Putting together the data jigsaw:

The extent of the linking of administrative datasets on children by local authorities in England. Responses to a Freedom of Information request

Sarah Martin-Denham, Nathan Scott, Karen Horridge, Anne Pinney

May 2023

University of Sunderland

Summary

Across England, administrative data is routinely collected by the NHS, education and social care services. Multiple national reviews have advocated for better sharing of data by local agencies to improve outcomes for children and their families. The aim of the research was to investigate the extent of data linkage of children and young people's data in Local Authorities (LAs) across England. The research questions were explored using a Freedom of Information (FOI) request submitted to LAs, including, Metropolitan Districts (36), London boroughs (32), City of London (1), Unitary authorities (55), and County Councils (24). The response rate was 91/148 requests. The results show that there is limited linking of health, educational and social care data on children by LAs in England.

This policy briefing presents interim findings, which will be updated when results from a similar request to integrated care systems (ICS) and outstanding responses from LAs are received. Local research is also underway to explore barriers to data linkage and enabling factors. The research team are also exploring policy solutions with local and national officials; initial recommendations, informed by this research, are shown at the end of this briefing.

Policy context

The Special Educational Needs Disability (SEND) review (Her Majesty's Government, 2022) proposed to make better use of effective data sharing, particularly across education and health partners and 'to promote this more widely across the system' (p. 70). The SEND and Alternative Provision (AP) Improvement Plan (HMG, 2023) proposes national and local inclusion dashboards to present data across education, health and care 'and provide a basis for measuring whether we are achieving our mission of improved outcomes, better experiences and a financially sustainable system' (p. 71).

The Independent Review of Social Care (MacAlister, 2022) recommended that the Department for Education (DfE) should have a proactive strategy on making better use of data in children's social care, including a strategy for data linking with other sources such as education, hospital and justice data. The review called for a 'consistent identifier' to ensure that data could be 'easily, quickly and accurately linked' (p. 62). Debate on the Health and Care Act (2022) advocated for the NHS number as a proposed solution, the Government is due to report soon on their proposals to improve children's data-sharing and to explore the implementation of a consistent identifier. The DfE strategy (2023) Stable homes, built on love: children's social care strategy and consultation recognises that data is not consistently available to support practitioners, setting out their intention to deliver social care dashboards by the end of this Parliament. This recognises that there are 'data gaps' and burdens on councils that prevent better use of data, with a strategy to address this to be published by the end of 2023.

Benefits of linking data

Linking data from public services can enhance the provision of education, healthcare, and social care (Downs et al., 2016). The breadth and depth of local data sets offers opportunities to determine the impact of interventions to improve outcomes for children (ADR UK, 2023). Data linkage can be useful in identifying those at risk of negative outcomes to enable early intervention and gathering information about specific groups (Atherton et al., 2015) from a wider population (Harron et al., 2016). The Supporting Families programme has promoted data-sharing and linkage 'to help local services to identify, understand and better support children and families at the right time, to prevent them reaching crisis point' (Ten years of Supporting Families: annual report, 2022-2023). Opportunities to understand public health needs can be provided by linking data, which helps to identify those in need and to inform interventions (Sohal et al., 2022). Linking data between healthcare and education records provides useful ways of looking at impacts of mental health and illness on academic achievement (Downs et al., 2019), for example.

Enhance the provision of education, healthcare, and social care

Useful in identifying those at risk of negative outcomes

Better support children and families at the right time

Barriers to data linking

The first barrier to meaningful data linkage is inadequate data capture in individual services e.g., in health, not having the facility to easily capture data at each point of care; in education, a strong focus on the perceived primary need of each child, instead of all their needs. Different categories used by each service to describe children's needs also present an obstacle (Pinney 2017, p.9, p.19). Barriers to data linkage include issues of consent surrounding data sharing, difficulty performing required analysis, and quality of data (Atherton et al., 2015) and digital maturity (National Data Strategy, section 2.4, DCMS 2020). To effectively link data, a unique personal identifier is needed to make it possible to link data from different administrative sources with marginal error (Ludvigsson et al., 2009).

There are social benefits of linking data, however, it requires considerable organisation and is demanding in terms of leadership and co-operation (Holman et al., 2008). Co-ordinated action and national leadership is needed by Government departments to support LAs, education, health and police to achieve 'frictionless data sharing' (MacAlister, 2022, p. 83).

Aim and objectives

To investigate the extent of data linkage of children's data at the local authority level in England:

1. Determine which LAs are linking data sets
2. Determine which datasets LAs are linking
3. Identify how LAs are using linked data

Methods

Between February and April 2023, freedom of information (FOI) requests were sent to Metropolitan Districts (36), London Boroughs (32), City of London (1), Unitary Authorities (55) and County Councils (24) (for the questions see Appendix 1). The response rate was 91/148 for complete responses (Table 1). There were six partial responses that were not included in the analysis. The data was analysed using descriptive statistics. For ease of reading, 'children' is used to include 'children and young people'.

Table 1. Responses to FOI by LA type as of (10.05.2023)

	Unitary Authorities	Metropolitan Districts	London boroughs + City of London	County councils	Total
Total FOIs attempted	55	36	33	24	148
Full responses	33	23	21	14	91
Partial responses	3	0	3	0	6
Responded with queries	5	1	1	1	8
Awaiting response	14	11	7	9	41
No way of contacting for information[1]	0	1	1	0	2

[1] The FOI request could not be sent to Hackney Borough Council and Wirral Borough council, due to a website fault, and a restrictive character limit respectively.

Results

The results are presented in order of the FOI questions. Percentages are rounded to one decimal place.

❶ Do you currently link your education and social care data for children and young people?

As shown in Figure 1, 67 (73.6%) reported that they were linking education and social care data. These datasets will be held internally by LAs. Seventeen (18.7%) said linking these datasets was in development and seven (7.7%) were not yet linking data. The LAs that do not link their data were predominantly in the South of England. The most common reasons given were that their internal systems were not compatible with the linking of data.

Key: Bubble size representative of number of responses ● Yes ● In development ● No

Figure 1. Number of LAs linking education and social care data

❷ Which of your datasets are you linking?

The LAs were asked to report the education and social care data sets they were linking (school census, Children in Need (CiN) census, Children looked after return or other). Figure 2 illustrates that the most likely dataset to be linked was school census data (70.3%), followed by the looked-after children return (59.3%) and the CiN census (54.9%). Fifty-three LAs (58.2%) reported linking other datasets such as early help and youth offending. On analysis, the other datasets largely related to subsets found within school census, like children on Free School Meals (FSM), and CiN census data, like children known to social care.
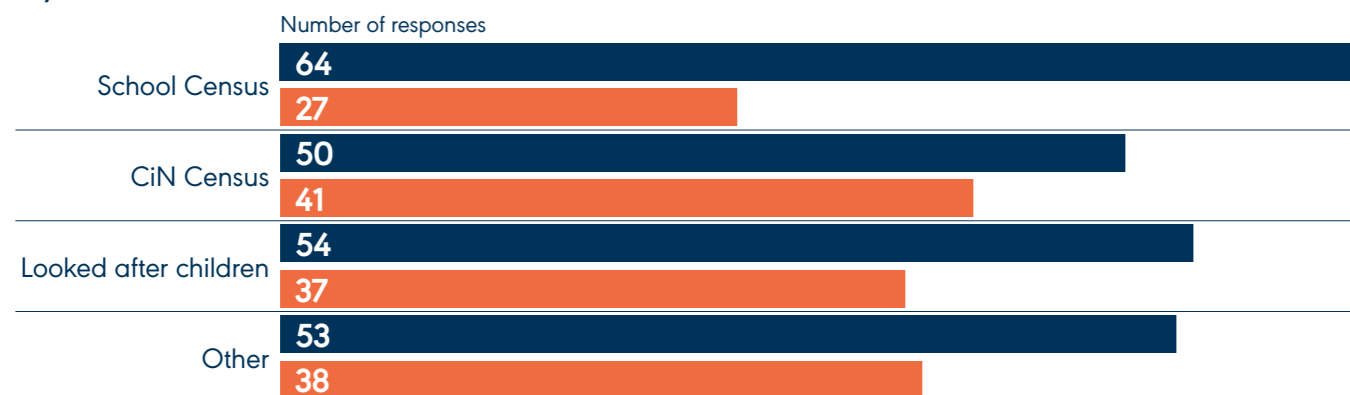
Key: ■ Yes ■ No

Number of responses

School Census	64
	27
CiN Census	50
	41
Looked after children	54
	37
Other	53
	38

Figure 2. Education and social care datasets councils are linking

❸ Which groups of children and young people do you link data on?

Figure 3 illustrates that, while the majority of LAs were linking data on certain groups of children, less than half (49.5%) were linking data on all children in the LA. LAs were predominantly linking data for children with SEND (82.4%), looked after children (82.4%), those on the CiN census (80.2%), and children on Free School Meals (FSM) (70.3%).
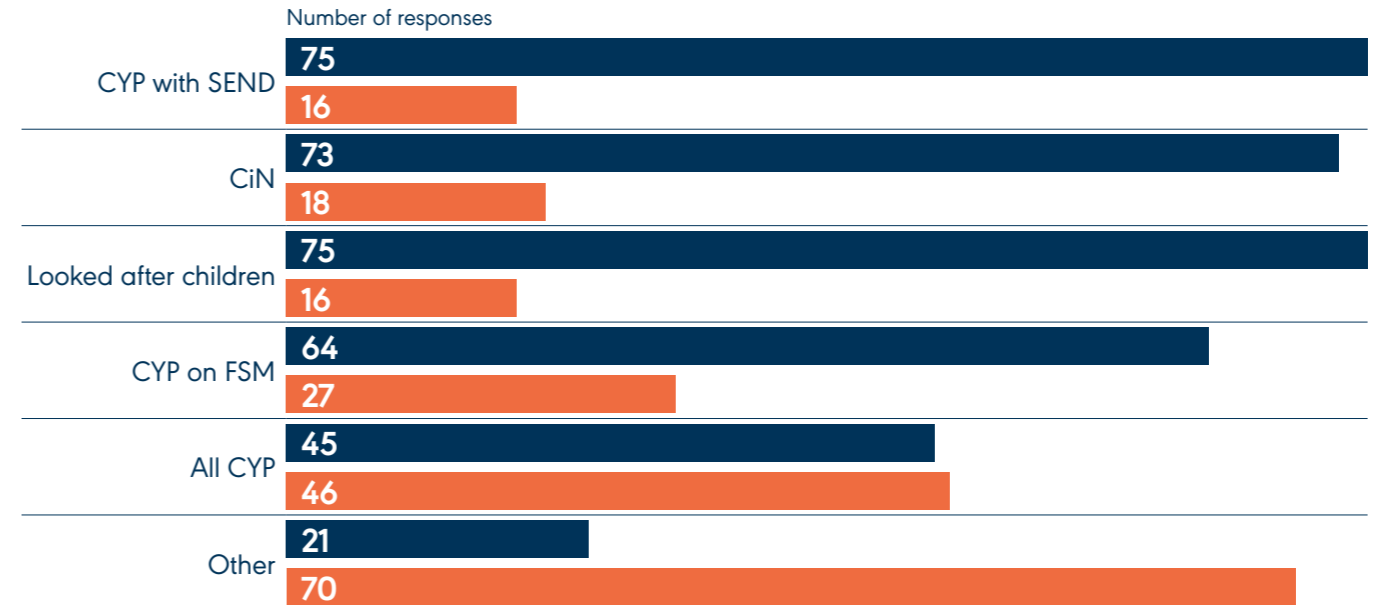
Key: ■ Yes ■ No

Number of responses

CYP with SEND	75
	16
CiN	73
	18
Looked after children	75
	16
CYP on FSM	64
	27
All CYP	45
	46
Other	21
	70

Figure 3. Groups of children that LAs are linking data on

❹ How are you using the linked data?

Data dashboards (74.7%) and strategic planning (70.3%) were the most common uses of linked data. Figure 4 shows that LAs are less likely to use linked data for joint commissioning decisions (41.8%), service review (49.5%) or targeted interventions% (52.75). Thirty-two LAs (35.2%) also shared 'other' uses of linked data. Most commonly this was for the Supporting Families programme or to complete statutory returns.
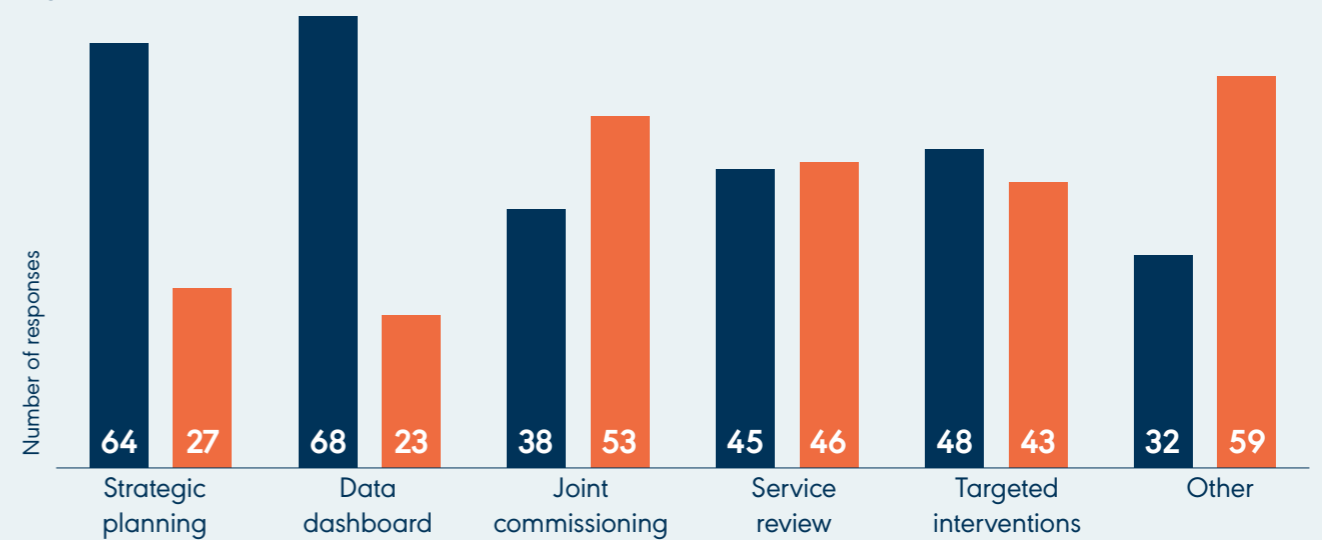
Key: ■ Yes ■ No

Strategic planning	Data dashboard	Joint commissioning	Service review	Targeted interventions	Other
64 / 27	68 / 23	38 / 53	45 / 46	48 / 43	32 / 59

Figure 4. How LAs are using linked data

5 **Do you currently link LA data sets for children and young people with health data sets?**

Figure 5. indicates that only 22 (24.2%) are linking LA data with health datasets, 20 (22.0%) are in development, and 4 (53.8%) are not linking any health datasets with LA data.

Key: ■ Yes ■ In development ■ No

Number of responses

| 22 (24.3%) | 20 (24.3%) | 49 (53.8%) |

Figure 5. Number of LAs linking education, social care and health datasets

6 **Which health data sets are you linking with education and or social care?**

Although some LAs responded that they were linking health, education and social care datasets, they were only linking limited health data sets. For example, only three LAs (3.2%) link health data in the Community Services Data Set (CSDS) (paediatric disability data) with education/social care data (which has the facility to include data about the multifaceted needs of children, including health conditions, diagnoses, parent reported needs, technology dependencies and need (or not) for round the clock care/continuous supervision). This is the key dataset for understanding trends in childhood disability, in line with the expectations set in the SEND Green Paper. Paediatric disability data is only in the CSDS if clinicians have the facility to report this, ideally at the point of care, which most do not (Horridge and Harvey et al., 2016). The CSDS is, however, the only potential source of such health data, unless it is present in GP data also.

Key: ■ Yes ■ No

Number of responses

Primary care: 11 / 80
CSDS*: 3 / 89
CAHMS: 10 / 81
CYPS: 8 / 83
A&E: 8 / 83
SUS / HES: 2 / 89
Child Health Records: 8 / 83
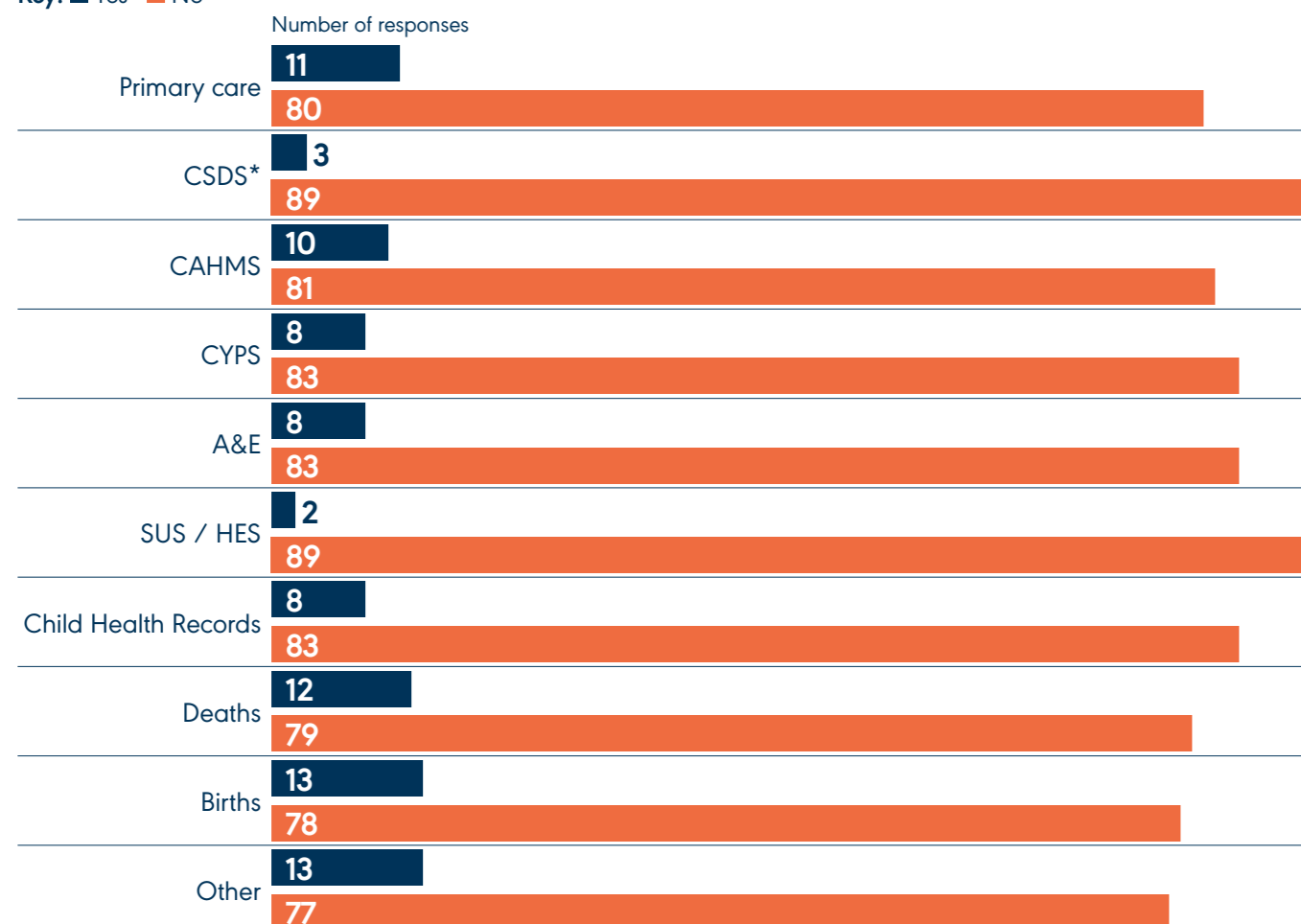Deaths: 12 / 79
Births: 13 / 78
Other: 13 / 77

Figure 6. Which health data sets are currently being linked with education and social care datasets

*CSDS: Community Services Dataset, CAMHS: Child and Adolescent Mental Health Service, CYPS: Children and Young People's Service, A&E: Accident and Emergency, SUS: Secondary Uses Services, HES: Hospital Episode Statistics.

Table 2 summarises the health datasets that 22/148 LAs are linking, including the CSDS. Only Manchester, North Tyneside and Luton reported linking CSDS; there is insufficient information to know if these data include any diagnostic information or whether the quality of the data is satisfactory. More encouragingly, some areas are linking a range of health data, most commonly, births and deaths, GP primary care and CAMHS data.

Table 2. Linked health datasets by type of council

Name of council	Health datasets	CSDS linked Y/N
Unitary authorities		
Cheshire East	Primary care (GP)	● N
Hartlepool	Primary care (GP), Accident and Emergency (A&E)	● N
Leicester City Centre	Children and Young People's Service (CYPS), child health records, births and deaths	● N
Luton	Primary care (GP), CSDS, Child and Adolescent Mental Health Service (CAMHS), CYPS	● Y
North Somerset	Births and deaths	● N
Northumberland	Deaths	● N
Portsmouth	School nursing and health visiting data	● N
Reading	Births and deaths	● N
Redcar	Child health records, CAMHS	● N
Swindon	CYPS, child health records	● N
Wiltshire	CAMHS, A&E, Secondary Uses Services/Hospital Episode Statistics	● N
London boroughs		
Newham	Primary care (GP), CYPS, A&E, child health records, births and deaths	● N
Metropolitan districts		
Birmingham	Births, NHS child protection	● N
City of Leeds	Primary care (GP) and child protection	● N
Manchester	Primary care (GP), CSDS, CAMHS, births and deaths	● Y
North Tyneside	Primary care (GP), CAMHS, A&E, CSDS and Child Health Records	● Y
Rochdale	Births and deaths	● N
Solihull	CAMHS	● N
County councils		
Derbyshire	CAMHS	● N
Hampshire	Primary care (GP), CAMHS, CYPS, births and deaths	● N
Norfolk	CAMHS, A&E data, births and deaths	● N
Somerset	CAMHS, A&E data (limited), primary care (GP) (limited)	● N

Discussion

The aim of this research was to investigate the extent of data linkage of children and young people's data at the local level.

Determine which LAs areas are linking data sets

Currently, 27.5% of LAs are not linking internally held education and social care data. Less than half of those that are linking data, were linking data on all children in their local area. Instead, they opted to link data on specific groups of children, such as those identified to have SEND rather than those who actually have SEND. This could indicate that LAs do not have the capacity or systems to link data or they don't see the value. However, rationale provided in the FOI responses for not linking data was scarce.

75.8% of LAs were not yet linking their internal datasets with external health datasets. Only 22 LAs were linking to health data and 20 reported it was in development. This could be due to capacity, a lack of funding, not knowing how to begin to work with data controllers in health, information governance barriers to link data or not foreseeing the benefit. Also, our local research (on-going) indicates that health reorganisation (with the move to ICS) is also creating challenges (e.g., change of personnel, larger geographic footprint).

One of the LAs stated that their local government 'provides datasets to health for their needs but does not receive health datasets in return'. This suggests that ICBs may be performing data linkage for children in certain areas rather than the LAs. This will be addressed in an upcoming publication, as the researchers sent a similar FOI request to all NHS ICBs.

Determine which datasets LAs are linking

The most likely dataset to be linked was school census data, followed by the looked-after children return and the CiN census. Of the LAs that were linking social care and education data, the majority were linking all three of the aforementioned datasets. Some LAs noted that they were not linking specific datasets such as the children looked after data return (also known as SSDA903). Rather, they were undertaking data linkage on all children known to social care, which would include this subset of looked after children. This

was also the case for linking data on specific groups of children, where some LAs were linking data for all children on the school census, rather than those on FSM specifically.

Regarding health datasets, only three LAs were linking the CSDS. There is insufficient information to know if these data include any diagnostic information or to determine the quality of the data. The most common datasets to be linked were births and deaths, followed by primary care (GP), and CAMHS. The discrepancy between data linkage for CSDS and other datasets may indicate that this dataset is not flowing, or LAs are linking data through local data sharing agreements and are not identifying CSDS as an essential dataset for their purposes. It could also be the case that LAs don't know the content of individual datasets and therefore the value they can bring in improving outcomes for children.

Identify how LAs are using linked data

LAs most commonly use linked data to create data dashboards and for strategic planning. However, it is unclear what the outcomes of strategic planning were, as most LAs (in their answers to the FOI) indicated they were not performing joint commissioning or service reviews. Targeted interventions were also a less common use of linked data than data dashboards or strategic planning. There was some indication that respondents did not understand what joint commissioning entailed, which is surprising given the expectations established in section 26 of the Children and Families Act 2014. A frequent response under the 'other' category was that LAs were using linked data to identify families eligible for the Supporting Families programme. Guidance on this programme has encouraged data sharing and linkage for meaningful outcomes measurement, which could be emulated by those Government departments responsible for children's health, education and care (e.g., Supporting Families programme guidance 2022-25, Chapter 2 – Data Transformation).

National recommendations (initial)

During 2022-23, the SEND and AP Review, the Independent Children's Social Care Review and debate on the Health and Care Act recommended data-sharing to improve outcomes for children and young people.

A joined-up approach is urgently needed at the national level, working together to realise the benefits of data-sharing and linkage and to tackle the barriers to data-sharing and linkage that exist at the local level, particularly between health and local authority services.

We recommend a national strategy to tackle the barriers to linking children and young people's data at the local level, especially between local authority and health services, including:

a. A consistent unique identifier to support efficient and effective data-sharing and linkage

b. Opportunities for LAs and health partners to share learning on how they have successfully linked data to improve outcomes for children

c. Best practice guidance and templates to make it easier for LAs and health partners to develop data-sharing protocols and processes

d. A grants programme to stimulate and evaluate progress

e. A consistent data capture interface to improve reporting of paediatric disability data in the Children's Community Services Data Set, a key dataset for understanding trends in SEND and the multi-faceted needs of disabled children and their families.

f. Data capture at all points of care and services across agencies needs to be in place, to ensure that quality data are available to link.

The national strategy should draw on learning from the Supporting Families programme, which has successfully promoted data-sharing and data linkage to identify and support families with complex needs holistically and to capture outcomes.

Future research

Future studies should attempt to gain greater insight into the specifics of LA data linkage practices. For example, it would be beneficial to clarify whether LAs are linking data systematically (e.g., automatically linking any new data entered), or whether they are performing post-hoc data linkage as required (e.g., one-off linking a child's social care and education data to undertake a targeted intervention). The authors are currently involved in SEND data linkage projects in the Northeast of England.

Limitation

There are a number of potential limitations of this research. First, fifty-seven councils have not yet provided a complete response to the FOI request. It is expected that a number of these responses will be received after this report is published, and an updated report will be published at a later date. A higher response rate will allow a more comprehensive understanding of data linkage practices in England.

Second, verification of the data was not undertaken, though clarity from councils was sought where responses were ambiguous. It was not possible to assess the quality of data where LAs responded that they were linking certain datasets to their internal data. Future research may benefit from querying the quality of data in the linked datasets.

Third, due to an oversight, not all councils were contacted; Isles of Scilly, Dorset and West Northamptonshire were omitted from the contact list. This has been rectified and these LAs will be included in the full publication (if a response is received).

Fourth, some of the language used in the FOI request may have lacked clarity. For instance, based on extra information supplied by LAs and some queries that were received, it became clear that some LAs were unsure whether post-hoc linking of data on an as-required basis constituted data linkage. There was also considerable overlap between responses to questions 2 and 3 e.g., 'Children on FSM' as a dataset that LAs are linking (Question 2), but also a group of children that LAs are linking data on (Question 3).

Fifth, based on queries that were received, it appears that the health dataset known to the University of Sunderland research team as CYPS data is not known to many LAs, and the data may be collected under a different name, or may be in scope of an existing dataset. This may reflect variations in local service configuration in different regions.

References

ADR UK (2023) *Linked local data on children and young people*. Available at: https://www.adruk.org/our-work/browse-all-projects/linked-local-data-on-children-and-young-people/

Atherton, I.M., Lynch, E., Williams, A.J. and Witham, M.D. (2015) 'Barriers and solutions to linking and using health and social care data in Scotland', *The British Journal of Social Work*, 45(5), pp. 1614-1622.

Department for Education (2023) *Stable homes, built on love: children's social care reform consultation*. London: DfE. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1147317/Children_s_social_care_stable_homes_consultation_February_2023.pdf

Department for Levelling up Housing and Communities and Department for Education (2022) *Supporting Families programme guidance 2022 to 2025*, Chapter 2, section on Data Transformation. Available at: https://www.gov.uk/government/publications/supporting-families-programme-guidance-2022-to-2025/chapter-2-delivering-supporting-families#data-transformation

DLUHC (March 2023) Ten years of Supporting Families: Supporting Families programme Annual Report 2022-23. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1140924/Supporting_Families_Annual_Report_2023.pdf

Downs, J., Gilbert, R. and Hayes, R.D., Hotopf, M. and Ford, T. (2017) 'Linking health and education data to plan and evaluate services for children', *Archives of Disease in Childhood*, 102, pp. 599-602.

Downs, J.M., Ford, T., Stewart, R., Epstein, S., Shetty, H., Little, R. and Hayes, R. et al. (2019) 'An approach to linking education, social care and electronic health records for children and young people in South London: a linkage study of child and adolescent mental health service data', *British Medical Journal*, 9(1), pp. 1-12.

Harron, K., Dibben, C., Boyd, J., Hjern, A., Azimaee, M., Barreto, M.L. and Goldstein, H. (2017) 'Challenges in administrative data linkage for research', *Big data & society*, 4(2).

Health and Care Act 2022 Part 6, S.175 - https://bills.parliament.uk/bills/3022

HM Government (2022) *SEND Review: Right support, right place, right time Government consultation on the SEND and alternative provision system in England*. London: HM Stationery Office.

HM Government (2023) *Special educational needs and disabilities (SEND) and alternative provision (AP) improvement plan: Right support, right place, right time*. London: HM Stationery Office.

Holman, C.D.A.J., Bass, J.A., Rosman, D.L., Smith, M.B., Semmens, J. B., Glasson, E.J. and Stanley, F.J. et al. (2008) 'A decade of data linkage in Western Australia: strategic design, applications and benefits of the WA data linkage system', *Australian Health Review*, 32(4), 766-777.

Horridge, K.A. et al. (2016) 'Quantifying multifaceted needs captured at the point of care. Development of a Disabilities Terminology Set and Disabilities Complexity Scale', *Dev Med Child Neurol*. Available at: https://pubmed.ncbi.nlm.nih.gov/27009933/.

Ludvigsson, J.F. et al. (2009) 'The Swedish personal identity number: Possibilities and pitfalls in healthcare and medical research', *European Journal of Epidemiology*, 24, p. 659–667.

MacAlister, J. (2022) *The independent review of children's social care: final report*. Available at: https://webarchive.nationalarchives.gov.uk/ukgwa/20230308122535mp_/https://childrenssocialcare.independent-review.uk/wp-content/uploads/2022/05/The-independent-review-of-childrens-social-care-Final-report.pdf.

Olsen J, Bronnum-Hansen H, Gissler M, et al. (2010) 'High-throughput epidemiology: Combining existing data from the Nordic countries in health-related collaborative research', *Scandinavian Journal of Public Health* 38, pp. 777–779.

Pinney, A. (Feb 2017) *Understanding the needs of disabled children with complex needs or life-limiting conditions*. Available at: https://pubmed.ncbi.nlm.nih.gov/27009933/.

Sohal, K., Mason, D., Birkinshaw, J., West, J., McEachan, R.R., Elshehaly, M. and Wright, J. et al. (2022) 'Connected Bradford: a Whole System Data Linkage Accelerator', *Wellcome Open Research*, 7(26), pp. 26.

Appendix 1

1. Do you currently link your education and social care data for children and young people?

Y/N/in development

2. Which of your datasets are you linking? please state all the apply

- School census
- Children in Need (CiN) census
- Children looked after return SSDA903 data collection
- Other (please specify)

3. Which groups of children and young people do you link data on? Select all that apply

- Children and young people with Special Educational
- Needs and Disabilities
- Children in Need
- Children looked after (care experienced)
- Children and young people on Free School Meals
- All children and young people
- Other (please specify)

4. How are you using the linked data?

- Strategic planning
- Data dashboard
- Joint commissioning
- Service review
- Targeted interventions (individuals)
- Other (please specify)
- We aren't using linked data

5. Do you currently link council data sets for children and young people with health data sets? Y/N/in development

6. Which health data sets are you linking with education and or social care? Select all that apply

- Primary care (GP data)
- Community Services Data Set (CSDS)
- Secondary care (CAMHS)
- Children and Young Peoples Service (CYPS)
- Accident and Emergency data
- Secondary Uses Services (SUS) also known as HES (Hospital Episode Statistics)
- Child Health Records
- Deaths
- Births
- Other (please specify)