

User Guide: Children Looked After Datasets - Wales



Authors: Gemma Allnatt, Martin Elliott , Laura Cowley, Alex Lee, Laura North, Karen Broadhurst, Lucy Griffiths

Date: V1.1 May 2025

This guide is for new and existing researchers to assist on working with the Children Looked After Dataset - Wales.

It is intended as a starting point to help guide researchers who are looking to use the Looked After Children datasets as part of a research project

This document is correct at the point of publication.

Introduction to the dataset(s)

The dataset is accessible to accredited research users from the [SAIL Databank](#) either as a stand-alone dataset or as part of a pre-defined project linkage or datasets.

This guide summarises the content and potential uses of the Children Looked After (CLA) Census dataset. This is the primary dataset relating to children who are looked after in Wales (and its subsets: the adoptions of children looked after dataset; the care leavers aged 16 and over dataset; the care leavers at 19th birthday dataset; and the educational qualifications of care leavers dataset). This document aims to guide researchers who are already researching or are interested in using these datasets to investigate the experiences of children and young people in care in Wales.

This guide has been produced by researchers working together within the ADR Wales Social Justice team at Swansea University, the Children's Social Care Research and Development Centre (CASCADE) and the now completed Family Justice Data Partnership (FJDP). It provides a concise narrative of the current learning about the CLA data. The information presented in this guide has been reviewed by ADR Wales colleagues with expertise within this thematic area and is accepted to be accurate at the point of publication.

How is this data collected?

In Wales, children's social care services are delivered at local government level, via local authorities. Local authorities are obligated to deliver to the Welsh Government yearly individual level information about those who are supported by children's social services. The Welsh Government collect and maintain several datasets relating to children receiving services from social care, which includes those focusing on children who are looked after.

These datasets contain rich situational and individual-level data items. The information collected in the CLA (formerly Looked After Children) Census dataset allows the trajectories of individual children through the care system to be objectively described in detail. Subsets of the main CLA dataset also provide snapshots of information about the educational achievements of individual children, and their experiences upon leaving care if they are adopted or age out of the social care system. Each subset dataset must be linked to the main CLA Census dataset to retrieve child characteristics (week of birth, gender, anonymised linking field).

Structure of the datasets

1) Children Looked After Census dataset

The CLA Census dataset within the Secure Anonymised Information Linkage (SAIL) Databank spans from 1 April 2002 – 31 March 2021 and the data continues to be collected annually. Its purpose is to provide a complete record of all individual 'episodes' of care that a child has experienced throughout a year and record any changes in the child's placement and legal status. The data available for this period consists of approximately 310,550 recorded episodes, reflecting data for 37,775 unique children.

This dataset has an 'episodic' structure, with the period during which a child is looked after by the local authority being broken down into 'episodes' of care. An episode is when a child is in the care of a local authority for more than 24 hours. A new episode begins every time a child starts to be looked after; there is a change in the child's legal status; and/or the child's placement changes. Episodes are recorded in consecutive date order. Where episodes are consecutive, the new episode will start on the same day as the previous episode finished. A 'period' of care may consist of only one 'episode' or may consist of multiple episodes.

If an episode of care is still open on 31 March in a census year, then the episode end date and the reason for the end of the episode are left blank, and the following year's record starts with the same episode details. In cases where an open episode of care continues unchanged from year-to-year, a record is still submitted every year for that child. These are essentially duplicates of the episodes recorded in the previous year(s). Therefore, it is necessary to remove duplicate records and restructure the data, so that episodes of care for each child can be properly sequenced. Variables available within the main CLA Census dataset include:

- Child characteristics:
 - Gender (pre-2016, child sex was recorded, recent collections record child gender identity at the census date)
 - Unique Pupil Number (UPN), predominantly recorded for children aged over four years old
 - Ethnicity (broader categories used following the Social Services and Well-being (Wales) Act (2014))
 - Disability status
 - Asylum-seeking status
 - Week of birth
 - Lower Super Output Area (LSOA) of home postcode (for anonymisation reasons)
 - Local authority responsible for the child.

- Care episodes:
 - Episode start date
 - Episode end date
 - Reason for episode starting
 - Reason for episode finishing
 - Category of need (most relevant category, recorded at the beginning of each care period)
 - Legal status
 - Placement type
 - LSOA of placement postcode (for anonymisation reasons).

2) Adoptions of children looked after dataset

At the time of writing, this dataset includes details of all adoptions of children looked after in Wales from 1 April 2002 – 31 March 2021 and is continuously updated each year. There are 5,255 records for 5,240 children.

Variables include the date that the local authority decided adoption was in the child's best interests; the date that the child was matched with prospective adoptive parents; the date of placement for adoption; and the date that the child was adopted and ceased to be looked after. This dataset also records information about the characteristics of the adopter(s), including gender; the number of adopters; their legal status (i.e. marital status); and their ethnicity. There is also a flag to indicate whether the child has been adopted by their former foster carers.

3) Care leavers aged 16 and over dataset

This dataset includes all children who ceased to be looked after during a collection year if they were aged 16 years or older at the time of care ending, and currently covers 1 April 2016 – 31 March 2021. Only the last date that care ceased is included. There are 3,315 records reflecting 3,240 individual children. There are more records than children because a small minority of children in this dataset leave care in one financial year, and return to care in the next financial year, following a brief being of not being looked after.

This dataset contains information about whether the care leaver was assigned a personal adviser or 'Pathway Plan' (setting out needs, views, future goals and support), if they were eligible; the type of accommodation they were living in on the census date; and the suitability of the accommodation.

4) Care leavers at 19th birthday dataset

This dataset includes all care leavers who turned 19 years old between 1 April 2002 and 31 March 2016. There are 4,940 records for 4,890 children. This dataset is no longer updated.

Information recorded in this dataset includes whether the care leaver was in contact with the local authority responsible for them when they left care; their involvement in employment, education, or training activities; and the type and suitability of their accommodation.

5) Educational qualifications of care leavers dataset

This dataset includes all children who ceased to be looked after during a collection year if they were aged 16 years or older at the time of care ending, and covers 1 April 2002 – 31 March 2016. There are approximately 7,150 records reflecting 7,080 individual children. This dataset is no longer updated.

The main purpose of this dataset was to record the qualifications that care leavers had gained at the point they ceased to be looked after. Variables include the number of GCSEs attained, by grade range; the number of General National Vocational Qualifications (GNVQs) attained; the number of other educational and vocational qualifications attained; and, if no qualifications were attained, the reason for this (such as illness or disability). Although this dataset is no longer updated, it is possible to retrieve information about the number of GCSEs attained by an individual and the grades they achieved in each subject. This is done by linking to the Pupil Level Annual School Census (PLASC) dataset, using the Anonymised Linking Field (ALF) or the Individual Record Number (IRN) from the primary CLA Census dataset.

Data Information

Further information on the data can be found here: [Dataset - Health Data Research Gateway](#)

Data Limitations

One significant limitation of the CLA Census dataset is the lack of routinely collected information on the individual-level socioeconomic circumstances of children and families requiring state intervention. Such information would provide a more holistic picture of the circumstances and backgrounds of the children who require care.

A further limitation is that the data do not contain information about children post-adoption.

When an adoption order is granted, a child's identifiers are changed. This includes their User Principal Name, unique ID within a local authority, and their NHS number. The impact of this change on the SAIL Anonymised Linkage Field (ALF) and the ability to link pre- and post-adoption records is currently unknown, but it is likely that it will not be possible to investigate incidence of return to care, or the educational outcomes of this group.

As with most administrative datasets there is some missing information relating to, for example, home and placement postcodes, primarily of those in older years. There is also some variation in recording between local authorities, and local authority child identifiers do not continue with the child if the responsibility for their care is transferred to another local authority. This can make it challenging to link records for individual children who have

this experience. Researchers can mitigate this by using the Individual Record Number and the SAIL ALF to identify children who have multiple local authority identifiers and reunify their care records. In addition, this means that it is necessary to use both the child and the local authority identifiers for each child in order to track individual children within the CLA data.

For the CLA return, local authorities are required to choose only one 'category of need' code from a list that best describes the main reason for the child receiving care at the time the period of being looked after began. This code remains the same for as long as the period of care continues. Therefore, it is not possible to examine whether children enter care for multiple reasons, or to determine whether their reasons for being in care change over time, unless they exit and re-enter the care system.

Researchers within SAIL Databank at Swansea University continue to complete data quality checks to aid further use of the data for all.

Research questions could focus on:

- Health and education outcomes, including the comparison of children in different service categories (child protection, looked after, others)
- Comparisons of outcomes for different durations and intensities of social care involvement (e.g. using placement data from the looked after children return)
- Care experiences
- The possible impact of policy changes (e.g. further exploring the interesting finding in this paper about GP contact)

Accessing the data

The data are accessed via the [SAIL Databank](#) based in Swansea University Wales.

The following video gives a full overview of the SAIL application process. [Accessing SAIL](#)

ADR Wales does not hold datasets. This data is held securely and anonymously within the SAIL Databank. Subject to safeguards and approvals, data held within the SAIL Databank can be linked together to address important research questions. A list of the anonymous data held within the SAIL Databank can be found here: [Dataset Catalogue](#)

Prior research: Publications Register - [SAIL Publication Register](#)

Data User Register (Active projects) - [Data use register - SAIL Databank](#)

Acknowledgements

We would like to thank Welsh Government and local authorities for collecting the Children Looked After data and allowing researchers to access it.

This research has been carried out as part of the ADR Wales programme of work, as part of the Social Justice thematic research area. The thematic research areas are aligned to the priority themes as identified in the Welsh Government's Programme for Government.

ADR Wales brings together data science experts at Swansea University Medical School, staff from the Wales Institute of Social and Economic Research and Data (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. ADR Wales is part of the Economic and Social Research Council (part of UK Research and Innovation)-funded Administrative Data Research UK (ADR UK).

This work is supported by ADR UK. ADR UK is a partnership transforming the way researchers access the UK's wealth of public sector data, to enable better informed policy decisions that improve people's lives.

References

- Allnatt G, Elliott M, Scourfield J, Lee A, Griffiths L. Use of linked administrative children's social care data for research: a scoping review of existing UK studies. *Br J Soc Work*. 2022; 52(7): 3923.
- Allnatt G., Lee, A., Scourfield, J., Elliott M., Broadhurst, K., Griffiths L (2022) Data Resource Profile, children looked after administrative records in Wales. *Int J Popul Data Sci*. 2022; 2;7(1):1752.
- Bywaters P, Scourfield J, Jones C, Sparks T, Elliott M, Hooper J, et al. Child welfare inequalities in the four nations of the UK. *J Soc Work*. 2020; 20(2):193.
- Elliott M. (2020) 'Child Welfare Inequalities in a Time of Rising Numbers of Children Entering Out-of-Home Care', *The British Journal of Social Work*, 50 (2), pp. 581–97
- Hodges H & Bristow D. Analysis of the factors contributing to the high rates of care in Wales. 2019. Wales Centre for Public Policy.
- Hodges H. Children looked after in Wales: Flows into and out of the care system. 2020a. Wales Centre for Public Policy.
- Hodges H. Children looked after in Wales: Trends. 2020b. Wales Centre for Public Policy.
- Jay MA, Woodman J, Broadhurst K, Gilbert R. Who cares for children? Population data for family justice research. 2017. Available from: <http://wp.lancs.ac.uk/observatory-scopingstudy/>
- Lee A, Elliott M, Scourfield J, Bedston S, Broadhurst K, Ford DV, Griffiths LJ. Data resource: Children receiving care and support and children in need, administrative records in Wales. *Int J Popul Data Sci*. 2022 May 9;7(1):1694.

Produced by ADR Wales

Website: ADRWales.org

Twitter: [@ADR_Wales](https://twitter.com/ADR_Wales)

ADR Wales brings together specialist teams, data science experts, and statisticians as part of the Economic and Social Research Council (part of UK Research and Innovation) funded ADR UK. Our team is made up of specialists in their field from Swansea University Medical School, the Wales Institute of Social and Economic Research, Data and Methods (WISERD) at Cardiff University and the SAIL Databank at Swansea University with statisticians, economists and social researchers from Welsh Government. Together ADR Wales develops new evidence which supports the Welsh Government's national strategy, Prosperity for All to improve the lives of people in Wales.

For further information please contact
Cathrine.E.Richards@Swansea.ac.uk



ADR Wales Partners

