



Llywodraeth Cymru
Welsh Government

RESEARCH, DOCUMENT

Estimating numbers of children not in state education using linked administrative data

This research report explores the feasibility of providing an estimate of children potentially missing from state education using linked administrative data in the SAIL Databank.

First published: 18 September 2024

Last updated: 20 November 2024

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data-html> for the latest version.

Get [information on copyright](#).

Contents

[Summary](#)

[Introduction](#)

[Approach](#)

[Method](#)

[Findings](#)

[Conclusions](#)

[Acknowledgements](#)

[Footnotes](#)

[Contact details](#)

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data-html> for the latest version.

Get [information on copyright](#).

This report was revised on 20 November 2024, after initial publication on 18 September 2024.

Since publication a problem emerged with an underlying table used to produce the estimates. This meant the resident population was over-estimated, as well as the number of children missing from state education. This was particularly the case for Gwynedd. A new version of the underlying table has been used in the revised analysis.

Furthermore, a detailed review of the data used for the estimate indicated that not all relevant records were being selected from the PLASC/EOTAS data. This also led to a slight over-estimate of the number of children missing from state education. A small amendment to the method was required to correct this.

The estimates have been re-produced using the new underlying table and the amended method. The revised estimates of the resident child population and those missing from state education are both slightly lower overall, whilst both are notably lower for Gwynedd.

1. Summary

The aim of this study was to explore the feasibility of providing an estimate of children potentially missing from education using linked administrative data in the SAIL Databank.

The approach found that approximately 6.0% of children (24,000) in the dataset of GP registrations could not be found in the Pupil Level Annual School Census (PLASC) or Educated Other Than At School (EOTAS) data on 20 April 2021. This is likely to represent an upper band estimate of children missing from state

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

education. Reasons for children not found in PLASC or EOTAS data may include the following:

- Children educated in independent schools (approx. 8,000)
- Electively home-educated children
- Children educated in England
- Data linkage issues, possibly due to discrepancies in recording of name, address or date of birth
- Over-coverage in GP registration data (e.g. due to children moving away from Wales but not de-registering with their GP)
- Other reasons

The approach suggests there could be a wide variation of children missing from the education data between local authorities, potentially between around 3.0% and 14.9%. These differences may be partly explained by the presence of independent schools and proximity to schools in England of some local authorities. The approach also suggests a difference between school ages (possibly between about 4.5% to 8.1% increasing with age). These are estimates only given the limitations of the approach.

This research has been undertaken as part of [Administrative Data Research \(ADR\) Wales](#) planned [programme of work 2022-2026](#) funded by the Economic Social Research Council.

2. Introduction

Section 7 of the Education Act 1996 (“the 1996 Act”) places a duty on parents to ensure that a child of compulsory school age receives efficient full-time education suitable to the child’s age, ability and to any additional learning needs the child may have, either by regular attendance at school or otherwise. Section 436A of the 1996 Act states that local authorities (LAs) must make

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

arrangements to identify children not receiving suitable education and to have regard to guidance issued by the Welsh Ministers.

To support LAs to fulfil their duty, the Welsh Government is proposing to make regulations to require LAs to establish a database of children missing education and issue statutory guidance to LAs using existing law. It is envisaged that the database would be produced by linking education data to GP registration data to identify those children missing from state education.

To test the feasibility of this approach, we have attempted to replicate this data linkage using de-identified data in the Secure Anonymised Information Linkage (SAIL) Databank. This report sets out our approach and findings.

3. Approach

Our approach utilises administrative data linking techniques using individual-level health and education datasets.

A dataset of GP registrations, called the Welsh Demographic Service Dataset (WDSD), described below, is used as the base data. Using linked individual-level data allows us to get a more precise estimate than comparing aggregate figures because, if we simply subtract the number of children in the education data from the number of children from a population dataset such as the WDSD, the number produced will reflect the number of children missing from both data sources but it would not be possible to determine how many children were missing from the education data, which we are specifically interested in. An aggregate comparison may also underestimate the true number of missing children. For example, if there were 10 children missing from the education data and 10 children missing from the WDSD (whether the same or different children) an aggregate comparison would indicate that there were zero missing children.

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data-html> for the latest version.

Get [information on copyright](#).

The method was developed in Structured Query Language (SQL) DB2 software and has been reviewed, modified and refined where appropriate.

4. Method

Health and education datasets for each child in Wales were de-identified and deposited to SAIL via a 'Split-File' method using Digital Health and Care Wales (DHCW) as a trusted third party. Via this process, personal identifier information was removed and replaced with an Anonymised Linking Field (ALF), whilst address details (where these had been provided) were removed and replaced with a Residential Anonymised Linking Field (RALF). For each pupil an individual record number (IRN) was generated to facilitate linked data analysis between the education datasets once uploaded. In both health and education datasets date of birth was replaced with week of birth (WOB) to further de-identify the data. Once in SAIL valid ALF, RALF and IRN information was encrypted to further protect the data before being made available for analysis.

The WDSO contains individual-level demographic and multiple address registrations for all individuals of any age who have ever registered with a GP in Wales. This includes both residential and non-residential addresses (i.e. communal establishments, which include boarding schools) along with start and end dates of residence. This health dataset was linked to education data, firstly using the ALF and WOB to link to a core table which contained all recorded ALFs for pupils in Wales. Using the encrypted IRN this core table was then linked to two other education tables, namely the Pupil Level Annual School Census (PLASC) and Educated Other Than At School (EOTAS) datasets.

The PLASC dataset covers all pupils registered in local authority maintained schools on the annual school census date.^[footnote 1]

The EOTAS dataset covers pupils that local authorities place in Pupil Referral Units (PRUs), or other forms of alternative provision referred to as Educated Other Than at School.^[footnote 2] It does not include electively home

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

educated children.

As the WDS D can be updated continuously based on GP registration dates, any dates can be specified as 'cut-off dates' to select data to gain a snapshot estimate of the number of people resident in Wales at a particular point in time. By also specifying birth dates to include it is possible to gain an estimate of residents for particular age groups.

The PLASC date of 20 April 2021 was chosen as the base date for the estimate. PLASC is usually taken in January but in 2021 was postponed until April due to the pandemic. This change in timing may have had an effect on the match rate between the WDS D and PLASC and EOTAS data but it is not possible to determine whether this would have been a positive or negative effect. Using data for 2021 allowed comparability with a parallel exercise undertaken by the Office for National Statistics (ONS) that involved linking PLASC data to the 2021 Census. This date was used as the cut-off date so that only records for residents registered by that date were selected for use in the analysis. Records relating to registrations ending before 20th April 2021 were excluded. Only those who were resident in Wales were included, based on 2011 Lower Super Output Area (LSOA) codes.

WOB was also used to calculate age at the start of the academic year (i.e. on 31 August 2020) for each resident. Cut-off dates were then applied to WOB data so that only records for those who were of statutory school age at the start of the 2020/21 academic year were selected for use in the analysis (i.e. aged 5 to 15 years on 31 August 2020). WDS D records relating to residents whose age was above or below these ages at that point were excluded.

By applying these cut-off dates to the linked WDS D, PLASC and EOTAS datasets it was possible to develop an estimate of residents of statutory school age who were in state education in 2021. A subtraction of this estimate from that for all statutory school-age residents yielded an indication of the number of children potentially missing from the education system on 20 April 2021.

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

Records categorised as missing from PLASC or EOTAS data (and therefore potentially missing from state education) were categorised by whether a RALF was recorded as resident at a residential address. Records for which no RALF had been found in the WSDS dataset were categorised as having a non-residential address.

4.1 Comparison with ONS estimates

A parallel exercise to estimate children missing from education has been undertaken by ONS, the aim of which was to validate the ADR Wales estimate using other data sources. Under the legal gateway provided by Section 45A of the Statistics and Registration Services Act 2007 (as inserted by Section 79 of the Digital Economy Act 2017), the Welsh Government shares education data with the ONS to support the development of an administrative data-based census. We asked ONS to link this with 2021 Census data as an alternative approach to estimating the number of children not in state education.

ONS published their findings on 17 April 2024 (see [CT21_0206_Census 2021 – \(ONS\) Statistics \(ons.gov.uk\)](#)).^[footnote 3] Due to the core purpose of the Census (which is to count the ‘usually resident’ population) ONS used different definitions to produce the estimate to those used by ADR Wales. The differences in the definitions used by ONS and ADR Wales are explained below.

Estimating children missing from education: ONS and ADR Wales definitions

Core population dataset

ONS

2021 Census

ADR Wales

WDS

Study population

ONS

Not in PLASC or England School Census data in 2020-2021 academic year

ADR Wales

Not in PLASC or EOTAS data in 2020-2021 academic year

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

Reference date (i.e. date child is observed in the respective population dataset)

ONS

21 March 2021

ADR Wales

20 April 2021

Residence duration

ONS

12 months or more

ADR Wales

Any

Age reference date

ONS

21 March 2021

ADR Wales

31 August 2020

Birth date range

ONS

1 September 2004 to 31 August 2016

ADR Wales

1 September 2004 to 31 August 2015

Resultant age range at age reference date

ONS

4 to 16

ADR Wales

5 to 15

Some testing of the ADR Wales method was carried out using some of the ONS definitions (e.g. re-running to extract records just for those resident for 12 months or more and using the ONS age range). An adapted version of the SQL

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data-html> for the latest version.

Get [information on copyright](#).

code was created and run using these definitions to provide an adjusted estimate as a form of sensitivity analysis. This was summarised and extracted from SAIL, then compared with the estimate from ONS. Results are shown in Tables 5 and 6.

5. Findings

5.1 Main results

Table 1 shows the estimate from the ADR Wales approach for all resident children on school census day 2021 broken down by single year of age. These figures include all resident children of statutory school age on 31 August 2020, whether permanent or temporary residents on school census day.

Note that all figures have been rounded to the nearest ten. As age and local authority breakdowns are aggregated independently the totals vary between Tables 1 and 3.

Table 1: Estimated number of children resident in Wales, by PLASC/EOTAS registration status and age, 20 April 2021

Age at 31 August 2020	All Resident Children (WSDS)	Previously Recorded in PLASC or EOTAS (but not on 20th April 2021)	No PLASC / EOTAS record	% No record or Previously in PLASC / EOTAS
5	35,700	500	1,200	4.8%
6	35,030	580	1,000	4.5%

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

Age at 31 August 2020	All Resident Children (WDSO)	Previously Recorded in PLASC or EOTAS (but not on 20th April 2021)	No PLASC / EOTAS record	% No record or Previously in PLASC / EOTAS
7	35,960	780	1,000	4.9%
8	37,050	810	980	4.8%
9	37,810	990	960	5.2%
10	37,240	1,080	970	5.5%
11	37,390	1,540	970	6.7%
12	37,620	1,660	960	7.0%
13	36,050	1,710	850	7.1%
14	35,400	1,730	1,010	7.7%
15	35,070	1,800	1,030	8.1%
All aged 5 to 15	400,320	13,180	10,930	6.0%

Source: SAIL Databank

Note: Children previously recorded in PLASC or EOTAS or without a PLASC or EOTAS record (i.e. the third and fourth columns) may be in independent schools, electively home educated or educated in England. Some of these

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

children may also have PLASC or EOTAS records but could not be linked to the WDSO due to discrepancies in recording of name, address or date of birth.

Table 1 shows that 6.0% of children on the WDSO were not recorded in PLASC or EOTAS data on 20 April 2021.

Around 3% of children on the WDSO could not be found in PLASC or EOTAS data at all, with little variation by age.

It was not possible to identify children living in Wales who attended schools in England, nor children who attend independent schools. The 'Previously in PLASC or EOTAS' category figures for ages 11 and over are higher than those for younger ages so it is possible that some children attend primary school in Wales followed by a secondary school in England. This may be due to localised patterns of education provision and/or transport networks, or other factors. On the other hand, at these ages children may have moved to an independent school for secondary education for similar reasons. Again, it is not possible to determine this using available data.

For some of those children who could not be found in PLASC or EOTAS data it is possible that they are not truly missing but their records in the education data could not be linked to the WDSO (e.g. due to discrepancies in recording of name, address or date of birth).

There may also be an element of over-coverage in the WDSO. Following the 2011 Census, **the ONS undertook some exploratory work to explore future options for producing population and small area socio-demographic statistics for England and Wales.**^[footnote 4] Although this work is now over 10 years old, it contains some findings that are relevant to our analysis. One of the data sources that the ONS explored was the NHS Patient Register, which contained a record for every person registered with an NHS GP in England and Wales. In this respect it is broadly equivalent to the WDSO but covers England as well as Wales. The ONS found that the NHS Patient Register was subject to

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

both over-coverage and under-coverage. The ONS identified various possible reasons for coverage issues, with the ones listed below likely to be relevant to our analysis:

- Multiple area registrations – where someone may be registered with more than one different GP practice and also hold more than one NHS number (likely impact: over-coverage)
- Duplicate NHS numbers – where two or more people have the same NHS number (likely impact: under-coverage)
- Lags in the recording of deaths (likely impact: over-coverage)
- Emigration from the UK without de-registering from the NHS (likely impact: over- coverage)
- Immigration to the UK without registering with a GP (likely impact: under-coverage)
- Registering with a private GP only (likely impact: under-coverage)
- Definition differences – with the ONS definition only counting ‘usual residents’ defined as those who, on Census Day, were in the UK and had stayed or intended to stay in the UK for a period of 12 months or more or had a permanent UK address and was outside the UK and intended to be outside the UK for less than 12 months (likely impact: over-coverage)

The ONS found that, in April 2011, the NHS Patient Register record count was 4.3% greater than the 2011 Census population estimate of England and Wales. This suggests that, whilst there are elements of under-coverage in GP registration data, the impact of over-coverage is likely to be greater.

The ONS is currently undergoing a transformation programme on the production of population and migration statistics using administrative data. This work has involved assessing the suitability of various datasets for providing population and migration statistics.^[footnote 5] As part of this, the ONS has studied the Personal Demographics Service (PDS). The PDS contains demographic data for those who have interacted with an NHS Service in England, Wales, and the Isle of Man, including through GP practices and

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

hospital visits, so it has slightly broader coverage than the WDS and the NHS Patient Register. As was the case with the NHS Patient Register, the ONS found evidence of both over and under coverage. The ONS found a decrease in GP registrations and address changes from April 2020 (the start of the coronavirus (COVID-19) pandemic) followed by an increase in the first half of 2021, rising above pre-coronavirus pandemic levels. The ONS suggested this may be because of a backlog of people returning to the GP after lockdown, updating their details because of the vaccination programme, or people moving house and changing their address as the property market reopened. This has relevance for our analysis because it may suggest a large number of people and families registered with a GP in 2021, leading to a higher estimate in the WDS.

Taking all these factors into consideration suggests that the overall percentage of children missing from the education data (6.0%) is likely to represent an upper band estimate of children missing from state education. For instance, Section 5.2 notes that there are around 8,000 children aged 5 to 15 attending independent schools in Wales. If we assume that all these children are resident in Wales and subtract this estimate from the number missing from the education data, this would bring down the percentage to 4.0%. Taking into account the approximately 4,000 children aged 5 to 15 that are known to be electively home educated would bring the percentage down to 3.0%. It is more difficult to quantify the effects of attending school in England, matching failures and over-coverage in the GP registration data but were we able to do so, this would bring the percentage down even further.

Table 2 provides a breakdown of the number of children who could not be found in the PLASC / EOTAS data at all who were recorded at a residential or non-residential address for each age. Non-residential addresses are communal establishments and could include care homes, homeless hostels, healthcare setting or boarding in an independent school.

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

Table 2: Number of children who could not be found in PLASC / EOTAS by residential versus non-residential address and age, 20 April 2021

Age at 31 August 2020	No PLASC / EOTAS record: all	No PLASC / EOTAS record: at a residential address	No PLASC / EOTAS record: at a non residential address
5	1,200	1,070	130
6	1,000	910	90
7	1,000	910	90
8	980	900	80
9	960	870	90
10	970	880	90
11	970	880	90
12	960	850	110
13	850	740	110
14	1,010	850	160
15	1,030	840	190

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data-html> for the latest version.

Get [information on copyright](#).

Age at 31 August 2020	No PLASC / EOTAS record: all	No PLASC / EOTAS record: at a residential address	No PLASC / EOTAS record: at a non residential address
All aged 5 to 15	10,930	9,700	1,230

Source: SAIL Databank

Around 89% of children who could not be found in PLASC / EOTAS data were resident at a residential address in 2021. For all ages, fewer than 20% of children who could not be found in PLASC / EOTAS data were living at a non-residential address.

Table 3 shows the estimated resident school-aged children broken down by local authority in Wales. Note that local authority is based on residential address, not school address. These figures have been aggregated from recorded LSOA code by local authority for each resident, so PLASC figures shown here are likely to differ from published PLASC counts by local authority for 2021.

Table 3: Estimated number of children resident in Wales, by PLASC/EOTAS registration status and local authority, 20 April 2021

Local authority	All resident children (WSD)	Previously recorded in PLASC or EOTAS (but not on 20 April 2021)	No PLASC / EOTAS record	% No record or previously in PLASC / EOTAS
Isle of Anglesey	8,480	200	220	5.0%

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

Local authority	All resident children (WDS)	Previously recorded in PLASC or EOTAS (but not on 20 April 2021)	No PLASC / EOTAS record	% No record or previously in PLASC / EOTAS
Gwynedd	14,340	400	430	5.8%
Conwy	13,580	420	400	6.0%
Denbighshire	12,570	460	320	6.2%
Flintshire	20,310	920	750	8.2%
Wrexham	18,240	870	600	8.1%
Powys	15,040	920	710	10.8%
Ceredigion	7,740	330	210	7.0%
Pembrokeshire	15,290	570	420	6.5%
Carmarthenshire	23,400	700	580	5.5%
Swansea	29,730	690	510	4.0%
Neath Port Talbot	18,090	350	230	3.2%
Bridgend	18,690	370	320	3.7%

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

Local authority	All resident children (WDS)	Previously recorded in PLASC or EOTAS (but not on 20 April 2021)	No PLASC / EOTAS record	% No record or previously in PLASC / EOTAS
Vale of Glamorgan	17,870	490	380	4.9%
Rhondda Cynon Taf	31,670	530	420	3.0%
Merthyr Tydfil	8,160	160	160	3.9%
Caerphilly	23,430	420	290	3.0%
Blaenau Gwent	8,450	160	100	3.1%
Torfaen	12,390	250	200	3.6%
Monmouthshire	10,980	800	840	14.9%
Newport	22,530	790	720	6.7%
Cardiff	49,320	2,370	2,090	9.0%
All Local Authorities	400,300	13,170	10,900	6.0%

Source: SAIL Databank

Note: Children previously recorded in PLASC or EOTAS or without a PLASC or

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

EOTAS record (i.e. the third and fourth columns) may be in independent schools, electively home educated or educated in England. Some of these children may also have PLASC or EOTAS records but could not be linked to the WDS due to discrepancies in recording of name, address or date of birth.

The percentage of resident children with no PLASC or EOTAS record, or previously recorded in either dataset, and therefore potentially missing from state education, are highest in Monmouthshire and Powys. Those with no PLASC or EOTAS record at all make up no more than 8% of resident children in these authorities, however.

It should be noted that the caveats discussed in relation to Table 1 all apply to Table 3 too. The caveats around children attending independent schools or schools in England are particularly relevant to local authority level results as these factors have geographical patterns. For example, Monmouthshire and Powys share a border with England whilst Monmouthshire also has a relatively high number of children attending independent schools. This is discussed further in Sections 5.2 and 5.4.

Table 4 provides a breakdown of the number of children who could not be found in the PLASC / EOTAS data at all who were recorded at a residential or non-residential address for each age.

Table 4: Number of children who could not be found in PLASC / EOTAS by residential versus non-residential address and local authority, 20 April 2021

Local authority	No PLASC / EOTAS record: all	No PLASC / EOTAS record: at a residential address	No PLASC / EOTAS record: at a non-residential address
Isle of Anglesey	220	180	40

Local authority	No PLASC / EOTAS record: all	No PLASC / EOTAS record: at a residential address	No PLASC / EOTAS record: at a non-residential address
Gwynedd	430	370	60
Conwy	400	370	30
Denbighshire	320	290	30
Flintshire	750	670	80
Wrexham	600	560	40
Powys	710	550	160
Ceredigion	210	180	30
Pembrokeshire	420	370	50
Carmarthenshire	580	540	40
Swansea	510	470	40
Neath Port Talbot	230	210	20
Bridgend	320	300	20
Vale of	380	340	40

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data-html> for the latest version.

Get [information on copyright](#).

Local authority	No PLASC / EOTAS record: all	No PLASC / EOTAS record: at a residential address	No PLASC / EOTAS record: at a non-residential address
Glamorgan			
Rhondda Cynon Taf	420	390	30
Merthyr Tydfil	160	150	10
Caerphilly	290	270	20
Blaenau Gwent	100	100	0
Torfaen	200	180	20
Monmouthshire	840	630	210
Newport	720	640	80
Cardiff	2,090	1,920	170
All Local Authorities	10,900	9,680	1,220

Source: SAIL Databank

In most local authorities, the percentage of children who could not be found in PLASC / EOTAS data who were living in a non-residential address was fewer than 20%. In Powys and Monmouthshire, the percentages were 23% and 25%

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data-html> for the latest version.

Get [information on copyright](#).

respectively.

Table 5 shows the results of applying the ONS definitions to WDS data. The age reference and reference date are both 21 March 2021 (as opposed to 31 August 2020 and 20 April 2021 above) for those born between 2004/05 and 2015/16 academic years so includes ages 4 to 16. We have also attempted to align with the ONS 'usual residency' definition by restricting our cohort to those children registered with a GP in Wales for 12 months or more prior to the reference date. It should be noted that this is a narrower definition than the ONS one. The ONS definition includes those who had stayed or intended to stay in the UK for a period of 12 months or more or had a permanent UK address and was outside the UK and intended to be outside the UK for less than 12 months.

Table 5: Estimated number of children resident in Wales for 12 months or more, by PLASC/EOTAS registration status and age, 21 March 2021

Age at 21 March 2021	All resident children (WDS)	Previously recorded in PLASC or EOTAS (but not on 20 April 2021)	No PLASC / EOTAS record	% No record or previously in PLASC / EOTAS
4	14,080	120	610	5.2%
5	31,050	360	1,070	4.6%
6	31,390	510	920	4.6%
7	31,910	640	890	4.8%
8	33,120	760	840	4.8%

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

Age at 21 March 2021	All resident children (WDS)	Previously recorded in PLASC or EOTAS (but not on 20 April 2021)	No PLASC / EOTAS record	% No record or previously in PLASC / EOTAS
9	34,050	830	840	4.9%
10	34,910	950	870	5.2%
11	34,150	1,190	840	5.9%
12	34,210	1,520	810	6.8%
13	34,110	1,570	750	6.8%
14	32,940	1,570	750	7.0%
15	32,340	1,640	810	7.6%
16	18,080	930	450	7.6%
All Aged 4-16	396,340	12,590	10,450	5.8%

Source: SAIL Databank

Note: Children previously recorded in PLASC or EOTAS or without a PLASC or EOTAS record (i.e. the third and fourth columns) may be in independent schools, electively home educated or educated in England. Some of these children may also have PLASC or EOTAS records but could not be linked to the WDS due to discrepancies in recording of name, address or date of birth.

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

Despite the broader age definition, the overall resident child population in Table 5 is around 4,000 lower than that in Tables 1 and 3 due to the more restricted residency definition. The PLASC and EOTAS counts are similarly lower.

The proportion of those potentially missing from state education is slightly lower in Table 5 at 5.8% than that in Tables 1 and 3 (6.0%), with a similar distribution across the categories to that shown in Table 1. There is some variation in the distribution by age of those with no PLASC or EOTAS record. The proportion of those with no PLASC or EOTAS record at all is around 3%, as before.

Tables 6 and 7 provides a comparison of the ADR Wales estimate based on ONS definitions, and the published ONS figures themselves, broken down by age.

Table 6: Estimated number of children resident in Wales for 12 months or more, by 2021 Census versus WDS datasets, 21 March 2021

Age	2021 Census: all resident children	WDS: all resident children
4	14,605	14,080
5	32,890	31,050
6	32,605	31,390
7	32,990	31,910
8	34,395	33,120
9	35,380	34,050

Age	2021 Census: all resident children	WDS: all resident children
10	35,495	34,910
11	35,265	34,150
12	35,090	34,210
13	35,160	34,110
14	33,970	32,940
15	33,300	32,340
16	18,645	18,080
All aged 4 to 16	409,790	396,340

Source: ONS and SAIL Databank

The comparison of the 'All resident children' totals in Table 6 shows that the ONS estimate of number of usual resident children from the 2021 Census is around 13,000 higher than the estimate from the WDS when restricting to children who have been registered with a GP for 12 months or more. It should be noted that unlike the comparison between the WDS and PLASC/EOTAS undertaken above, the comparison between the WDS and Census estimates of all resident children is not based on a linked data analysis. Therefore, we cannot say there are 13,000 children 'missing' from GP registration data because some children may be missing from both sources. It is also possible that some children are in the GP registration data but not in the Census, thus

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

leading to a reduction in the difference between the two sources. Furthermore, whilst in Tables 5, 6 and 7 we have adjusted the residency definition to be more in line with that of ONS, the definitions do not exactly match. As stated above, our definition here is a narrower one. If we removed the 'resident for 12 months or more' restriction from our estimate of the number of children from the WDS then we find more children than the Census estimate.^[footnote 6]

Table 7: Estimated number and percentage of resident children not in state school, by 2021 Census versus WDS datasets, 21 March 2021

Age	2021 Census: not in state school	2021 Census: % not in state school	WDS: no record or previously in PLASC / EOTAS	WDS: % no record or previously in PLASC / EOTAS
4	635	4.3%	730	5.2%
5	1,400	4.3%	1,430	4.6%
6	1,200	3.7%	1,430	4.6%
7	1,395	4.2%	1,530	4.8%
8	1,385	4.0%	1,600	4.8%
9	1,485	4.2%	1,670	4.9%
10	1,655	4.7%	1,820	5.2%
11	1,770	5.0%	2,030	5.9%

Age	2021 Census: not in state school	2021 Census: % not in state school	WDS: no record or previously in PLASC / EOTAS	WDS: % no record or previously in PLASC / EOTAS
12	2,040	5.8%	2,330	6.8%
13	2,175	6.2%	2,320	6.8%
14	2,440	7.2%	2,320	7.0%
15	2,665	8.0%	2,450	7.6%
16	1,715	9.2%	1,380	7.6%
All aged 4 to 16	21,960	5.4%	23,040	5.8%

Source: ONS and SAIL Databank

Despite differences in the numbers potentially missing from the state education system it is notable that the estimate of usually resident children for whom this is the case under either approach is around 5% or 6%.

5.2 Independent schools

Published figures indicate that in 2021, around 8,000 children aged 4 to 16 attended independent schools in Wales (although not all schools provided data for 2021 so the actual figure is likely to be higher).^[footnote 7] Within this, pupil counts tend to increase in the 9 to 12 age group, possibly due to some children

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

moving to the independent sector for secondary education. The age profile of pupils previously in PLASC or EOTAS shown in Tables 1 and 5 would indicate agreement with this, with the biggest increase in the numbers of children previously recorded in PLASC or EOTAS but not in 2021, occurring at age 11 or 12. This suggests that many are not missing from the education system completely but are moving between sectors for secondary education. This may to some extent explain the proportionally higher numbers for some local authorities in Table 2 which tend to have higher proportions of children attending independent schools, for example Monmouthshire and Cardiff.

5.3 Electively home educated children

In the 2020-21 school year, 4,342 children were known to be electively home educated in Wales ^[footnote 7]. The numbers tend to increase at around age 11, again suggesting that elective home education may be another contributing factor to the increase in numbers of children previously recorded in PLASC or EOTAS but not in 2021, occurring at age 11 or 12.

5.4 Children resident in Wales educated in England

Children resident in local authorities in Wales which have a border with England can attend school either locally or across the border. This could be the case at any point within statutory school age, but local provision may mean it is practical to attend a primary school in Wales followed by a secondary school in England. This may be seen partly in Table 3 where for most authorities those children who only appear in PLASC or EOTAS prior to April 2021 account for fewer than 5% of resident children. The exceptions are Monmouthshire (7.3%) and Powys (6.1%), both of which share a border with England. Wrexham and Flintshire, the other two local authorities sharing a border with England also have a higher percentage of such children than the Wales average of 3.3%, at 4.8% and 4.5% respectively though Cardiff and Ceredigion also have rates above 4%.

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

As shown in Section 4, ONS holds education data for England which they were able to incorporate into their linked data analysis. This means that, if a child resident in Wales attended school in England they would not be flagged as missing from state education. Whilst the ONS analysis estimated a lower percentage of children missing from education for Wales overall than the SAIL analysis, this was particularly the case for the four local authorities that share a border with England.

5.5 In PLASC/EOTAS 2020, but missing in 2021

Around 3,700 children aged 5 to 15 resident in Wales in 2021 were last recorded in PLASC/EOTAS in 2020. This suggests that they left the state school system in Wales during the pandemic. Reasons for this may include:

- Relocating outside Wales but without re-registration/de-registration with a GP
- Transferring to an independent school in 2021 (as discussed above)
- Continuing home schooling following the pandemic

Media coverage during the pandemic suggested new and emerging trends in education, including migration, working from home and the take-up of home schooling. It is not possible to further explore the possible impact of these trends after 2020 using the datasets available currently.

6. Conclusions

The aim of this study was to explore the feasibility of providing an estimate of children potentially missing from education using health and education data.

The estimate of children missing from education is based on records for April

2021 and should be considered as indicative. More recent estimates may provide different results. Within this approach alternative sub-methods may be applied e.g. linking the datasets using the ALF and week of birth (as has been done here) or just on ALF. This alteration was made to make more certain that the correct records were matched between health and education datasets. However slight adjustments such as this can cause variations and re-categorisation of records for those potentially missing from education.

ONS and ADR Wales estimates suggest that between 5% and 6% of children could not be found in the education data in March / April 2021. In both analyses, Monmouthshire was found to be a local authority with a high percentage of such children. The ADR Wales analysis was also able to distinguish between those previously recorded in the education data (i.e. prior to 2021) and those who cannot be found in the data at all. Around 3% of children (around 10,900 children) could not be found in the education data at all.

Both the ADR Wales and ONS estimates provide an indication of children who are potentially missing from the education system, but neither approach takes account of those educated in the independent school sector. It is also important to note that some children counted as missing in either the ADR Wales or ONS approach may simply be missing from the linked data, i.e. their education records cannot be linked to their population records due to inconsistencies in recording of linking fields. In this respect, the estimates presented here are likely to be upper band estimates.

The ONS estimate of children missing from the education data was lower than that of ADR Wales. This may be because individuals recorded their identifiable data more accurately in the Census than when registering for a GP. It may also be because Census records can be definitively tied to a particular timepoint, i.e. Census Day, whereas GP registration data suffers from time lags due to delays in people updating their records. These factors may have led to ONS obtaining a better linkage rate between the population and education records. As discussed above, there may also be some over-coverage in GP registration data leading to

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

our analysis finding more children who could not be found in the education data than the ONS.

Furthermore, ONS also had access to education data for England which meant pupils living in Wales but educated in England were not flagged as missing in their analysis. On the other hand, ONS did not have access to EOTAS data so such children in this dataset would have appeared as missing in the ONS analysis.

Nevertheless, the fact that the ONS estimate, although lower, is not too dissimilar to the ADR Wales one does suggest that linking education data to GP registration data can provide a reasonable estimate of children missing from the state education system. Furthermore, in their assessment of administrative data sources used to develop their Statistical Population Dataset for England and Wales, the ONS assessed the PDS (NHS registration data) as having excellent coverage.^[footnote 8] It could be argued that this provides further support for using NHS registration data (in the form of GP registrations in our case) as a way of flagging children who may be missing education.

However, there is likely to be an element of over-coverage in GP registration data which may lead to an over-estimate of 'missing' children. Moreover, many children missing from the education data may be attending independent schools or schools in England. This could be determined by local authorities through follow-up with families of children flagged as 'missing'. It should also be noted that a small number of children may be missing from both GP registration and education datasets.

Acknowledgements

The data analysis was carried out in the Wales Institute of Social and Economic Research and Data (WISERD) Education Data Lab project 'Participation in and

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

Progression Through Education in Wales' within the SAIL Databank, funded under the [Administrative Data Research \(ADR\) Wales Education theme ES/W012227/1 \(UK Research and Innovation\)](#) and [WISERD ES/S012435/1 \(UK Research and Innovation\)](#). This work uses data provided by patients and collected by the NHS as part of their care and support.

ADR Wales brings together data science experts at Swansea University Medical School, staff from the WISERD at Cardiff University and specialist teams within the Welsh Government to develop new evidence which supports the Programme for Government by using the SAIL Databank at Swansea University, to link and analyse anonymised data. The [ADRW Programme of Work 2022-2026](#) outlines the ten thematic areas that the ADR Wales team will focus their research on to help government address the most pressing issues facing society.

ADR Wales is part of ADR UK and funded by the Economic and Social Research Council (part of UK Research and Innovation).

This study acknowledges the [peer review and academic support of Dr Alex Sandu, Dr Jen Keating, Dr Katy Huxley, and Dr Rob French](#).

Footnotes

[1] [Schools' census results \(Welsh Government\)](#) July 2024

[2] [Pupils educated other than at school \(Welsh Government\)](#) September 2024

[3] [CT21_0206_Census 2021 \(ONS\)](#) April 2024

[4] [Beyond 2011: Administrative Data Sources Report: NHS Patient Register \(ONS\)](#) November 2012

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data.html> for the latest version.

Get [information on copyright](#).

[5] **Administrative sources used to develop the Statistical Population Dataset for England and Wales: 2016 to 2021 (ONS)** March 2023

[6] Although Tables 1 and 3 show lower numbers of children from the WDSD than the ONS figures overall, this is because we have used a narrower age range than ONS. Comparing numbers by age in Table 1 of our analysis and Table 1 in the ONS publication finds lower numbers in the ONS estimates for every age from 5 to 15.

[7] **Pupils educated other than at school (Welsh Government)** September 2024

[8] **Quantitative Quality Indicators produced for administrative sources used to develop Statistical Population Dataset version 4.0 (ONS)** March 2023

Contact details

Views expressed in this report are those of the researchers and not necessarily those of the Welsh Government.

For further information please contact:

Tony Whiffen

Email: adrwales@gov.wales

Social research number: 70/2024

Digital ISBN 978-1-83625-606-9

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data-html> for the latest version.

Get [information on copyright](#).



Dadansoddi ar gyfer Polisi



Analysis for Policy

This document may not be fully accessible.

For more information refer to our [accessibility statement](#).

This document was downloaded from GOV.WALES and may not be the latest version.

Go to <https://www.gov.wales/estimating-numbers-children-not-state-education-using-linked-administrative-data-html> for the latest version.

Get [information on copyright](#).