Oxford Review of Education

ISSN: 0305-4985 (Print) 1465-3915 (Online) Journal homepage: www.tandfonline.com/journals/core20

Children's early care experiences and their educational attainment: a population data-linkage study in Wales

Emily Lowthian, Stuart Bedston, Alex Lee, Ashley Akbari, Lucy Griffiths, Tom Crick & Donald Forrester

To cite this article: Emily Lowthian, Stuart Bedston, Alex Lee, Ashley Akbari, Lucy Griffiths, Tom Crick & Donald Forrester (12 Sep 2025): Children's early care experiences and their educational attainment: a population data-linkage study in Wales, Oxford Review of Education, DOI: 10.1080/03054985.2025.2540036

To link to this article: https://doi.org/10.1080/03054985.2025.2540036

OXFORD REVIEW OF EDUCATION
https://doi.org/10.1080/03054985.2025.2540036

Routledge
Taylor & Francis Group

Children's early care experiences and their educational attainment: a population data-linkage study in Wales

Emily Lowthian [a], Stuart Bedston [b], Alex Lee [b], Ashley Akbari [b], Lucy Griffiths [b], Tom Crick [a] and Donald Forrester [c]

[a]Department of Education and Childhood Studies, School of Social Sciences, Swansea University, Swansea, UK; [b]Population Data Science, Data Science Building, Swansea University, Swansea, UK; [c]CASCADE, Cardiff University, Cardiff, UK

ABSTRACT

Care-experienced children are at risk of lower educational attainment. Duration of care is related to attainment, as is the type of placement(s) (e.g. foster care). To determine 'what works' for care-experienced children, our research examined: i) profiles of children on their care experiences, and ii) how these relate to educational attainment at age seven. Using anonymised, linked records in Wales, United Kingdom, we constructed an e-cohort of children born between September 2000 and August 2003. Data sources included births, children's social care, primary health care, demographics and education. We conducted latent class analysis using a three-step approach, summarising social care experiences with attainment at age seven in English/Welsh and mathematics as a distal outcome. Seven profiles best fitted the data, using data on placement types, duration and age on entry. For the first six years of life, those who experienced foster care which progressed to adoption showed the highest attainment (~1.00 masked), whereas those children who entered foster care from their fourth birthday had the lowest attainment intercept (0.40, 0.13–0.68). From this, we argue that stakeholders should develop additional support for children whose placement is largely foster care, as this group was most at risk of low attainment.

KEYWORDS

Children in care; education; social care; administrative data

Background

In the UK, children in care are those who have been provided with out-of-home accommodation by children's social care services for more than 24 hours. The most common reason for a child being in care is because they have experienced or are thought to be at risk of significant harm. The number of children who are in care remains a substantial social issue in Wales, which is a devolved area of policy since 1999 (Scourfield et al., 2008). Wales has a comparatively high rate of care-experienced children (1.1%) and the early years are said to be a prominent time for entry into care (Elliott, 2018). Research suggests

CONTACT Emily Lowthian e.m.lowthian@swansea.ac.uk Department of Education and Childhood Studies, School of Social Sciences, Swansea University, Swansea SA2 8PP, UK

Supplemental data for this article can be accessed online at https://doi.org/10.1080/03054985.2025.2540036

that care-experienced children are at greater risk of mental health issues, hospital admissions and lower educational attainment (Berridge et al., 2020; Forrester et al., 2009; Lowthian et al., 2022; McAuley & Davis, 2009; Sutcliffe et al., 2017). A UK Government report for England showed a 26-percentage point difference in care-experienced children meeting the expected level for writing and mathematics by age seven, compared to children who have not experienced care (Department for Education, 2020). The Education Endowment Foundation (2017) states that this 'attainment gap' begins early, is evident when children begin school at age five, and grows wider with each year.

Research has continued to identify factors which contribute to this attainment gap experienced by care-experienced children, potentially identifying areas for intervention. A systematic review by O'Higgins et al. (2017) identified that age of care entry, number of placements, placement duration, and placement type were related to educational attainment for children, along with individual characteristics such as gender and ethnicity. First, there was some consensus that the earlier a child entered care, the better the educational outcome (McClung & Gayle, 2010; Sebba et al., 2015), noting that earlier entry was often related to abuse or neglect and later (teenage) entry was related to behavioural problems and family breakdown (O'Higgins et al., 2017). Second, a higher number of placements was found to be related to lower attainment in several studies (Australian Institute of Health and Welfare [AIHW], 2011; Sebba et al., 2015). However, this finding was not consistent, which may be due to the complex nature of number of placements as a measure (O'Higgins et al., 2017). Third, children who experience more stability in their care placement(s) are at less risk of lower educational attainment (Goyette et al., 2021). This may be a causal factor, with higher level of placement stability allowing children to perform better educationally, however it is also possible that other factors, such as behavioural challenges, contribute to both reduced stability and poorer educational performance. For placement type, fewer differences are generally observed between kinship care (i.e. children who are placed with someone they have a kinship bond with, e.g. a grandparent) and foster care in terms of educational attainment (Winokur et al., 2014), although high instability is related to lower grades for children in foster care compared to kinship care (Sebba et al., 2015). Within the United States, Font (2014) found that children in kinship care had lower attainment over time compared to children in foster care. Lastly, adopted children tend to have higher attainment compared to children with other care experiences; however, generally it is still lower than for children with no care experience (Brown et al., 2017).

A methodological limitation of some of this previous work is that these factors are often considered individually or estimated separately in their statistical analyses. Yet, they occur together and interrelate, forming a complex sequence of events that ultimately shape children's opportunities and foundations for achieving their educational potential. This complexity has been explored using sequence analysis (McGrath-Lone et al., 2020), but this method cannot identify subgroups of experiences, nor has it been used in relation to understanding the impact on educational outcomes. To address this gap, our study incorporates the factors of age of care entry, number of placements, placement duration and placement type into a single analysis to determine specific care profiles. These care profiles were then used to estimate educational attainment at age seven while adjusting for key individual characteristics, as recommended by O'Higgins et al. (2017). This builds on research that has explored group and individual movements using sequence analysis

(McGrath-Lone et al., 2020), as our analysis enables the identification of data-driven care profiles which can be regressed on to educational outcomes adjusted for sociodemographic characteristics. This responds to existing limitations in the field, where care experience is identified by event flags, placement type, length or re-occurrence, either individually or in limited combinations, in relation to educational outcomes (Sebba et al., 2015; Tessier et al., 2018).

We used population-level administrative data linked across demographic, health, education and social service sectors to form an electronic cohort of children in Wales. With these data, we aimed to answer three main research questions to address existing gaps in the literature:

(1) How can the variables of placement type(s) and number, duration in care, and age first entered care be used to develop distinct groups of care experiences?
(2) What is the educational attainment of the care-experience groups, when adjusted for key covariates?
(3) Which demographic factors are associated with each care-experience group?

To address this gap, our study looked to use anonymised, linked, administrative records on the population of Wales to ascertain the impact of care experiences on educational attainment by age seven. Specifically, we aimed to answer the following three research questions:

(1) How can we best summarise and identify distinct groups of care experiences before the first major educational assessment?
(2) After considering key confounders, to what extent do these groups of care experiences meet the required attainment of English/Welsh and mathematics by age 7?
(3) What demographic characteristics are associated with the groups of care experiences?

Method

We built a retrospective, observation cohort study of children who were born in Wales between September 2000 and August 2003, using anonymised, individual-level, population-scale linked data sources. We built this electronic cohort using records on all births in Wales, UK, and if children had moved to Wales in their first year of life. The data were accessed via the Secure Anonymised Information Linkage (SAIL) Databank, the national Trusted Research Environment for Wales (Ford et al., 2009; Jones et al., 2014; Lyons et al., 2009; Rodgers et al., 2009, 2012). The administrative data sources included national birth records, population demographics, health records, education records and social service records; for additional information on these, see online Supplementary Table S1.

Similarly to Melis et al. (2023), due to the lower linkage rate for the data sources which hold the children in care data, we employed additional linkage processing by using the anonymised version of a child's Unique Pupil Number (UPN) in care data sources with the same week of birth and sex to improve linkage; see published paper on upgraded algorithm (Bailey et al., 2025). As births in Wales are recorded across three main data sources, we used these to create a single cohort of children who had care experience. Once we had developed this, we began to apply the criteria shown in Figure 1 to ensure

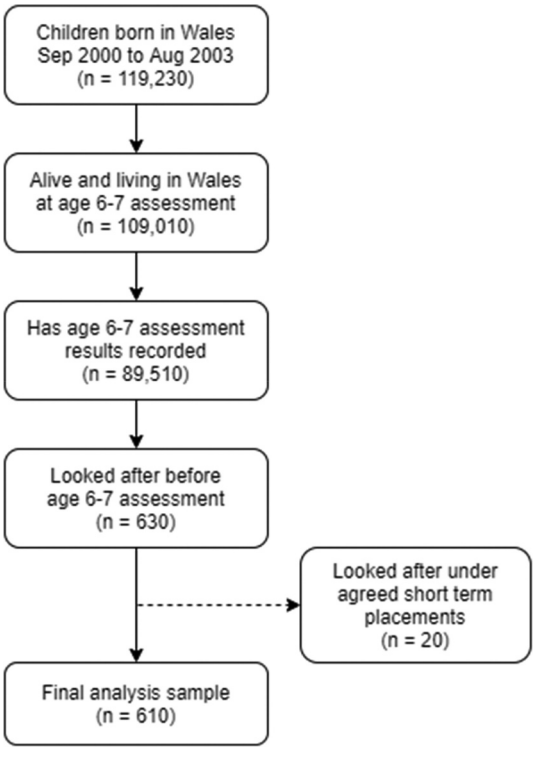

Figure 1. Consort diagram showing sample selection process.

the data integrity, and appropriate records were then selected for analysis. Our final sample consisted of 610 children who had full Welsh care history data that was able to be linked with education records.

Public patient involvement and engagement (PPIE)

Throughout our study we worked with PPIE groups who were related to our research and had the possibility of being affected by its outputs (Aiyegbusi et al., 2023). First, we worked with *CASCADE*[1] *Voices*, an advisory group of care-experienced young people who support research ideas and development. We worked with the young people in the first stage to explore the relevance of different placement types in the development of profiles shown in this paper. Young people discussed which placement types they thought were the most important to include for analysis; we included all of these in the analysis and only excluded if the sample was too small, e.g. residential care or mother-baby units. We also communicated our results to the young people, before any publication or dissemination of results, to understand the impact of our results. From this, we recognised that our measures of educational outcomes were crude and obscured wider educational outcomes felt and understood by care-experienced children. Our research is not intended to be stigmatising but rather to help improve identification of which groups need greater support to reach their potential. Second, before beginning the project, we worked with the SAIL

Consumer Panel group in order to understand wider public perceptions of our study's aims and impacts; from this, we gained insights on how to effectively communicate with stakeholders. Thirdly, we worked with primary school teachers to understand professional perspectives and understandings of care-experienced children in the initial stages of our work; much of this discussion provided real-world context to our initial research aims. Lastly, we worked with an advisory group compromising academics, government officials and social care associates to provide feedback on our research design, statistical analysis and dissemination strategies.

Ethics

Ethical approval was granted for this study on 18 November 2022 by Swansea University (SU-Ethics-Staff-181122/541). This included the use of data from SAIL (along with their own governance procedures), and working with the public patient involvement groups mentioned above.

Measures

The main exposure was children's social service information before their teacher assessment when aged 6–7 years old – termed Key Stage 1 (KS1) at the date of sample (National Assembly for Wales, 2007). From this, we derived their type of placement: foster care, kinship care, adoption, and residing with parents, which is defined as being 'placed with own parents or other person with parental responsibility' (Statistics for Wales, 2021, p. 9). Kinship care is defined as care provided by a relative or friend, whereas foster care is care provided by others including foster carers, agencies (private organisations) and parent-child foster placements; both are managed by the local authority (small governance areas in Wales). We also captured reasons why the placement ceased, including adoption and returning home; all placement variables were binary (0 and 1+) in order to achieve model convergence. We then derived the age of the child when they first entered care, along with the duration of time they spent in care in total. We excluded those with short-term breaks.

The outcome was educational attainment, KS1, at age 6–7 years. Children of this age in Wales are nationally expected to achieve a level 2 (Strand, 2002) in English or Welsh and mathematics; if they achieved this, they were considered as meeting national expectations. For context, it is identified that 83 per cent of children in the general population met this (*StatsWales: Results, by Local Authority and Level*, 2011).

The analysis was adjusted for key sociodemographic covariates, including sex (male being the reference class), ethnic group (the ethnic majority in Wales – White, was the reference class), mother's age (seven categories) and the Welsh Index of Multiple Deprivation at birth (in deciles, with increases meaning less deprivation) (Welsh Government, 2019a). We also used birth characteristics such as whether the baby was born pre-term (<36 weeks), of low birthweight (≤2500 g), and the year the child was born, termed as birth cohort (2000, 2001, 2002, 2003). To define learning difficulty, we used read codes from a variety of sources identified in published papers (Madley-Dowd et al., 2023; Welsh Government, 2019b).

Statistical analysis

Data management was conducted using SQL (via the IBM DB2 database in SAIL) and *R* (*R: A Language and Environment for Statistical Computing*, 2020), and we conducted latent class analysis with a distal outcome and covariates using the Manual 3-step approach (Nylund-Gibson et al., 2019) in *Mplus* (Muthén & Muthén, 2017). We registered our statistical analysis plan on the Open Science Framework first on 1 March 2023 and made revisions in January 2024 and again in July 2025 (Lowthian et al., 2023). First, we conducted the unconditional latent class analysis model using the exposure measures. When running the models, it was confirmed that the log likelihood replicated, with double and then triple the random starts for reliability and stability. Then, we assessed the model fit using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), where the lowest value indicates an improved model, Entropy (0–1.00), with larger values indicating less classification error (Collins & Lanza, 2009) alongside the Vuong-Lo-Mendell-Rubin likelihood ratio test (VLMRT), Lo-Mendell-Rubin adjusted test (LMRLRT) and Bootstrapped Likelihood Ratio Test (BLRT), which tests if the new model is a better fit, e.g. three classes compared to two (Weller et al., 2020). Classes were added until the model no longer converged or the log likelihood would no longer replicate, or the model fit criteria suggested extracting more classes was no longer efficient, which was balanced with theoretical interpretability (Melendez-Torres et al., 2018).

Once we had decided on the best unconditional model, we began the process of the three-step manual approach as suggested by Nylund-Gibson et al. (2019). We added in the covariates and educational attainment outcome as auxiliary variables and then used the classification probabilities logits as values in the third step. We regressed the covariates onto the classes and onto the educational outcomes. Each class had an estimated intercept (the mean) for educational attainment, along with odd ratios (ORs) for how covariates predicted the class. Lastly, covariates were used to predict educational attainment, with the classes not allowed to vary freely.

Results

Our sample consisted of 610 looked after children, for which 380 (62.3%) had passed the Key Stage 1 assessment (Table 1). Most were born in the year 2002 ($n = 220$), with fewer children born in the year 2000 ($n = 60$) (due to the e-cohort beginning on 1 September 2000). Nearly a quarter of children were born to mothers under 19 years of age (22.2%), with most being born to mothers aged 20–24 years (31.7%). In terms of perinatal characteristics, 14.8 per cent of the children were born pre-term, 16.4 per cent were born with a low birthweight, and there were fewer female than male children (42.6%). A third of children were living in areas classified as being in the most deprived decile (31.7%), with the vast majority of children living in the fifth deprived decile or greater (79.3%). Few children in our sample were of an ethnic minority group (4.9%; Asian, Black, Mixed or Other groups). Furthermore, there was an over-representation of children who had been diagnosed with a learning difficulty in our sample (3.3%). Most children had experienced a foster care placement (80.6%), around a quarter had experienced kinship care (26.2%), a quarter had been placed with parents (24.2%), and one in twenty had been placed for adoption (4.9%). In terms of care outcomes, over a third (34.4%)

Table 1. Descriptive statistics; counts rounded to the nearest 10 to meet disclosure requirements.

Variable	n (%)
Total	610 (100.0%)
Birth year	
2000	60 (9.8%)
2001	170 (27.9%)
2002	220 (36.1%)
2003	160 (26.2%)
Pre-term birth	
No	520 (85.2%)
Yes	90 (14.8%)
Low birthweight	
No	510 (83.6%)
Yes	100 (16.4%)
Welsh Index Multiple Deprivation	
1 – Most deprived	200 (31.7%)
2	110 (17.5%)
3	90 (14.3%)
4	50 (7.9%)
5	50 (7.9%)
6	30 (4.8%)
7	30 (4.8%)
8	30 (4.8%)
9	20 (3.2%)
10 – Least deprived	20 (3.2%)
Sex	
Male	350 (57.4%)
Female	260 (42.6%)
Ethnic group	
White	580 (95.1%)
Ethnic Minorities	30 (4.9%)
Learning difficulty	
No	590 (96.7%)
Yes	20 (3.3%)
Mother's age at birth	
<19 years	140 (22.2%)
20–24 years	200 (31.7%)
25–29 years	120 (19.0%)
30–34 years	80 (12.7%)
35–39 years	50 (7.9%)
>40 years	20 (3.2%)
(Missing)	20 (3.2%)
Age first looked after	
Mean (SD)	2.9 (1.94)
Duration looked after (years)	
Mean (SD)	2.2 (1.90)
History of placements	
Adoption	30 (4.9%)
Foster care	490 (80.6%)
Kinship care	160 (26.2%)
Parents	150 (24.2%)
History of care outcomes	
Returned home	210 (34.4%)
Adopted	20 (3.3%)
Passed Key Stage 1 assessment	
No	230 (37.7%)
Yes	380 (62.3%)

experienced returning home, and a small proportion were adopted (3.3%). Children, on average, first went into care at nearly three years old (mean 2.9, SD 1.94) and their total duration in care was 2.2 years, on average (SD 1.90).

Latent class analysis

We conducted latent class analysis on the placement type, age first entered care, and duration of being in care. We attempted to use continuous counts of foster care and kinship care where possible, but the model would not converge, so binary variables were used; we also excluded the category of 'other' as it made no contribution to the model. We decided that the seven-class solution was the best solution. While it had a lower entropy than others (0.90), it had a lower BIC, and both the LRT and the BLRT suggested it was an improved fit compared to six classes. The eight-class solution was unable to replicate the best log likelihood, but it did have a lower AIC and BIC value; however, the LRT suggested that eight classes were not improved compared to seven. Classes two to six were not chosen as they did not have the lowest AIC/BIC and the (B)LRTs suggested more classes were improved to the previous (*k*-1); this is despite entropy being the highest with five classes (0.92), but not that much different to seven classes (0.90). See Table 2 for estimates in full.

The seven classes produced class probabilities to determine which placements had a high probability in each class, along with an average age of first placement and an average duration of being in care for each class. Table 3 shows the class probabilites, the average age first in care and the average duration.[2]

Class 1 was defined as ***Foster care to adoption***. This class identified children who had entered into care at around one year of age, were fostered and then continued

Table 2. Model fit criteria for latent class analysis.

Classes	AIC	BIC	Entropy	VLMR-LRT	LMR-LRT	BLRT
2	7721.35	7805.24	0.89	$p < 0.05$	$p < 0.05$	$p < 0.05$
3	7445.38	7569.01	0.88	$p < 0.05$	$p < 0.05$	$p < 0.05$
4	7191.85	7355.21	0.90	$p < 0.05$	$p < 0.05$	$p < 0.05$
5	7015.41	7218.51	0.92	$p < 0.05$	$p < 0.05$	$p < 0.05$
6	6929.42	7172.25	0.91	$p < 0.05$	$p < 0.05$	$p < 0.05$
7	**6859.13**	**7141.70**	**0.90**	$\mathbf{p < 0.05}$	$\mathbf{p < 0.05}$	$\mathbf{p < 0.05}$
8	6788.40	7110.70	0.90	$p = 0.05$	$p = 0.05$	*NA*

Table 3. Probabilities of being in each latent class by variable.

	Class 1 (3%)	Class 2 (8%)	Class 3 (19%)	Class 4 (21%)	Class 5 (17%)	Class 6 (17%)	Class 7 (14%)
Age first in care	1.05	5.01	0.74	5.22	1.41	2.32	3.94
Duration	2.43	0.85	5.26	0.44	0.60	3.41	1.98
Foster care	~1.00	<0.20	0.77	1.00	~1.00	0.69	~1.00
Kinship care	~0.00	0.77	0.36	~0.00	0.21	0.43	0.12
Parents	~0.00	0.21	0.47	~0.00	0.11	0.45	0.20
Adoption	~1.00	0.00	<0.10	0.00	0.00	<0.10	~0.00
Returned home	~0.00	<0.20	0.15	0.37	~1.00	0.26	0.16
Ceased to be adopted	~1.00	0.00	<0.10	0.00	0.00	0.00	0.00

to be in care for nearly two and a half years until care ceased due to adoption. Over this time, it was unlikely that children returned home or stayed with parents; most children had an adoption placement (~1.00). This class accounts for 3 per cent of the sample ($n = 20$).

Class 2 was defined as ***Kinship care entrants.*** This class identified children who had entered care at around five years of age and were largely in care for just under a year (0.85). Most of the children had a kinship care placement (0.77), around a fifth resided with parents (0.21), and fewer had the probability of being in foster care or returning home. This class accounts for 8 per cent of the sample ($n = 52$).

Class 3 was defined as ***Infant-entry mixed-placement***. This class identified children who had entered care before their first year of life and were in care the longest compared to other classes (5.26). Most children had a foster care placement (0.77) and the likelihood was relatively high for a placement with parents (0.47) or kinship care (0.36). Some children also returned home (0.15). This class accounts for 19 per cent of the sample ($n = 119$).

Class 4 was defined as ***After age 5 foster care***. This class identified children who had entered care late, after their fifth birthday, and therefore had the shortest duration of care (0.44); however, the length of care could be artificially shortened in this class as the exam-period date ended the cohort exposure time. All children had entered foster care (1.00), with few experiencing other types of care. Some children returned home (0.37). This class accounts for 21 per cent of the sample ($n = 126$).

Class 5 was defined as ***Early fostering to reunification***. This class identified children who had entered care around one year and five months of age and stayed in care for around seven months. Most children had been placed in foster care (~1.00), and some had been placed in kinship care (0.21) or with parents (0.11). Children in this class had a high probability of returning home (~1.00) compared to any other class. This class accounts for 17 per cent of the sample ($n = 104$).

Class 6 was defined as ***Toddler-entry mixed care***. This class identified children who had entered care around two years and four months of age and stayed in care for around three years and five months. Children had a high probability of foster care (0.69) and somewhat for kinship care (0.43) or residing with parents (0.45). Children had a lower likelihood of returning home with parents (0.26) or being adopted (0.00). This class accounts for 17 per cent of the sample ($n = 105$).

Class 7 was defined as ***Pre-school foster care***. Children in this class had entered care at almost four years of age, with an average duration of almost two years. Children in this class had a high probability of foster care (~1.00) and had a lower likelihood of kinship care (0.12) or residing with parents (0.20). Some children returned home (0.16). This class accounts for 14 per cent of the sample ($n = 85$).

Covariates predicting latent classes

Odd ratios were interpreted to understand the likelihood increase or decrease for the covariates predicting each class as an outcome; *Foster care to adoption* was the reference class. Most covariates did not have a strong association with being in each class compared to the reference class. However, a unit increase in the WIMD decile (i.e. being more affluent) decreased the likelihood of being in the *Early fostering to reunification* class by

0.80 (CI 95%, 0.67–0.96). In addition, a unit increase in the WIMD decile decreased the likelihood of being in the *Pre-school foster care* class by 0.82 (CI 95%, 0.68–1.00). See Supplementary Table S2 for full estimates.

Educational attainment: predictors and outcomes

In the third step of the model, class changes can occur which can be problematic for interpretation and stability. However, in our model, changes were within a 1 per cent difference and the entropy lowered to a still acceptable value of 0.83. Table 4 shows the intercepts of KS1 for each class with statistical significance and 95% confidence intervals. Across these intercepts, the *Foster care to adoption* class had the highest KS1 intercept (~1.00), followed by *Toddler-entry mixed care* (0.66), then *Early fostering to reunification* (0.65). Both *Kinship care entrants* and *Infant-entry mixed-placement* had the same intercept (0.63), which was then followed by *After age 5 foster care* (0.58). The lowest intercept was 0.40, which was attained by the *Pre-school foster care* class.

Table 5 shows factors predicting KS1 results (age 6/7 educational attainment) in the context of the latent classes, however the number in the sample is smaller ($n = 596$) due to missing data in the covariates. Female children had a higher chance of attaining the nationally expected level of education (0.14, 0.06–0.21). Those with low birthweight had a lower chance of attaining the required education level (−0.19, −0.33–−0.06). All other covariates showed wider confidence intervals and were not statistically significant; we were unable to include learning difficulty in this model due to convergence issues.

Pairwise difference tests were conducted to compare statistically significant differences across the classes. Table 6 shows the comparison of the KS1 intercepts of the classes with the estimate and 95% confidence intervals in brackets. Notably, those who were in the *Foster care to adoption* class had statistically significantly higher KS1 intercepts than all

Table 4. KS1 intercepts for educational attainment by latent classes with *p*-value and 95% confidence intervals.

	KS1 intercept	*p*-value	Lower CI	Upper CI
Foster care to adoption	~1.00	<0.05	–	–
Toddler-entry mixed care	0.66	<0.05	0.39	0.93
Early fostering to reunification	0.65	<0.05	0.37	0.93
Kinship care entrants	0.63	<0.05	0.34	0.92
Infant-entry mixed-placement	0.63	<0.05	0.36	0.92
After age 5 foster care	0.58	<0.05	0.30	0.86
Pre-school foster care	0.40	<0.05	0.13	0.68

Table 5. Regression analysis showing covariates predicting KS1 across latent classes.

	Coefficient	*p*-value	Lower CI	Upper CI
Female = yes	0.14	<0.05	0.06	0.21
Marginalised ethnic group	0.07	0.43	−0.10	0.24
Mother's age	−0.02	0.14	−0.05	0.01
WIMD	0.01	0.37	−0.01	0.02
Pre-term	0.06	0.41	−0.08	0.19
Low birthweight	−0.19	<0.05	−0.33	−0.06
Birth year	−0.03	0.12	−0.07	0.01

Table 6. Pairwise comparison of KS1 intercepts across latent classes, with *p*-values in bold and 95% confidence intervals.

	Foster care to adoption	Kinship care entrants	Infant-entry mixed-placement	After age 5 foster care	Early fostering to reunification	Toddler-entry mixed-placement
Kinship care entrants	$p < 0.05$					
Infant-entry mixed-placement	$p < 0.05$	0.00 (−0.18–0.24)				
After age 5 foster care	$p < 0.05$	0.05 (−0.13–0.17)	0.05 (−0.08–0.12)			
Early fostering to reunification	$p < 0.05$	−0.02 (−0.20–0.15)	−0.02 (−0.15–0.12)	−0.07 (−0.21–0.06)		
Toddler-entry mixed care	$p < 0.05$	−0.03 (−0.22–0.43)	−0.03 (−0.19–0.39)	−0.08 (−0.23–0.35)	−0.02 (−0.17–0.42)	
Pre-school foster care	$p < 0.05$	**0.23** (0.03–0.19)	**0.23** (0.07–0.08)	**0.18** (0.00–0.14)	**0.25** (0.08–0.44)	**0.26** (0.08–0.44)

other classes. In reverse, all classes were statistically significantly higher than the *Pre-school foster care* class. No other classes had statistically significant differences, which is unsurprising given their similar intercepts.

Discussion

Our results suggest that several care-related factors contribute to educational attainment in the early years of childhood. Using information on care placement type, duration in care, and age first entering care, we were able to distinguish seven latent classes. These classes were: *Foster care to adoption*; *Toddler-entry mixed care*; *Early fostering to reunification*; *Kinship care entrants*; *Infant-entry mixed-placement*; *After age 5 foster care*; and *Pre-school foster care*. All classes were distinct and offered insights into how children have experienced the care system. We found overall that the *Foster care to adoption* class showed the highest proportion meeting expected educational levels, whereas the *Pre-school foster care* class showed the lowest (40%). We consider our findings in light of other studies and theory to develop the meaningfulness of these results in relation to knowledge, policy and practice.

First, one of our main findings included the *Foster care to adoption* class having the highest attainment outcome on average compared to other classes. This is in line with the wider literature that suggests children who are adopted often have a higher chance of attaining the nationally expected grades compared to children in other placement types (Brown et al., 2017; Sutcliffe et al., 2017). Welsh statistics suggest that 83 per cent meet the expected level of education (*StatsWales: Results, by Local Authority and Level*, 2011) and administrative studies using sub-samples suggest anywhere between 81.9–92.8 per cent (Evans et al., 2019, 2020; Rahman et al., 2018). Hence, we find that the adoption group are within nationally and administratively reported levels at age seven years. Further evidence suggests that the move from an often economically or socially deprived environment to a richer, more psychologically and emotionally adjusted environment often contributes to positive effects on adopted children's development (Brown et al., 2017; Collishaw et al., 1998; Van IJzendoorn et al., 2005). We theorise that, in addition to providing positive care, adoptive parents are more likely to be from more affluent positions (Maughan et al., 1998),

with a household richer in cultural capital (Lee & Bowen, 2006). Hence, adoptive families are more likely to possess the values, behaviours and economic decisions (e.g. reading at home, visiting educational sites), which are valued by educational systems and support higher educational performance (Lee & Bowen, 2006). Furthermore, children who are adopted are often placed earlier in life (Sebba et al., 2015), so they may have a lower risk of longer-term abuse and neglect, and fewer placement moves, leading to fewer developmental risks.

The other placement groups identified showed a higher proportion of children not meeting the nationally expected levels; on average those children attained 15–41 percentage points lower than the national average of 83 per cent noted previously. Sutcliffe et al. (2017) used English data in 2010 and found that 51 per cent and 61 per cent of looked-after children achieved the nationally expected levels of literacy and numeracy, respectively, at age seven years (KS1); this compared to 85 per cent and 88 per cent of all children in England, respectively. The classes *Toddler-entry mixed care*, *Early fostering to reunification*, *Kinship care entrants* and *Infant-entry mixed-placement* saw that between 63 and 66 per cent of children met the nationally expected levels. All these classes share a commonality of having some connection with their birth family and relations, irrespective of age on entry and duration. All groups (except *Kinship care entrants*) experienced early care entry (between one and two years of age), which is often identified as positive for educational attainment (Sebba et al., 2015); likewise those with mixed placements had longer stays in care, which is also identified as protective (Sebba et al., 2015). We found that the *Early fostering to reunification* profile was a high-achieving group compared to others (65 per cent on average), so, despite the short-term care stays which can be associated with risk, the placements are largely in foster care, and all children then return home, suggesting stability, as discussed in Sebba et al. (2015).

In terms of the *Kinship care entrants* profile, research shows that children who experience kinship care (which can include other parents) report more stable relationships and unconditional care (Rock et al., 2015), maintaining cultural and familial ties, thus retaining the established bond with the child (Font, 2014). However, wider literature remains unclear on kinship care's relationship with educational attainment, with a systematic review finding no difference to foster care (Winokur et al., 2014). Moreover, Font (2014) found that children in kinship care had higher performance than those in foster care at the baseline, but they declined over time, particularly those who were lower functioning at the baseline. This could be explained by kinship placements being less academically enriching settings, with Font (2014) theorising that kinship carers may have on average lower income and educational levels paired with a likely single-carer environment. Likewise, research in the UK shows that children who lived in the poorest 20 per cent of neighbourhoods were 2.25 times more likely to be in kinship care, when comparing to the richest 20 per cent of neighbourhoods (Nandy & Selwyn, 2013). While income levels are not directly comparable to educational attainment, being economically advantaged means being able to afford basics needs and material resources which could be educationally enriching.

Lastly, children in the *After age 5 foster care* and *Pre-school foster care* classes showed lower educational attainment on average: 58 per cent and 40 per cent, respectively. The *After age 5 foster care* class differed by its short duration of care (<6 months) and higher

likelihood of returning home. The likelihood of returning home suggests that the birth family have engaged with social workers and made necessary changes towards the home being a safe environment (Cheng, 2010; Kimberlin et al., 2009). Despite the necessary changes being made in the family unit, this group attained 25 percentage points lower than the national average of 83 per cent, suggesting that care, and likely pre-care factors, are associated with attainment at an early age. The *Pre-school foster care* class attainment was significantly lower (40%) than all other classes, suggesting that this group was the most vulnerable for low attainment. This class had a lower likelihood of residing with parents or kinship relatives, but a very high chance of foster care for around two years. In addition, this class was less likely to be born in an affluent decile, so poverty could have been an additional, complexly related factor for both not returning home and lower attainment.

Reflecting on wider research allows for the theorising of explanations for why these groups both had foster care as their dominant placement type, and low attainment; the environment in which the children reside is our first consideration. Foster care entails several factors such as the presence of other children, personal attributes of the carer, the carer–child relationship, and motivations to provide a home (Rock et al., 2015). A systematic review found that stability, confidence and encouragement were important factors for foster carers to have to support children's educational attainment (Rock et al., 2015; Sinclair et al., 2005). Interestingly, research by Sebba et al. (2015) in the UK did not point to foster carers having lower resources – only in kinship care was this the case. Alongside the environment, it is essential to consider why children have remained in foster care without parental or kinship care. It is likely that the family and extended support were considered not to be safe, which could be related to abuse, neglect and other complex factors including deprivation, which can have a cumulative effect on educational attainment (Mills, 2004, cited in Berridge, 2007). From this, we urge that further research should be considered on children in foster care, specifically in relation to their early educational outcomes and experiences.

This study benefits from high-quality population-level data on a rare sub-sample of the Welsh population, with coverage ranging from birth up to age seven, containing a wide range of information on health, education, and social care outcomes. However, we recognise that, due to this sample being small, which is to be expected due to the nature of the study's focus, some elements of the statistical analysis are under-powered. For example, only 30 children were adopted, and deriving population estimates using this data is challenging; this is due to our research having access to only a small proportion of the adopted children available due to a lack of linkage, and we were unable to use more recent data, which is better populated, due to poor records of education data during the Covid-19 period. Moreover, we recognise that numerous factors outside our chosen covariates contribute to educational outcomes, such as familial involvement and engagement, school attendance, special educational needs and pre-care experiences. Related to this, the read codes used to identify learning difficulties may not capture the full prevalence over the life course (Madley-Dowd et al., 2023; Welsh Government, 2019b).

In terms of care-related factors, we recognise that being a child 'in need' can involve being on the child protection register, or having experienced abuse, neglect, substance misuse, or disability, all of which are factors that contribute to attainment. We note that, in

our models, we were unable to ascertain clear re-entry when it was the same placement, e.g. entering foster care twice, due to a lack of model convergence outside of binary flags for placement (0 vs. 1+). Likewise, latent class analysis is a data-driven technique and, while balanced with theoretical interpretation, it is unlikely these groups would be closely replicated in other data, which limits generalisability. Further research needs to incorporate the complexity of why, when, where and for how long children are in care, adjusted for numerous demographic and school-related factors, to fully understand educational attainment. As imagined, this is a challenging area to develop research when faced with small sample sizes and data quality or linkage issues.

Overall, we find that seven distinct groups of care experience have differing chances of meeting the nationally expected attainment at age seven. In the Raising the Ambitions strategy for looked-after children in Wales, there was a focus on the use of data to understand educational attainment and how it could be used to guide good practice (Welsh Government, 2016). Our study addresses these gaps, building the evidence on care-experienced children's educational attainment, which is a continuing policy-focus. We find that the length of time in care and the age of entry into care are not the only variables associated with educational attainment, but also the placement type, as children placed in foster care later in childhood were at the greatest risk of lower attainment, specifically if their entry into care was just before school. We find that children placed with alternative members of their immediate family (e.g. grandparents, a parent not living in the main household, siblings, etc.) can be a protective factor for early years attainment – even when there are timing and length differences. Despite prior research suggesting that early, stable care has the most beneficial impact, we see that children in mixed placements are more likely to achieve the expected attainment levels compared to children who experienced only foster care for varying lengths of time. The question still exists as to whether educational attainment is related to the care placement or the childhood experiences before care occurred (O'Higgins et al., 2015). To add to this, educational experiences also contribute to attainment, with care-experienced children being specifically at risk of special educational needs, as discussed in O'Higgins et al. (2017), and, while not included in our study, we recognise the key contribution this has towards national measures of educational attainment.

From this, our study urges that further research must consider the complexity of care, including duration, age first entered, placement type, and pre-care experiences, to extend our findings and knowledge. It is clear that national cohort-scale administrative datasets have a key role in knowledge generation across the research–policy–practice interface. We recommend that research continues to investigate the complex and multi-faceted factors associated with educational under-performance, paying specific attention to genetic and prenatal factors, along with pre-intervention experiences such as neglect or victimisation where possible. Developing a clearer understanding of the pathways towards educational under-performance in care-experienced children will create avenues for improved support, including understanding the protective factors, the development of interventions, and practitioner guidance at multiple-levels (e.g. social care, education, health), to ensure a whole-service awareness.

Note

1. Children's Social Care Research and Development Centre.
2. Areas of this analysis had to be masked for statistical disclosure reasons, despite being a probabilistic method, '~' and '<' or '>' represent masked estimates and are indicative of the real estimate.

Acknowledgements

We thank the members of our Advisory Group who were involved in shaping this study's initial objectives and findings. We also thank the primary school teachers and social care professionals involved in this research, who provided their professional experience and opinions on this work. In terms of our public patient involvement and engagement, we firstly thank the young people at *CASCADE Voices*, whose input on this study in its initial stages were immensely helpful in shaping how we understood the educational system for care-experienced children; we are also grateful for their openness and critical mindset to our work, and we have tried to embed this in our communications to researchers and stakeholders. Lastly, we thank the SAIL IGRP (Information Governance Review Panel) Consumer Panel who provided helpful insights on our project in its initial stages, particularly in terms of dissemination.

This study makes use of anonymised data held in the Secure Anonymised Information Linkage (SAIL) Databank. We would like to acknowledge all the data providers who make anonymised data available for research.

Disclosure statement

No authors have any conflicts of interest to declare.

Funding

This research was supported by Health and Care Research Wales [SCG-21-1861], a networked organisation, supported by the Welsh Government, which brings together a wide range of partners across the NHS in Wales, local authorities, universities, research institutions, third sector and others. This research was supported by Administrative Data Research (ADR) Centre Wales. 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 [ES/W012227/1] is part of the Economic and Social Research Council (part of UK Research and Innovation)-funded ADR UK programme.

Notes on contributors

Emily Lowthian is a Lecturer in the Department of Education and Childhood at Swansea University. Her research focuses on understanding the impact of societal layers on child and adolescent health and wellbeing. Emily has published work examining the inequalities experienced by young individuals, including socioeconomic status, disability and gender disparities, and how these affect their mental health, educational outcomes, and health behaviours.

Stuart Bedston is a Research Officer and Data Scientist and specialises in techniques such as latent class analysis, survival analysis and epidemiology. His skills involve the management and preparation of big data to understand complex societal issues in the fields of health, education and social care. He has published an array of articles which explore the outcomes of families involved in the social justice system using population-level administrative data.

Alex Lee is the Head of Innovation in the SAIL databank at Swansea University within the Data Science Building. Alex has a comprehensive and detailed understanding of the Looked after Children dataset which led to the innovative boosting techniques used within the Looked after Children dataset. From this, Alex now provides data expertise and guidance for CASCADE and Welsh Government.

Ashley Akbari is a Senior Research Manager and Data Scientist at Swansea University and has experience of working on a variety of data linkage and research projects with a background in computing and various research disciplines which have evolved over time with his experience of working with anonymised data within SAIL Databank since 2008, as well as other data from the UK nations, and as part of projects with the NHS, government, policy and others.

Lucy Griffiths is a Professor of Paediatric Epidemiology and is Co-lead for the Environment and Health Research Centre (ENVHE) within Population Data Science at Swansea University. She currently works with the Nuffield Family Justice Observatory; Children's Social Care Research and Development Centre; Health Data Research UK Wales and Northern Ireland; Administrative Data Research (ADR) Wales and ADR England. Lucy's research largely uses survey and administrative information collected about children and young people, to study a range of individual, social and environmental circumstances that may affect health and health behaviours in these populations.

Tom Crick is Professor of Digital & Policy and Deputy Pro-Vice Chancellor (Civic Mission) at Swansea University. He holds a joint research and innovation appointment split between the Department of Education & Childhood Studies (Faculty of Humanities & Social Sciences) and the £32 m Computational Foundry (Department of Computer Science, Faculty of Science & Engineering). Whilst his disciplinary background is in computer science, his academic interests sit at the research/public policy interface with a focus on broad societal/policy impact: STEM/digital education, curriculum reform, science and innovation policy, data science, intelligent systems, smart cities, national infrastructure, software sustainability, digital transformation and skills/infrastructure for the digital/data economy.

Donald Forrester is the director of CASCADE, and has led projects and programmes worth over £20 million, including setting up the Tilda Goldberg Centre, academic lead for the Frontline social work programme, founding the What Works Centre for Children's Social Care and leading CASCADE Centre for Children's Social Care Research and Development. He is an ex-social worker and an expert in understanding outcomes and effective practice for children in receipt of social service intervention.

ORCID

Emily Lowthian http://orcid.org/0000-0001-9362-0046
Stuart Bedston http://orcid.org/0000-0002-5635-5957
Alex Lee http://orcid.org/0000-0001-7916-024X
Ashley Akbari http://orcid.org/0000-0003-0814-0801
Lucy Griffiths http://orcid.org/0000-0001-9230-624X
Tom Crick http://orcid.org/0000-0001-5196-9389
Donald Forrester http://orcid.org/0000-0002-2293-5718

Declaration of generative AI and AI-assisted technologies

During the preparation of this work, the first author used ChatGPT in order to support the naming of the latent classes available, i.e. *Infant-entry mixed-placement*; it was specifically used to help shorten the length of the names so they were better consolidated for readers. After using ChatGPT, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Data availability statement

This study makes use of anonymised data held in the Secure Anonymised Information Linkage (SAIL) Databank. For more information on accessing anonymised data see: https://saildatabank.com/.

References

Aiyegbusi, O. L., McMullan, C., Hughes, S. E., Turner, G. M., Subramanian, A., Hotham, R., Davies, E. H., Frost, C., Alder, Y., Agyen, L., Buckland, L., Camaradou, J., Chong, A., Jeyes, F., Kumar, S., Matthews, K. L., Moore, P., Ormerod, J., Price, G., … Calvert, M. J. (2023). Considerations for patient and public involvement and engagement in health research. *Nature Medicine*, *29*(8), 1922–1929. https://doi.org/10.1038/s41591-023-02445-x

Australian Institute of Health and Welfare. (2011). *Educational outcomes of children under guardianship or custody orders: A pilot study, stage 2*.

Bailey, G. A., Lee, A., Ahmed, S., Scanlon, I., Cowley, L. E., Stuart, A., Farr, I., Brooks, C., North, L., & Griffiths, L. J. (2025). Improving opportunities for data linkage within children looked after administrative records in Wales. *International Journal of Population Data Science*, *10*(1).

Berridge, D. (2007). And explanation in child welfare: Education and looked-after children. *Child & Family Social Work*, *12*(1), 1–10. https://doi.org/10.1111/j.1365-2206.2006.00446.x

Berridge, D., Luke, N., Sebba, J., Strand, S., Cartwright, M., Staples, E., McGrath-Lone, L., Ward, J., & O'Higgins, A. (2020). *Children in need and children in care: Educational attainment and progress* (Vol. 221). Nuffield Foundation.

Brown, A., Waters, C. S., & Shelton, K. H. (2017). A systematic review of the school performance and behavioural and emotional adjustments of children adopted from care. *Adoption & Fostering*, *41*(4), 346–368. https://doi.org/10.1177/0308575917731064

Cheng, T. C. (2010). Factors associated with reunification: A longitudinal analysis of long-term foster care. *Children and Youth Services Review*, *32*(10), 1311–1316. https://doi.org/10.1016/j.childyouth.2010.04.023

Collins, L. M., & Lanza, S. T. (2009). Latent class and latent transition analysis: With applications in the social. In *Behavioral, and health sciences*. Wiley.

Collishaw, S., Maughan, B., & Pickles, A. (1998). Infant adoption: Psychosocial outcomes in adulthood. *Social Psychiatry and Psychiatric Epidemiology*, *33*(2), 57–65. https://doi.org/10.1007/s001270050023

Department for Education. (2020). *Outcomes for children looked after by local authorities in England* (24), March 31, 2019.

Education Endowment Foundation. (2017). *The attainment gap*. https://educationendowmentfoundation.org.uk/support-for-schools/bitesize-support/closing-the-attainment-gap

Elliott, D. M. (2018). *The backgrounds of children entering public care in Wales*. Cardiff University. https://socialcare.wales/cms-assets/documents/Martin-Elliott-PhD-Edrych-ar-ôl-plant-yng-Nghymru-Dadansoddiad-o-gefndiroedd-plant-mynd-i-mewn-i-ofal-cyhoeddus.pdf

Evans, A., Dunstan, F., Fone, D. L., Bandyopadhyay, A., Schofield, B., Demmler, J. C., Rahman, M. A., Lyons, R. A., Paranjothy, S., & Ryckman, K. K. (2019). The role of health and social factors in education outcome: A record-linked electronic birth cohort analysis. *Public Library of Science ONE*, *14*(8), e0220771. https://doi.org/10.1371/journal.pone.0220771

Evans, A., Hardcastle, K., Bandyopadhyay, A., Farewell, D., John, A., Lyons, R. A., Long, S., Bellis, M. A., & Paranjothy, S. (2020). Adverse childhood experiences during childhood and academic attainment at age 7 and 11 years: An electronic birth cohort study. *Public Health*, *189*, 37–47. https://doi.org/10.1016/j.puhe.2020.08.027

Font, S. A. (2014). Kinship and nonrelative foster care: The effect of placement type on child well-being. *Child Development*, *85*(5), 2074–2090. https://doi.org/10.1111/cdev.12241

Ford, D. V., Jones, K. H., Verplancke, J.-P., Lyons, R. A., John, G., Brown, G., Brooks, C. J., Thompson, S., Bodger, O., Couch, T., & Leake, K. (2009). The SAIL databank: Building a national architecture for

e-health research and evaluation. *Biomed Central Health Services Research*, *9*(1), 157. https://doi.org/10.1186/1472-6963-9-157

Forrester, D., Goodman, K., Cocker, C., Binnie, C., & Jensch, G. (2009). What is the impact of public care on children's welfare? A review of research findings from England and Wales and their policy implications. *Journal of Social Policy*, *38*(3), 439–456. https://doi.org/10.1017/S0047279409003110

Goyette, M., Blanchet, A., Esposito, T., & Delaye, A. (2021). The role of placement instability on employment and educational outcomes among adolescents leaving care. *Children and Youth Services Review*, *131*, 106264. https://doi.org/10.1016/j.childyouth.2021.106264

Jones, K. H., Ford, D. V., Jones, C., Dsilva, R., Thompson, S., Brooks, C. J., Heaven, M. L., Thayer, D. S., McNerney, C. L., & Lyons, R. A. (2014). A case study of the secure anonymous information linkage (SAIL) gateway: A privacy-protecting remote access system for health-related research and evaluation. *Journal of Biomedical Informatics*, *50*, 196–204. https://doi.org/10.1016/j.jbi.2014.01.003

Kimberlin, S. E., Anthony, E. K., & Austin, M. J. (2009). Re-entering foster care: Trends, evidence, and implications. *Children and Youth Services Review*, *31*(4), 471–481. https://doi.org/10.1016/j.childyouth.2008.10.003

Lee, J.-S., & Bowen, N. K. (2006). Parent involvement, cultural capital, and the achievement gap among elementary school children. *American Educational Research Journal*, *43*(2), 193–218. https://doi.org/10.3102/00028312043002193

Lowthian, E., Bedston, S., Lee, A., Akbari, A., Forrester, D., Griffiths, L., & Crick, T. (2023). *Educational pathways and outcomes for children who are looked after: A population-scale data linkage study*. Open Science Framework. https://doi.org/10.17605/OSF.IO/ATK2H

Lowthian, E., Moore, G., Evans, A., Anthony, R., Rahman, M. A., Daniel, R., Brophy, S., Scourfield, J., Taylor, C., Paranjothy, S., & Long, S. (2022). *Associations between receipt of social services interventions*. *Educational outcomes and emergency hospital admissions: Longitudinal analyses of national administrative health, social care, and education data in Wales, UK* (SSRN Scholarly Paper No. 4204883). https://doi.org/10.2139/ssrn.4204883

Lyons, R. A., Jones, K. H., John, G., Brooks, C. J., Verplancke, J.-P., Ford, D. V., Brown, G., & Leake, K. (2009). The SAIL databank: Linking multiple health and social care datasets. *BMC Medical Informatics and Decision Making*, *9*(1), 3. https://doi.org/10.1186/1472-6947-9-3

Madley-Dowd, P., Thomas, R., Boyd, A., Zammit, S., Heron, J., & Rai, D. (2023). Intellectual disability in the children of the Avon longitudinal study of parents and children (ALSPAC) [version 2; peer review: 2 approved]. *Wellcome Open Research*, *7*(172).

Maughan, B., Collishaw, S., & Pickles, A. (1998). School achievement and adult qualifications among adoptees: A longitudinal study. *Journal of Child Psychology and Psychiatry*, *39*(5), 669–685. https://doi.org/10.1111/1469-7610.00367

McAuley, C., & Davis, T. (2009). Emotional well-being and mental health of looked after children in England. *Child & Family Social Work*, *14*(2), 147–155. https://doi.org/10.1111/j.1365-2206.2009.00619.x

McClung, M., & Gayle, V. (2010). Exploring the care effects of multiple factors on the educational achievement of children looked after at home and away from home: An investigation of two Scottish local authorities. *Child & Family Social Work*, *15*(4), 409–431. https://doi.org/10.1111/j.1365-2206.2010.00688.x

McGrath-Lone, L., Harron, K., Dearden, L., & Gilbert, R. (2020). Exploring placement stability for children in out-of-home care in England: A sequence analysis of longitudinal administrative data. *Child Abuse & Neglect*, *109*, 104689. https://doi.org/10.1016/j.chiabu.2020.104689

Melendez-Torres, G. J., Bourne, A., Reid, D., Hickson, F., Bonell, C., & Weatherburn, P. (2018). Typology of drug use in United Kingdom men who have sex with men and associations with socio-sexual characteristics. *International Journal of Drug Policy*, *55*, 159–164. https://doi.org/10.1016/j.drugpo.2018.01.007

Melis, G., Bedston, S., Akbari, A., Bennett, D., Lee, A., Lowthian, E., Schlüter, D., & Taylor-Robinson, D. (2023). Impact of socio-economic conditions and perinatal factors on risk of becoming a child looked after: A whole population cohort study using routinely collected data in Wales. *Public Health*, *224*, 215–223. https://doi.org/10.1016/j.puhe.2023.09.001

Muthén, L., & Muthén, B. (2017). *Mplus: User's guide* (8th ed.). Muthén & Muthén. https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf

Nandy, S., & Selwyn, J. (2013). Kinship care and poverty: Using census data to examine the extent and nature of kinship care in the UK. *British Journal of Social Work*, *43*(8), 1661. https://doi.org/10.1093/bjsw/bcs057

National Assembly for Wales. (2007). *Curriculum and assessment (2)*.

Nylund-Gibson, K., Grimm, R. P., & Masyn, K. E. (2019). Prediction from latent classes: A demonstration of different approaches to include distal outcomes in mixture models. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(6), 967–985. https://doi.org/10.1080/10705511.2019.1590146

O'Higgins, A., Sebba, J., & Gardner, F. (2017). What are the factors associated with educational achievement for children in kinship or foster care: A systematic review. *Children and Youth Services Review*, *79*, 198–220. https://doi.org/10.1016/j.childyouth.2017.06.004

O'Higgins, A., Sebba, J., & Luke, N. (2015). *What is the relationship between being in care and the educational outcomes of children?* University of Oxford.

Rahman, M. A., Todd, C., John, A., Tan, J., Kerr, M., Potter, R., Kennedy, J., Rice, F., & Brophy, S. (2018). School achievement as a predictor of depression and self-harm in adolescence: Linked education and health record study. *The British Journal of Psychiatry*, *212*(4), 215–221. https://doi.org/10.1192/bjp.2017.69

R: A language and environment for statistical computing [Computer software]. (2020). R Core Team. https://www.R-project.org/

Rock, S., Michelson, D., Thomson, S., & Day, C. (2015). Understanding foster placement instability for looked after children: A systematic review and narrative synthesis of quantitative and qualitative evidence. *The British Journal of Social Work*, *45*(1), 177–203. https://doi.org/10.1093/bjsw/bct084

Rodgers, S. E., Demmler, J. C., Dsilva, R., & Lyons, R. A. (2012). Protecting health data privacy while using residence-based environment and demographic data. *Health & Place*, *18*(2), 209–217. https://doi.org/10.1016/j.healthplace.2011.09.006

Rodgers, S. E., Lyons, R. A., Dsilva, R., Jones, K. H., Brooks, C. J., Ford, D. V., John, G., & Verplancke, J.-P. (2009). Residential anonymous linking fields (RALFs): A novel information infrastructure to study the interaction between the environment and individuals' health. *Journal of Public Health (Oxford, England)*, *31*(4), 582–588. https://doi.org/10.1093/pubmed/fdp041

Scourfield, J., Holland, S., & Young, C. (2008). Social work in Wales since democratic devolution. *Australian Social Work*, *61*(1), 42–56. https://doi.org/10.1080/03124070701818724

Sebba, J., Berridge, D., Luke, N., Fletcher, J., Bell, K., Strand, S. Thomas, S. Sinclair, I., O'Higgins, A., Thomas, S., Sinclair, I., & O'Higgins, A. (2015). *The educational progress of looked after children in England: Linking care and educational data*. Rees Centre, University of Oxford & University of Bristol.

Sinclair, I., Baker, C., Wilson, K., & Gibbs, I. (2005). *Foster children: Where they go and how they get on*. Jessica Kingsley Publications.

Statistics for Wales. (2021). *Children looked after census 2020-21*. Welsh Government.

StatsWales: Results, by local authority and level [Dataset]. (2011). https://statswales.gov.wales/Catalogue/Education-and-Skills/Schools-and-Teachers/Examinations-and-Assessments/Key-Stage-1/Results-by-LocalAuthority-Level

Strand, S. (2002). Pupil mobility, attainment and progress during key stage 1: A study in cautious interpretation. *British Educational Research Journal*, *28*(1), 63–78. https://doi.org/10.1080/01411920120109757

Sutcliffe, A. G., Gardiner, J., & Melhuish, E. (2017). Educational progress of looked-after children in England: A study using group trajectory analysis. *Pediatrics*, *140*(3), e20170503. https://doi.org/10.1542/peds.2017-0503

Tessier, N. G., O'Higgins, A., & Flynn, R. J. (2018). Neglect, educational success, and young people in out-of-home care: Cross-sectional and longitudinal analyses. *Child Abuse and Neglect*, *75*, 115–129.

Van IJzendoorn, M. H., Juffer, F., & Poelhuis, C. W. K. (2005). Adoption and cognitive development: A meta-analytic comparison of adopted and nonadopted children's IQ and school performance. *Psychological Bulletin*, *131*(2), 301–316. https://doi.org/10.1037/0033-2909.131.2.301

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020). Latent class analysis: A guide to best practice. *Journal of Black Psychology*, *46*(4), 287–311. https://doi.org/10.1177/0095798420930932

Welsh Government. (2016). *Raising the ambitions and educational attainment of children who are looked after in Wales*. https://gov.wales/sites/default/files/publications/2018-03/strategy-raising-the-ambitions-and-educational-attainment-of-children-who-are-looked-after-in-wales.pdf

Welsh Government. (2019a, January 23). *Welsh index of multiple deprivation (WIMD) index guidance, technical information and background documentation*. https://gov.wales/welsh-index-multiple-deprivation-index-guidance

Welsh Government. (2019b). *New GMS contract QAIF implementation dataset and business rules - learning disability indicator set (LD)*.

Winokur, M., Holtan, A., & Batchelder, K. E. (2014). Kinship care for the safety, permanency, and well-being of children removed from the home for maltreatment. *Cochrane Database of Systematic Reviews*, *2014*(1), 1. https://doi.org/10.1002/14651858.CD006546.pub3