financial year (1 April until 31 March). The annual dataset is delivered to the ONS in October, with the monthly data available shortly after the reference date. The reference period for the data aligns well with the reference date of the ABPEs.

12 . Emergency Department Data Set (Wales)

The Emergency Department Data Set (EDDS) is the national dataset for emergency care in Wales, capturing attendances at NHS Wales hospitals, including emergency departments (EDs) and minor injury units (MIUs). Data are submitted by NHS hospitals and processed by [Digital Health and Care Wales](#) (DHCW), part of NHS Wales.

Each record represents a single episode of care, and includes:

- clinical information (for example, treatment, outcome)
- patient demographics (for example, age, sex)
- administrative details (for example, attendance dates)
- geographical information (for example, treatment location, patient's locality)

An extract of the data (excluding medical details) is used by the Office for National Statistics (ONS) as a "signs of activity" source for the admin-based population estimates (ABPEs) to infer presence in Wales's resident population through interactions with NHS services.

Coverage

The EDDS captures data on all individuals who attend EDs and MIUs within NHS Wales hospitals. This includes patients of all ages and demographics, regardless of their residency, ensuring a comprehensive overview of emergency care activity. Patients from England or other countries that receive treatment within the Welsh NHS system are included in the data. The latter are filtered out when estimating the population stocks.

Similar to the ECDS, the EDDS may also include records for those who died in an ED or emigrated after interacting with the NHS. The latter population groups might introduce some overcoverage, which is accounted for in the ABPEs' estimation process.

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the ABPEs. Data are collected and coded at each hospital, then transferred to NHS Wales. While there have been concerns about inconsistencies in EDDS because of variability in data entry across sites, the data we use are judged to be generally accurate for ABPEs' purposes.

Timeliness

The EDDS data used in the 2024 ABPEs relate to the period 1 July 2023 to 30 June 2024, which aligns well with the ABPE's reference date of 30 June. The ONS received the data in October 2024.

13 . Admitted Patient Care, Outpatient and Critical Care Datasets for Wales

The Admitted Patient Care (APC), Outpatient (OP) and Critical Care (CC) datasets are collected, processed and disseminated by [Digital Health and Care Wales](#) (DHCW), part of NHS Wales. Together they comprise all inpatient, day case and critical care activity undertaken in NHS Wales, plus data on Welsh residents treated in English trusts. The APC and Outpatient data are the Welsh counterparts to NHS England's Hospital Episode Statistics (HES) and are essential for healthcare planning, policy and research.

Records reflect a care or appointment episode, and include:

- clinical information (for example, diagnosis)
- patient demographics (for example, sex, age)
- administrative details (for example, discharge and admission dates)
- geographical information (for example, treatment location, patient's locality)

An extract of APC, OP and CC (excluding medical details) is used by the Office for National Statistics (ONS) as a "signs of activity" source for the admin-based population estimates (ABPEs) to infer presence in Wales's resident population through interactions with NHS services.

Coverage

The datasets include all inpatient and day-case hospital interactions recorded by NHS Wales that are not captured by the [Hospital Episode Statistics](#) (HES), alongside data on Welsh residents treated in English trusts (further information is provided on the [Digital Health and Care Wales website](#)).

Similar to the EDDS, deceased individuals are filtered out when creating the ABPEs. However temporary workers living in Wales, students from other parts of the UK or abroad who temporarily reside in Wales, visitors and tourists who require medical care while in Wales, and patients from other parts of the UK receiving treatment in Welsh hospitals might be included in the APC, OP and CC data. The latter population groups might introduce some overcoverage, which is accounted for in the ABPEs' estimation process.

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the ABPEs.

The data undergo rigorous validation and quality checks led by DHCW. While these processes promote consistency, some variation may remain from differences in hospital coding practices.

Incomplete data on a patient at the time of the initial collection can be enhanced by additional information collected subsequently and provided in regular updates.

Timeliness

The reference period for the APC, OP and CC data used in the 2024 ABPEs is 1 July 2023 to 30 June 2024, which aligns with the reference date of the ABPEs. The data were received by the ONS in October 2024, within four months of the data's reference period.

14 . HMRC Corporate Data Frameworks

The Corporate Data Frameworks (CDF), commonly referred to as "Frameworks", is a relational database maintained by HM Revenue and Customs (HMRC). It contains both current and historical demographic data for individuals interacting with one or more HMRC systems, covering the period 2007 until 2025.

Frameworks draws from multiple live and historical HMRC systems, including Tax Credits, Student Loans, Child Benefit, Child Trust Fund, and the Integrated Debt Management System. The data cover both current and historical Pay As You Earn (PAYE), and Self-Assessment taxpayers, as well as individuals receiving benefits from HMRC.

The data include record-level demographic and address information, updated as individuals engage with HMRC services. Updates may occur through direct contact, employer submissions, or interactions with systems such as the National Insurance and PAYE service (NPS), Computerised Environment for Self-Assessment (CESA), and others. Address data are also validated using the Post Office Address File and third-party services like GB Group.

Frameworks data provide strong coverage of the usually resident, working-age population in England and Wales, and this source is used by the Office for National Statistics (ONS) as a "coverage source" for the admin-based population estimates (ABPEs) stocks.

Records are linked via [National Insurance numbers](#) (NINOs), enabling integration with other income-related datasets. This linkage supports the identification of signs-of-life, such as payments or pension activity, helping to determine whether individuals are likely residents in England and Wales, and should be included in the ABPEs.

Coverage

Frameworks data provide strong coverage of the usually resident, working population in England and Wales who have a NINO.

However, the data exclude individuals who have not interacted with HMRC including children who are not in the child benefit system, UK-born adults who have not worked or claimed HMRC benefits, and those without the right to work or claim benefits in the UK.

We are aware of a level of overcoverage as the data may include individuals who have moved out of England and Wales, or even the UK, without notifying HMRC, temporary workers who are no longer active, and duplicate entries that have not been resolved. The overcoverage is accounted for in the ABPEs estimation process.

Coverage can be affected by changes to the tax and benefits system. For example, the removal of child benefit for high earners (as per the [High Income CB Charge](#)) resulted in a reduction in the coverage of women not interacting with other HMRC services, children and new births (also see [Section 16: HMRC Child Benefit](#)).

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the ABPEs.

Demographic and address information in HMRC systems is updated through individual interactions with services such as National Insurance, PAYE, Computerised Environment for Self-Assessment (CESA), the Personal Tax Account (PTA) and employer submissions. HMRC uses the Post Office Address File (PAF) to validate address entries, ensuring they match official UK address formats.

Updates to Frameworks records rely on individuals or employers notifying HMRC of changes. Hence, update frequency varies across the population. Individuals who interact regularly with HMRC, such as those in continuous employment, are more likely to have current records, while others, including students, retirees, or more mobile groups, may have less up-to-date information.

Overall, the accuracy of age and sex data is expected to be high because of their reliance on official documents at the point of registration. However, address information may be less reliable for certain sub-groups, particularly younger or more mobile individuals who may not consistently update their records. This can lead to variation in data quality across geographic areas, with urban locations and areas with high population turnover more likely to experience discrepancies.

Timeliness

The data are supplied to the ONS as monthly extracts following the end of that month. This means that there is a very short lag between the reference period of the data and supply to the ONS, and that the data provided align well with the ABPEs' reference date.

15 . HMRC P14

P14 refers to an annual extract of tax year information providing a summary of each employee's annual earnings, Income Tax, National Insurance contributions (NICs) and other employment-related details. The data are derived from HM Revenue and Customs' (HMRC's) Real Time Information System (RTI) returns and P14 tax form returns (which since the tax year ending 2014 have largely been replaced by RTI returns). The data are used as a "signs of activity" source for the admin-based population estimates (ABPEs) to infer presence in England and Wales.

Coverage

P14 includes individuals in the UK paid through a Pay As You Earn (PAYE) scheme, covering a large share of the working population. However, it excludes volunteers, unpaid workers, the self-employed (for example, sole traders using Self-Assessment), and those whose employers are not required to [register for a PAYE scheme](#), namely where none of their employees earn more than a certain threshold, receive expenses or company benefits, have another job, or receive a pension or certain state benefits.

There may be an element of overcoverage as individuals who have migrated out of England and Wales might still appear in the data if paid from the UK.

Accuracy and completeness

The data have high completeness for the population covered and low missingness for data fields used in the derivation of ABPEs, because of mandatory employer reporting requirements. However, the accuracy and currency of these fields depend on the information provided by employees and maintained by employers. For example, although employers can update employee details throughout the tax year, updates such as a change in home address rely on employees notifying their employer and HMRC does not routinely validate this information.

Annual extracts are delivered to the Office for National Statistics (ONS), around five months after the end of the tax year. These include new and amended P14 returns, which can help improve the accuracy and completeness of the dataset.

Timeliness

The ONS receives annual data covering the tax year, with the tax year ending 2024 data received in August 2024. The tax year covers the annual period to 5 April, prior to the 30 June ABPEs reference period, therefore, the data used for the current year of the ABPEs do not include information for the May and June periods, which are not available until the following year.

16 . HMRC Child Benefit

[Child Benefit](#) (CB) data are sourced from HM Revenue and Customs' (HMRC's) Child Benefit System, which records the number of claimants and their children. CB is paid to one person per child. This person is responsible for a child aged under 16 years, or a qualifying young person (aged under 20 years and in [approved education or training](#)). It is usually paid every four weeks, though weekly payments are possible.

Only one person can claim CB for a child in any given week without a limit on the number of children a person can claim for. The child must be, or treated as present in the UK, and the claimant both present and ordinarily resident in the UK. Temporary absences, such as holidays or medical treatment abroad, do not affect eligibility. CB data were previously provided by the Department for Work and Pensions (DWP) as part of the Benefits and Income Dataset. Current and future deliveries come directly from HMRC and include additional variables.

CB data provide evidence for beneficiaries being resident in the UK and this source is thus used as a "signs of activity source" in the production of the admin-based population estimates (ABPEs).

Coverage

CB data have good coverage of the usually resident population aged 0 to 17 years, and some coverage of the 18 to 19 years age-group. Coverage in the latter age-group is affected by the eligibility rules for the benefit which exclude, for example, university students. Further undercoverage will result from higher-income families who opt out of claiming because of the [High Income CB Charge](#), recent migrants or newborns not yet registered, and foster children if the local authority is paying their accommodation or maintenance.

Non-usual residents are generally ineligible for CB and excluded from the data. However, exceptions exist, such as Crown servants posted overseas, or residents of certain countries, who may still appear in the data. This can lead to overcoverage, which we mitigate by filtering out ineligible records.

An important difference between the previous DWP supply and the current HMRC supply, is that the latter may include individuals who register for CB but do not receive payments because of the high-income cap.

As the data cover the entire UK, the Office for National Statistics (ONS) is using address information to only retain England and Wales residents, in line with the ABPEs.

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the ABPEs.

The Child Benefit statistics are of high quality and as such they are classified as [accredited official statistics](#). They are as close to real-time as possible and represent the complete picture at 31 August, including back-dated awards with a start date on or before 31 August.

The accuracy of the CB data is high since to make a successful Child Benefit claim with HMRC, individuals are required to provide accurate and complete information, including but not limited to the full legal name of the claimant and the child(ren), current address, child's birth certificate and National Insurance number (NINo) of the claimant. These are mandatory fields and are cross-checked with other government databases to verify identity and prevent fraud. Only a small proportion of records might be expected to have some inaccuracies because of name changes, for example, because of marriage or legal changes that are not promptly reported, or cultural variations in name formats.

However, despite the requirement for individuals to update their address as soon as they move, the accuracy depends on the members of the public interacting with HMRC, for example, by calling or writing to HMRC, using the HMRC app or the [Personal Tax Account](#).

Validation checks conducted by HMRC increase the reliability of the data supplied and include sense checks with comparisons across years. The [Child Benefit Statistics quality report](#) is an example of the statistical processing and quality management of HMRC-supplied datasets.

Timeliness

The data are supplied to the ONS on a quarterly basis, usually within two months of the quarterly reference period (for example, data to June 2024 were available by August 2024). Therefore, the data are timely and their reference period aligns well with the ABPEs' reference period.

17 . Ministry of Justice Prisoners

These data are supplied to the Office for National Statistics (ONS) by [HM Prison and Probation Service](#) (HMPPS), an executive agency sponsored by the Ministry of Justice (MoJ). The MoJ processes person-level prisoner data for law enforcement purposes including processing for execution of criminal offences as provided under Part 3 (Law Enforcement Directive) of the Data Protection Act 2018 (DPA).

The data cover all prison establishments in England and Wales, which are required to record prisoner details on the [Prison National Offender Management Information System](#) (p-NOMIS). They include personal information, such as prisoners' name, date of birth, age and sex, that allow them to be linked to other admin data sources.

The data provide valuable information on the current residence of this population group and is used as a "signs of activity source" in the production of the admin-based population estimates (ABPEs).

Coverage

The data provide excellent coverage of individuals held in custody, including all prisoners aged 15 to 21 years in public and private young offender institutions (YOIs), and adult establishments in England and Wales, under the following legal status groupings:

1. Sentenced
2. Recall
3. Remand
4. Indeterminate Sentence
5. Convicted Unsentenced
6. Immigration Detainee
7. Unknown
8. Civil Prisoner
9. Other

The data do not cover individuals held at [immigration removal centres](#) as these are non-criminal detainees (unlike immigration detainees housed in prisons).

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the ABPEs.

Data on prisoners are updated in real time by the various institutions involved in their collection and kept up to date, meaning that the data are accurate and do not suffer from significant lag effects.

Confidence in data accuracy and reliability is also believed to be high because of the status of its data subjects, as they are a population from which administrative information can be efficiently collected.

Timeliness

The annual extract from p-NOMIS reflects the data as of 30 June each year, aligning with the ABPEs' mid-year reference point. This timing ensures consistency across data sources and supports comparability within the ABPEs.

The 2024 p-NOMIS extract, used in the 2024 ABPEs, was delivered to the ONS in August 2024. This represents a short lag between the reference date and delivery, providing timely data for inclusion in the ABPEs.

18 . Births and deaths registrations

Births and deaths registrations data are sourced from the Local Registration Service in partnership with the [General Register Office \(GRO\)](#), which records all births and deaths in England and Wales. These data are used in admin-based population estimates (ABPEs) to estimate population change because of births and deaths. Death registrations are also used to filter the records forming the Statistical Population Dataset as described in [Section 2: Use of data sources](#).

Birth registrations are not currently used in stock creation, as babies aged under 1 year are effectively captured through the Personal Demographics Service (PDS). More information on the quality of the births and deaths registrations, can be found in the [Births Quality and Methodology Information](#) and [Mortality Statistics in England and Wales Quality and Methodology Information](#).

Coverage

The Civil Registration System provides very high coverage of births and deaths occurring in England and Wales. Elements of undercoverage resulting from births and deaths happening to usual residents while abroad are assumed to be offset by corresponding overcoverage from those events happening within England and Wales to people who are not usual residents.

The data used in the 2024 ABPEs relate to births and deaths occurring in the 12 months to end-June 2024 and registered at any time up to April 2025. This will effectively cover all births. A small adjustment is made to reflect the estimated number of deaths occurring in the ABPE reference period but not appearing in the registration data. More information on late registration of deaths is provided in our [Impact of registration delays on mortality statistics in England and Wales article](#).

Accuracy and completeness

The data are considered highly accurate and complete because of the legal framework and robust verification processes in place. Under the [Births and Deaths Registration Act 1953](#), there is a legal requirement in England and Wales for parents to register a birth within 42 days, and for deaths to be registered within five days unless the death is referred to a coroner or there are exceptional circumstances, for example, stillbirths. This legal obligation ensures a high level of compliance and data reliability.

While birth notifications - initial records typically completed by a midwife or doctor shortly after birth - are timelier, registrations are more complete and of higher quality. For birth registrations, the informant, usually the mother or both parents, must verify the information provided to the registrar. The data are entered into the Registration Online (RON) system, which includes built-in validation checks to reduce entry errors. We carried out additional quality assurance upon receipt of the data. Supplying false information is a criminal offence, hence birth registrations accuracy is generally believed to be high. The accuracy of sex and date of birth is particularly high, while names and addresses are generally reliable, though minor spelling or formatting inconsistencies may occur.

For death registrations, the informant, typically a close relative, hospital official, or coroner's officer, must verify the information before submission. Sex is recorded based on the Medical Certificate of Cause of Death (MCCD) and is considered highly reliable. Names and addresses are generally accurate, though address information may be affected by recent moves or temporary residence at the time of death.

Timeliness

Death registrations data are supplied to the Office for National Statistics (ONS) daily when a registration is made through the RON system. As described in the Coverage section, the registration data used in the ABPEs align very well with the reference period of the estimates, with a small adjustment made for very late registration of deaths.

19 . Related links

[Mid-year admin-based population estimates for England and Wales quality and methods guide](#)

Methodology | Released 30 July 2025

How we produced the statistics, their strengths and limitations and further quality information.

[Understanding mid-year admin-based population estimates for local authorities in England and Wales](#)

Methodology | Released 30 July 2025

This methodology brings together important information about our mid-year admin-based population estimates for England and Wales.

20 . Cite this methodology

Office for National Statistics (ONS), released 30 July 2025, ONS website, methodology article, [Quality overview of data sources used in mid-2024 admin-based population estimates for England and Wales](#)