Towards a Family Justice Observatory

A scoping study

Who cares for children? Population data for family justice research

Authors: Matthew A Jay,[1,2] Jenny Woodman,[3] Karen Broadhurst[4] and Ruth Gilbert[1,2]

1: Population, Policy & Practice Programme, UCL Great Ormond Street Institute of Child Health
2: Administrative Data Research Centre for England
3: Thomas Coram Research Unit, UCL Institute of Education
4: Centre for Child and Family Justice Research, Lancaster University

Please cite this as: Jay, M.A., Woodman, J., Broadhurst, K., Gilbert, R. (2017) *Who cares for children? Population data for family justice research*. Available at: http://wp.lancs.ac.uk/observatory-scoping-study/

Published October 2017

In all actions concerning children, whether undertaken by public or private social welfare institutions, courts of law, administrative authorities or legislative bodies, the best interests of the child shall be a primary consideration.

United Nations Convention on the Rights of the Child 1989, art 3(1)

When a court determines any question with respect to the upbringing of a child … the child's welfare shall be the court's paramount consideration.

Children Act 1989, s 1(1)

Summary contents

Abbreviations 8

Glossary 9

Acknowledgments 11

About the authors 11

Key points and recommendations 12

Executive summary 13

Introduction 15

1. What is the Family Justice System? 18

2. Family justice data 24

3. '*The child's present and future life as a human being*': a wider lens 28

4. Using family justice data 36

5. Next steps 57

References 65

Appendix I – Cafcass data profile 76

Appendix II – FamilyMan data profile 91

Appendix III – Legal aid data profile 95

Appendix IV – Children looked after and child in need data profile 100

Appendix V – Useful resources 112

Detailed contents

Abbreviations 8

Glossary 9

A note on referencing 10

Acknowledgments 11

About the authors 11

Key points and recommendations 12

Executive summary 13

Introduction 15

Who cares for children? 15

Towards a Family Justice Observatory 17

Locally-held local authority data 17

Methods and ethical approval for this project 17

1. What is the Family Justice System? 18

1.1. Deciding who cares for children 18

1.2. The family courts 20

1.3. Public and private family law 20

1.4. Children's social care 23

2. Family justice data 24

2.1. Core datasets 24

2.2. The complete picture: linking court and social care data 25

3. '*The child's present and future life as a human being*': a wider lens 28

3.1. Trajectories through the system 28

3.2. What data are available? 29

4. Using family justice data 36

4.1. A conceptual framework 36

4.2. Determinants of variation in practice and outcomes 38

4.3. Evaluating variation 40
4.4. The strengths and limitations of using linked cross-sectoral administrative data 43
4.5. Ethical, legal and governance issues 45
4.6. Linkage methods 46
4.7. Putting data into practice: three modes of action 49
4.7.1. An intermission: the place of evidence based on populations in legal scholarship 50
4.7.2. Research evidence to guide policy development, legislation and service planning 51
4.7.3. Research evidence to inform decisions about who cares for children 51
4.7.4. Using real time data to identify and approach specific children or families 55
5. Next steps 57
Conclusion 64
References 65
Appendix I – Cafcass data profile 76
A1.1. Overview 76
A1.2. Data completeness and accuracy 77
A1.2.1. Population 77
A1.2.2. Validation checks and missing data 78
A1.2.3. Changes over time 84
A1.3. Permission pathway 85
A1.4. Linkage 86
A1.5. Legal complexity and implications for research 86
A1.6. Literature review of studies using the Cafcass data resource 88
Appendix II – FamilyMan data profile 91
A2.1. Overview 91
A2.2. Data completeness and accuracy 92
A2.2.1. Population 92
A2.2.2. Validation checks and missing data 92
A2.2.3. Changes over time 92

A2.4. Permission pathway 92
A2.5. Overlap with Cafcass data 93
A2.6. Linkage 93
A2.7. Literature review of studies using the FamilyMan data resource 94
Appendix III – Legal aid data profile 95
A3.1. Overview 95
A3.2. Data completeness and accuracy 96
A3.2.1. Population 96
A3.2.2. Validation checks and missing data 96
A3.2.3. Changes over time 98
A3.3. Permission pathway 98
A3.4. Linkage 98
A3.5. Literature review of studies using the legal aid data resource 99
Appendix IV – Children looked after and child in need data profile 100
A4.1. Overview 100
A4.2. Data completeness and accuracy 102
A4.2.1. Population 102
A4.2.2. Validation checks and missing data 102
A4.2.3. Changes over time 104
A4.3. Permission pathway 104
A4.4. Linkage 105
A4.5. Data items not returned 106
A4.6. Literature review of studies using the CLA and CiN data resources 107
Appendix V – Useful resources 112

TABLES

Table 1: Overview of Family Justice System data sources 26
Table 2: Overview of population-level datasets from wider services 30

Table 3: Examples of studies using family justice data 41
Table 4: The Five Safes 45
Table 5: Availability of identifiers across the family justice datasets 47
Table 6: Missing data in Cafcass public law cases by year case started 80
Table 7: Missing data in Cafcass private law cases by year case started 81
Table 8: Summary of studies using record-level Cafcass data 88
Table 9: Proportion of closed cases with missing characteristics data in legal aid dataset, 2016/17 .. 97
Table 10: Summary of studies using record-level CLA and CiN data 107

FIGURES

Figure 1: Schematic representation of who cares for children 16
Figure 2: Support and interventions for vulnerable families 19
Figure 3: The relationship between private and public law proceedings 22
Figure 4: Overlap between Family Justice System datasets 27
Figure 5: The journey in and out of family courts 29
Figure 6: A conceptual framework linking child maltreatment to longer-term consequences 37
Figure 7: Hypothetical linkage between Cafcass data and the National Pupil Database 48
Figure 8: The interplay between data, research evidence and practice 50
Figure 9: Evidence-informed family justice 53
Figure 10: Simplified representation of a private law case and Cafcass' involvement 78
Figure 11: Cafcass research approval process 85
Figure 12: An overview of family law court orders 87
Figure 13: Potential linkage between Cafcass, CLA and CiN data 106

Abbreviations

ADRN	Administrative Data Research Network
Cafcass	Children and Family Court Advisory and Support Service
CiN	Child in need
CLA	Children looked after
CSC	Children's social care
DfE	Department for Education
EIFJ	Evidence-informed family justice
FJO	Family Justice Observatory
FJS	Family Justice System
LAA	Legal Aid Agency
MoJ	Ministry of Justice
NPD	National Pupil Database
SGO	Special guardianship order
TTP	Trusted Third Party
UPN	Unique Pupil Number

Glossary

A more comprehensive glossary of terms relating to the use of administrative data can be found on the Administrative Data Research Network's website.[1] Appendix V also contains links to other useful resources.

Administrative data. Information about people, businesses and other organisations that is collected by any government department or agency, for delivering their day-to-day services. It can include information such as tax records, school records, health information, etc.[1] Administrative data are not collected for research but can be put to such use as a secondary purpose.

Birth cohort. See Cohort studies.

Cohort studies. A type of investigation where a group of individuals are followed up over time with repeated data collections. The group followed up have a common inception point which could be, for example, birth or first use of a service. Studies which recruit at or near birth are known as birth cohorts. Those which recruit a sample of the population of different ages are sometimes known as panel surveys.

Longitudinal data. Data that capture information about individuals at multiple time points.

Panel survey. See Cohort studies.

Population-level data. Data from which it is possible to draw valid inferences about defined populations. The dataset could be a census of the entire population or a representative sample drawn randomly from it. The population might be the entire population of a country or a defined subset thereof, for example all school-aged children, all girls, all children in London, all births in 2004 or others.

Private law. Strictly speaking, the branch of law that deals with disputes between private parties (e.g., breach of contract or negligence). In this report, this term is used to mean private family law.

Private family law. The branch of family law that governs disputes between private individuals such as divorce or where and with whom a child should live.

Public law. Strictly speaking, the branch of law that deals with the exercise of public power and the constitution. In this report, this term is used specifically to mean public family law.

Public family law. The branch of family law governing state intervention in the upbringing of children, such as via care proceedings initiated by local authorities.

Record-level data. Sometimes referred to as microdata, record-level data are data at the smallest unit of analysis, such as each individual's record within a dataset.

Qualitative data. Unstructured information on data subjects collected using research methods, such as participant observation or case studies to record people's attitudes, feelings and behaviours in greater depth, which result in a narrative, descriptive account of a setting or practice which typically cannot be numerically measured.[1]

Quantitative data. Structured information on data subjects collected using research methods, such as surveys or questionnaires which allow for the measurement of variables, within a collection of people or groups, and resulting in numerical data which can be subjected to statistical analysis.[1]

Trusted Third Party. A Trusted Third Party (TTP) performs the matching of direct identifiers from different data sources, or the matching of direct identifiers of a single data source against an existing population spine.[1]

A note on referencing

We use both footnotes and endnotes in this report. Footnotes with symbols (*, †, etc) are used for legal citations as well as points of clarification that would otherwise interrupt the main text. When used in Tables, the note can be found at the end of the Table. All other citations (e.g., journal articles, books, reports and websites), are in the endnotes which are referenced with superscript Arabic numerals. The endnotes contain only the citation and no commentary and can therefore be ignored until a particular reference is sought.

Acknowledgments

We are grateful for the contributions of the following:

Stu Bedston	Lancaster University
Ross Black	Ministry of Justice
Richard Field	Legal Aid Agency
Louise Mc Grath-Lone	Administrative Data Research Centre for England
Jigna Patel	Children and Family Court Advisory and Support Service
John Simmonds	CoramBAAF
Linda Wijlaars	University College London Great Ormond Street Institute of Child Health

The Children and Family Court Advisory and Support Service (Cafcass) provided administrative data that form part of this report. We have also drawn information from a knowledge exchange event held in January 2017. A list of attendees and contributors to that event is available in a separate report.[2] We would like to thank the members of the wider Family Justice Observatory study team: Prof Judith Harwin, Claire Mason and Dr Laura Robertson (Lancaster University), Pippa Coutts and Jonathan Breckon (the Alliance for Useful Evidence), Susannah Bowyer and Julie Wilkinson (Research in Practice) and Prof Harriet Ward, Dr Lisa Holmes and Dr Georgia Hyde-Dryden (Centre for Child and Family Justice Research, Loughborough University). We are grateful to our policy and practice partners Dr John Simmonds (CoramBAAF) and Cathy Ashley (Family Rights Group) and project advisors Dr Julia Brophy (Association of Lawyers for Children) and Prof Judith Masson (Bristol University). Finally, we thank the Nuffield Foundation, in particular Teresa Williams and Tracey Budd, for funding, facilitation and guidance throughout this project.

About the authors

Matthew Jay is a social epidemiologist with a background in law. His interests include the interface between epidemiology and statistics and the study of law and legal processes. **Jenny Woodman** is a public health researcher with training in epidemiology and qualitative methods whose research focuses on the contribution that healthcare services can make to the health and well-being of vulnerable children and families. **Karen Broadhurst** is a Professor of Social Work with interests in family justice including the operation of the courts and more broadly in society. **Ruth Gilbert** is a Professor of Clinical Epidemiology whose contributions have included advances in evidence-based child health, studies on child maltreatment and methodologies for the use of clinical and administrative data. The Administrative Data Research Centre for England is funded by the Economic and Social Research Council (ESRC ES/L007517/1). https://adrn.ac.uk/centres/england.

Key points and recommendations

Key points	Recommendations (audience of recommendation in brackets)
Linkage across Family Justice System administrative data could transform our understanding of services and long-term outcomes. These data have not yet been linked for research.	1. Routine linkage is needed between the family justice data held by Cafcass, the Ministry of Justice, the Department for Education and local authorities as are mechanisms to authorise re-use of linked data for approved research projects (data providers).
There are serious limitations in the Family Justice System data that reduce their utility for research and operational purposes.	2. To follow up children over the life course within datasets, identifiers should be consistently allocated and linked, including for adopted children (data providers). 3. To understand who benefits from services, each child record in national administrative data should contain information about the child and family and services offered and received (data providers).
Data held by the Ministry of Justice are not currently available for external research use.	4. The court and legal aid data held by the Ministry of Justice should be made available for approved research projects (data providers).
To understand the lives of families in and through the system family justice data need to be linked to other services such as education and health. Few studies have done this.	5. Data providers and the Family Justice Observatory should advocate for linkage, including considering how the Digital Economy Act 2017 and existing frameworks can enable cross-sectoral linkages by trusted third parties (data providers and the Family Justice Observatory).
There is a lack of transparency about how data providers process and link data prior to release for analysis which limits the validity of findings.	6. It is necessary to improve data providers' understanding of the needs of researchers and their capacity to evaluate and report their methods for data processing (data providers and researchers).
Using family justice administrative data requires expertise in quantitative and qualitative methodology as well as an understanding of the law and the system but few research teams bridge these disciplines.	7. Research funders and the Family Justice Observatory should support capacity building and interdisciplinary collaboration between those with appropriate skills and knowledge (research funders and the Family Justice Observatory).
A failure to distinguish between the use of anonymised administrative data for research and the use of identifiable records for case management could threaten public support for research.	8. An examination of how research evidence can be put into practice is required, particularly at the level of individual decisions, taking account of the fact that using research evidence is separate from using real-time identifiable client data (services and researchers).
The family justice research community is emerging and to date there has been limited use of administrative data in the Family Justice System.	9. The Family Justice Observatory should facilitate the establishment of a family justice research community (the Family Justice Observatory).

Executive summary

A primary objective of the Family Justice System (FJS) is to make authoritative decisions about who cares for children. Such decisions might be in the context of disputes between private parties or between the state and carers. The former, referred to as private family law, typically occurs in the context of relationship breakdown when disputes arise about with whom a child should live. The latter, public family law, is concerned with safeguarding the welfare of children from neglect and abuse. There is widespread agreement that there is currently insufficient generation and use of research evidence within the FJS. This report describes the population data sources (data from which it is possible to draw valid inferences about a defined population) available for England that concern state-authorised decisions about who cares for children and how these data could be used to improve the effectiveness and safety of decision-making and quality of services. This report is part of a wider scoping review that will make recommendations about how a new organisational structure (a 'Family Justice Observatory') might best effect a step-change in the generation and use of research evidence within the FJS. A separate report sets out findings from a case study into locally-held data sources.[3]

We conceptualise the FJS as incorporating the family courts and wider legal and social services. Children's social care (CSC) are central as it is the local authority, through its CSC department, that makes decisions to place children in care (or apply to the court for an order to do so). The Children and Family Court Advisory and Support Service (Cafcass) is also central as that organisation represents children in public family law cases and is often the only safeguarding agency involved with children who go through the private family law system. All of these agencies—Cafcass, the courts and CSC—generate administrative data that can be used to understand who comes into the system and what happens to them before and after. To get a complete picture of such trajectories over time, it is crucial to link these datasets up as the same child may appear in any number of them.

These core FJS datasets have significant limitations that need to be addressed. These include:

- Limited data and follow up on adopted children and private law cases.
- No or limited information on the composition of the household of each child.
- No data on private arrangements made by families without recourse to public services.
- Limited information on the actual services offered by CSC and uptake of those service.
- Ministry of Justice data are not currently available to researchers for approved projects.
- Poor quality identifiers in some cases, and inconsistency in availability of identifiers across datasets, making linkage difficult.
- Changes in coding practices without strict audit of changes.

The lives of children and families are complex with events playing out over time. Whether a particular child comes into contact with the FJS is influenced by a host of factors, such as health, social and economic circumstances. Decisions made in the FJS also have long-lasting consequences in these domains. It is therefore crucial to adopt a broader life course perspective and we therefore consider other data sources:

- Primary and secondary healthcare
- Education
- Employment and income
- Crime and criminal justice
- More detailed case and child data held by local authorities
- Research data resources

These data, especially when linked to the core FJS data, could add particular value to the study of children through the system by enabling researchers to quantify specific risk factors and outcomes that are associated with entry into, and the long-term consequences of, decisions made in the FJS.

Using population data is a complex endeavour. Examining outcomes requires the specification of robust and thorough analysis plans. A range of different studies, including quantitative and qualitative approaches, are needed. Research must also utilise linkage and must address important ethical, legal and governance issues. All of this requires the application of epidemiological and statistical methods with close collaboration with experts in the field.

Going beyond analysis, putting research evidence into practice is crucial. Evidence can be used at the policy and planning level and at the level of individual decisions. The former is relatively uncontentious (though precise details may be fiercely debated) but the latter is particularly controversial. We propose a model of evidence-informed family justice but this requires further work to pilot and test ways of introducing research evidence into case level decision-making that are specific to family justice.

We make a series of recommendations for data providers, research funders, researchers and the Family Justice Observatory. It is hoped that the use of population data will be transformative of the FJS but so far very limited use has been made of them. There are significant challenges to be overcome but doing so is critical to ensure that the best interests of the child are treated as paramount.

Introduction

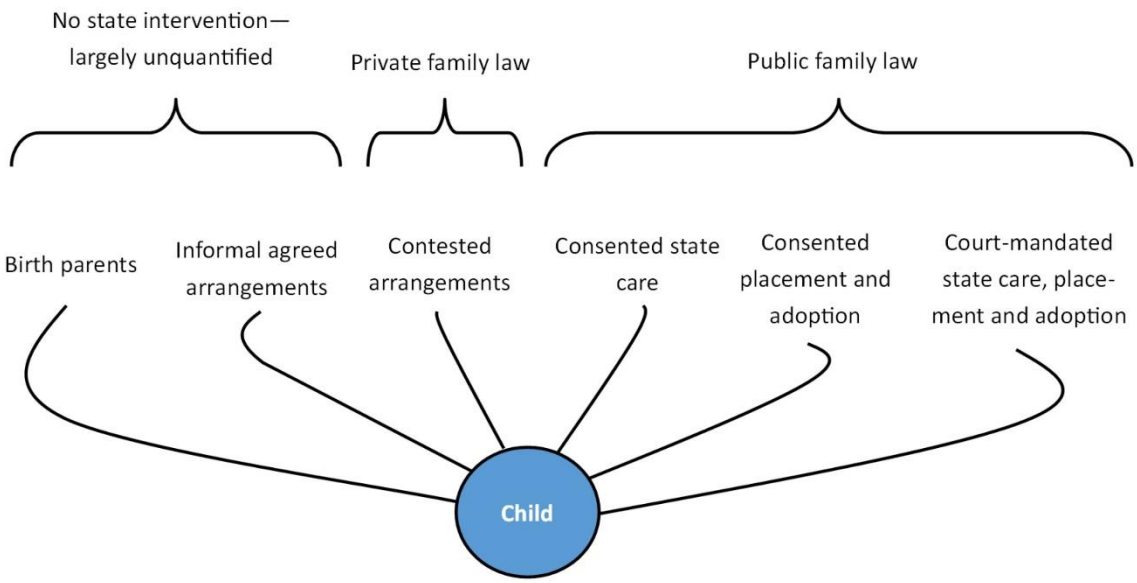
Who cares for children?

For children, who cares for them is a critical, fundamental aspect of their sense of self, wellbeing and development. By who cares for the child, we mean who looks after her (e.g. biological parents, siblings, other family members or strangers) and where she lives. A child's life starts with her birth parents but events may intervene that necessitate alternative care arrangements. The decision about who cares for children depends on whether the state has responsibility for caring for the child, as occurs for example, in circumstances of maltreatment (e.g. neglect or abuse) or high welfare need (e.g. disability or parental incapacity or death). State care is administered through children's social care (CSC) services. If agreed to voluntarily by those with parental responsibility, arrangements remain within CSC. If state care is mandated, the decision will be made by the family court in public family law proceedings. For many children, state care is not needed but the court is involved in the decision about who cares for them because of disputes between carers. The family court is involved in these cases through private family law proceedings.

Possible arrangements for who looks after children are shown in Figure 1. On the left-hand side are the birth parents and other arrangements made informally. Decisions as to who cares for children here are made without state intervention (though there may be state support for example for children in need and through the social security system) and are largely unquantified. Next, private family law (principally section 8 of the Children Act 1989) is used to settle disputes between private parties. These can be resolved by mediation or the family courts. Some children are cared for by the local authority with the parents' consent (section 20 of the Children Act 1989, which also covers cases where there is no person with parental responsibility). Parents can also agree for their children to be adopted through the Adoption and Children Act 2002. Finally, a court can order that a child be received into care and/or adopted, even if the parents do not consent. The principle mechanisms for doing so are care orders under section 31 of the Children Act 1989 and placement and adoption orders under the Adoption and Children Act 2002. The justification for such a decision is that the child is suffering or is likely to suffer significant harm attributable to the care of the parents and that separation of the family unit is a justifiable interference with the family's right to private home and family life.[*]

[*] European Convention on Human Rights, art 8.

Figure 1: Schematic representation of who cares for children



For many children, decisions about who cares for them can be made at multiple points during their childhood, as circumstances change. Some children move between voluntary and mandated state care and some move between state care and private arrangements. From the perspective of the decision makers, who decides what for which types of children varies over time, between local CSC departments and courts, and between individual social workers and judges. We see the Family Justice System (FJS) as encompassing all these different state-authorised decisions. To be able to monitor who is deciding, for whom, and whether children and their families are helped more than harmed, data need to be combined from all state-authorised decisions about children.

This report describes the population data sources available for England about state-authorised decisions about who cares for children and how these data could be used to improve the effectiveness and safety of decision-making and quality of services. These sources included administrative data and research data resources. By administrative data, we mean information about people, businesses and other organisations that is collected by any government department or agency, for delivering their day-to-day services.[1] Research data resources include studies set up to collect new data on large, representative samples of the population followed up over time.

We start in chapter 1 by providing an overview of the services that make decisions about who cares for children: the FJS. In chapter 2, we describe the data resources generated by these services. Chapter 3 considers the child's entire life course holistically and describes other sectors which collect data on outcomes and risk factors for the children and families subject to the FJS. In chapter 4, we discuss how family justice data can be used. We develop a conceptual framework for researching the FJS and we

give examples of the types of studies that can be employed. Legal, ethical and governance issue are covered as are methods of data linkage. Here we also discuss the strengths and limitations of using administrative data and how research evidence can be translated into practice at policy and individual decision-making levels. Each chapter is prefaced with a brief summary and a list of recommendations drawn from it. These recommendations are further detailed in chapter 5 where we recommend next steps to make effective use of population data. A series of appendices (I to IV) provides detail on the core FJS datasets and can be consulted for more detailed information. Finally, Appendix V lists some useful resources pertaining to administrative and population data.

Towards a Family Justice Observatory

This report is part of a larger scoping review which will make recommendations about how a new organisational structure (a 'Family Justice Observatory') might best effect a step-change in the generation and use of research evidence within the FJS. We address the challenges and opportunities of using administrative and population-level data for generating a useful and robust evidence-base in this area. Other components of the scoping review will be published on the study website.[4]

Locally-held local authority data

This report concerns national data—data that cover the whole of England. A companion report by Holmes and others[3] that details findings from a case study on locally-held data will also be available on the study website.[4]

Methods and ethical approval for this project

We used a combination of published literature, interviews with key informants identified through professional networks and a knowledge exchange event with experts held in January 2017[2] to gather data for this report. The study was approved by the Chair of the University College Research Ethics Committee (ref 9775/001) and is covered by University College London Data Protection Registration (ref Z6364106/2016/10/20). Use of data supplied by Cafcass was covered by the UCL ethics approval and was released following approval by the Cafcass Research Governance Committee (ref 180187).

1. What is the Family Justice System?

Summary
This chapter defines the Family Justice System and delineates the primary scope of this report. The Family Justice System is more than just the court. It includes all state-authorised decisions about who cares for children, some of which are made within children's social care. We also discuss the blurred boundaries between private and public family law and why attention must be focused on both.

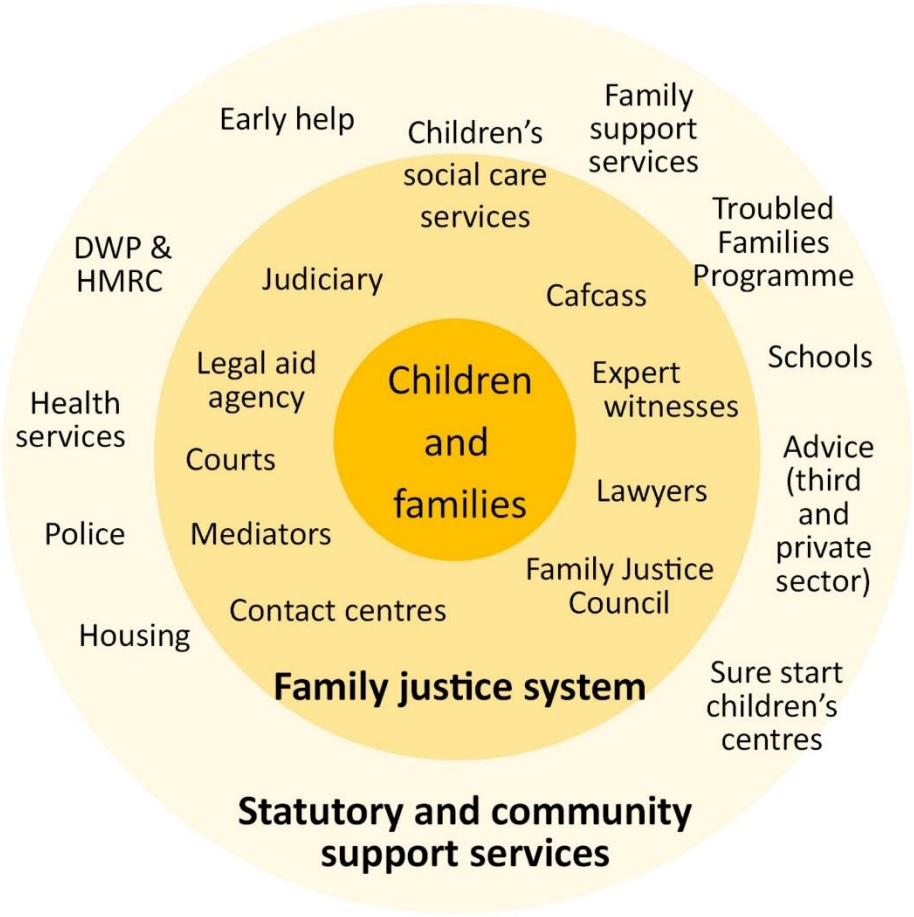
1.1. Deciding who cares for children

Family courts provide independent and authoritative adjudication of disputed decisions concerning who cares for children. Such disputes can arise between private parties, such as divorcing parents, or between the state and private individuals in cases concerning the protection of children from harm and neglect. Courts are crucial in ensuring fairness in the application of the law relating to who cares for children—principally the Children Act 1989 and the Adoption and Children Act 2002. However, the law does not operate in a vacuum and as noted by the Family Justice Review's Interim Report,[5] the courts exist amongst a complex web of other services that provide support or interventions for the welfare of vulnerable children and families.

We therefore conceptualise the Family Justice System (FJS) within this broader context and being more than just the decisions made by the courts (Figure 2). Community and statutory support services influence the number and type of children in contact with the family courts, either because they influence whether children are exposed to risk factors for maltreatment or care need or because of their response to it.[*] Local authorities, for example, are responsible for making decisions to receive children into care, either with parental consent,[†] or by a care order.[‡] The broader services may also influence who comes to court by preventing repeat proceedings. It is estimated that a quarter of mothers who have a care case in the courts experience repeat proceedings, usually for subsequent children (24% of 43,541 women who had an index care proceeding between 2007 and 2014 had a repeat care proceeding within 7 years, with most women having short intervals between: median interval 17 months).[6] The amount and type of support from community and statutory services during and after court will affect the risk of these mothers (and indeed fathers) returning. These broader services—especially children's social care (CSC)—are therefore seen as more than merely ancillary to the FJS but a crucial component of it.

[*] This aspect of service provision is considered in more depth in sections 4.1 to 4.3.
[†] Children Act 1989, s 20.
[‡] Ibid, s 31.

Figure 2: Support and interventions for vulnerable families



Contact centres: A safe and neutral place for children to see non-resident parents/family, sometimes supervised. Families can self-refer or be referred by the courts, social workers or Cafcass.
Early help: Support provided by local authorities to families with needs such as learning difficulties, financial problems, drug misuse and housing.
Expert witnesses: Professionals who provide specialist knowledge/opinion during legal proceedings, for example a child and adolescent psychiatrist or social worker
Family Justice Council: A non-statutory advisory body promoting an interdisciplinary approach to family justice and monitoring the system.
Family Support Services: These services might be family support workers (provided by the local authority, publically funded) or by charities working alongside social workers to support parents and families in their own home, for example through assessing children returning home after being in care or working with parents to improve their parenting or home management skills.
DWP & HMRC: The Department for Work and Pensions and Her Majesty's Revenue & Customs. Claims for various benefits and tax credits can be made through these two bodies.
Mediators: Independent professionals who help parents to work out agreements on issues such as financial support and settling with whom a child will live and spend time.
Sure Start Children's Centres: A publically funded service in the UK to integrate and provide early childhood services aiming to improve outcomes for young children and families and reduce inequalities.
Troubled Families Programme: A programme aimed at shifting support for families with multiple and complex needs from reactive services to preventative services.

Adapted from the Family Justice Review Interim Report.[5]

1.2. The family courts

As noted above, the family courts make a major intervention in families' lives by resolving significant disputes. They adjudicate local authority intervention to protect children, parental disputes over the upbringing of children, adoption, financial support for children after divorce or relationship breakdown, some aspects of domestic violence and petitions for divorce. This report, however, is only concerned with those areas of law that determine who cares for children and therefore does not cover all these aspects.

1.3. Public and private family law

Broadly speaking, there are two types of family cases concerning children: public and private. Public family law cases are brought by local authorities in order to protect children from harm and neglect and include matters such as:

- care orders, which give parental responsibility for the child to a local authority[*]
- supervision orders, which place the child under the supervision of their local authority[†]
- emergency protection orders, which are used to ensure the immediate safety of a child by taking them to (or preventing them being removed from) a place of safety.[‡]

Adoption,[§] whereby parental responsibility is completely transferred to adoptive parents, and special guardianship,[**] which was intended as an alternative to adoption, are also relevant and are often used as child protection mechanisms.

Private family law cases are brought by private individuals and mainly include orders used to settle where and with whom a child should live, with whom they will spend time and other specific disputes (child arrangements orders, prohibited steps orders and specific issue orders).[††]

This report is concerned with both public and private family law for three reasons. Firstly, child welfare is a concern in many private law cases and many are hybrids of public and private family law.[5,7] As Figure 3 shows, private law proceedings are sometimes driven by the local authority and cases may move between the private and public domains during the course of proceedings. Currently there is no research telling us how often, for whom or with what impact this happens. However, a significant

[*] Children Act 1989, s 31.
[†] Ibid.
[‡] Ibid, s 44.
[§] Adoption and Children Act 2002, s 46.
[**] Children Act 1989, s 14A.
[††] Ibid, s 8.

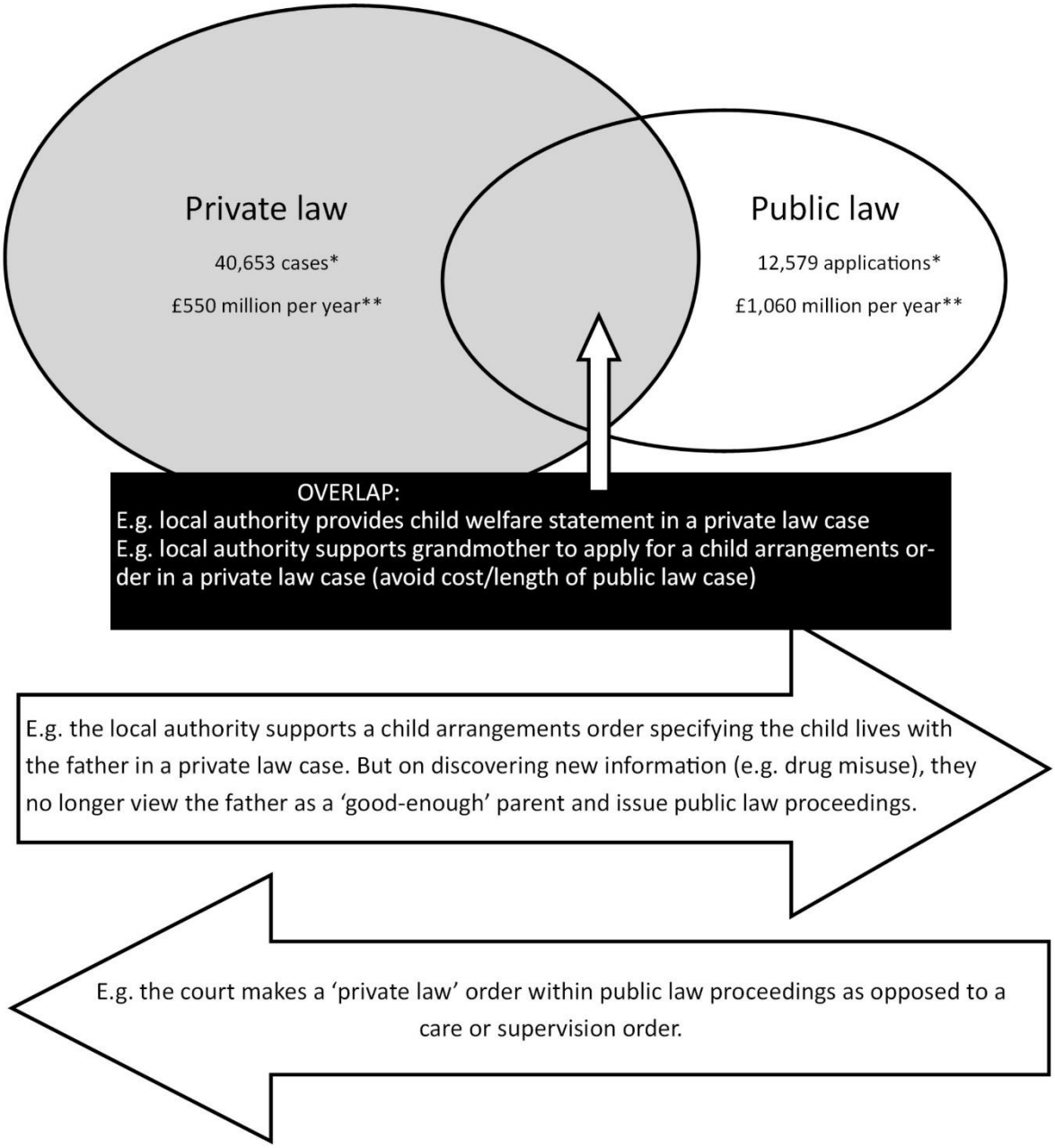
number of 'private law' orders are made in public law cases on the basis that such orders may be a more proportionate and appropriate response to the family's difficulties.*

Secondly, although there might be more obviously high need among children who undergo public family law proceedings, there is research evidence indicating that wellbeing is compromised in children who experience post-separation carer conflict that is resolved by the courts. The two studies that investigated outcomes for children experiencing private family law proceedings estimated that levels of psychological distress among this group are similar to children experiencing care proceedings. Bream and others[8] conducted a small cohort study of families subject to welfare reports. Data on emotional and behavioural problems was collected on 56 children at baseline and 47 at one year follow up. They found that emotional distress was more than twice as high as children in community samples, even more so where domestic violence was an issue in proceedings. Levels of distress were the same as those undergoing care proceedings. In a study of 250 parents and their children who had undergone in-court conciliation, 47% of whom were re-interviewed two years after proceedings, Trinder and colleagues[9] found that high levels of distress are still present two years after the proceedings. Findings such as these should be considered in the overall context of the volume of private family law work: over three times as many children experienced private family law proceedings in 2016/17 compared to public law proceedings (Figure 3).

Thirdly, there is limited research evidence about who comes into contact with the family courts, how often and what happens to them afterwards. This is particularly true for private family law.[10] In summary, we argue that the evidence-base needs to be improved across the whole of family law, private and public, to better understand who cares for children and their short, medium and long-term outcomes as well as to improve the services that make such decisions.

* See Figure 12 in Appendix I, section A1.5.

Figure 3: The relationship between private and public law proceedings



Based on Bainham.[7]
* Cafcass care and private law demand statistics for 2016/17.[11,12]
** Estimated cost to courts, legal aid, Cafcass and local authorities in 2009/10. [13]

1.4. Children's social care

CSC services are crucial because they influence the types and numbers of children and parents that come into contact with the FJS or become looked after via non-judicial routes. They act as the 'gatekeepers' to the care system. In 2016/17, the median rate of care order applications was 12.5 per 10,000 children across 152 local authorities in England.[11] Between local authorities, this rate varied from 0 to 47.1 care applications per 10,000 children.[11] At least some of this variation is likely driven by differences in local decision-making about when to start care proceedings. A study by Bywaters and colleagues[14] which used local authority data from the four UK countries found evidence to support an 'inverse intervention law' whereby a child from a very deprived sub-area of a generally non-deprived local authority was more likely to have a child protection plan or be looked after than a child from a similarly deprived sub-area in a highly deprived local authority.[14] In other words, it seems as if thresholds for child protection plans and out-of-home care are lower in less deprived local authorities.*

Further studies that use data routinely collected from both CSC and the family courts can begin to quantify and understand local variation in rates of service provision and/or child outcomes and inform decisions about whether specific local practices should be more widely promoted and with what likely consequences. Because CSC services are central in deciding who cares for children (whether court-mandated or by non-judicial decisions) we treat them as central to the FJS alongside the courts themselves.

* Variation is considered in depth in sections 4.1 to 4.3.

2. Family justice data

Summary
Here we outline the four core datasets generated by the Family Justice System. These are the datasets generated by the Children and Family Court Advisory and Support Service, the family courts, the Legal Aid Agency and children's social care. As children move through the system, information about them is collected in these different datasets which means that linking them together is crucial to get a complete picture.

Recommendations
1. Routine linkage is needed between the family justice data held by Cafcass, the Ministry of Justice, the Department for Education and local authorities as are mechanisms to authorise re-use of linked data for approved research projects.
2. To follow up children over the life course within datasets, identifiers should be consistently allocated and linked, including for adopted children.
3. To understand who benefits from services, each child record in national administrative data should contain information about the child and family and services offered and received.
4. The court and legal aid data held by the Ministry of Justice should be made available for approved research projects.

2.1. Core datasets

An overview of administrative family justice datasets, as well as their limitations, is given in Table 1. The Children and Family Court Advisory and Support Service (Cafcass) is a public body that represents children in family court cases.[15] It carries out welfare checks and advocates for children in court to safeguard their welfare. It routinely collects case management data on court cases in which it is involved and the individuals to whom those data relate. Cafcass is involved in all public law cases and all private law cases involving children, though Cafcass is involved up to the first hearing only in the majority of these and therefore holds limited data on them. Cafcass only holds data on court processes in its administrative database: no administrative data are held on work by other bodies up to and after court though detailed information is available in case files. Details on the dataset are in Appendix I.

The Ministry of Justice (MoJ) is the ministerial department that oversees Her Majesty's Courts & Tribunals Service, which is responsible for the administration of courts and tribunals in England and Wales.[16,17] It holds data on family court cases in the FamilyMan database. This includes divorce petitions, financial remedies, domestic violence remedies and cases of female genital mutilation as well as disputes about children. There is therefore partial overlap in terms of the population and cases covered by Cafcass and FamilyMan, though the two datasets are held by distinct entities (Cafcass and the MoJ) each with their own data collection, recording and access policies, and the two datasets have different data items and cover different time periods. Further, FamilyMan is not currently available

for research use. This report only considers the cases relating to children. Details on FamilyMan are in Appendix II.

Legal aid, administered by the Legal Aid Agency, an executive agency of the MoJ, helps individuals fund legal advice and/or representation. It is available in all public family law cases. In 2016, public family law comprised 61% of all civil legal aid expenditure (across all areas of law).[18] In private family law, legal aid is available following a means and merits test for mediation and for cases involving domestic violence and child abuse (prior to the implementation of the Legal Aid, Sentencing and Punishment of Offenders Act 2012 in April 2013, means-tested legal aid was available in all private family law cases). The availability of legal aid may influence whether disputes reach court and, if they do, whether the parties are represented. Information on legal aid data, which could be used to enhance the core datasets, is given in Appendix III.

Data from children's social care (CSC) are submitted by local authorities to, and held by, the Department for Education (DfE), the government department responsible for children's services and education.[19] There are two core CSC datasets: children looked after and the children in need. These, respectively, cover all children who enter care, whether judicially mandated or not, and all children referred to CSC, whether any further action is taken or not. Details on these two datasets can be found in Appendix IV. Other data are generated by local authorities and not submitted to the DfE. Some more information on these is given in Table 2 of section 3.2 and are detailed in a report by Holmes and others.[3]

2.2. The complete picture: linking court and social care data

There is a great deal of overlap in terms of the children who appear in the FJS datasets as some children will be children in need, subject to care proceedings and looked after. Some will be looked after without any court intervention and some will be children in need only, with no judicial involvement. This is presented in Figure 4. As each dataset contains different data items, linking them up is crucial in getting a full picture of a child's and family's journey through the FJS. Methods for doing so are described in section 4.6.

Table 1: Overview of Family Justice System data sources

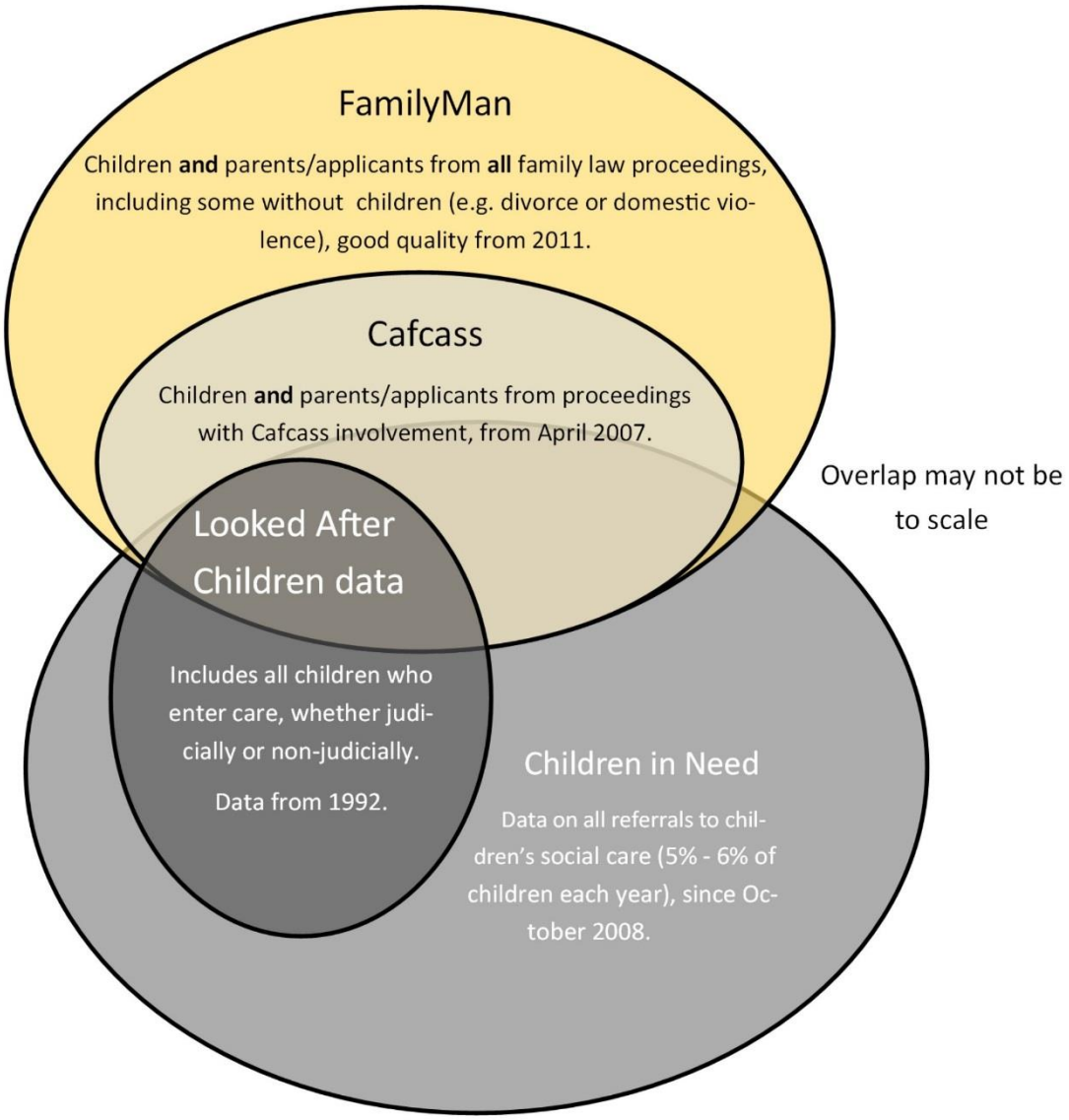
Data resources		Data provider and dates	Brief description	More detail
Family Courts	Children and Family Court Advisory and Support Service (Cafcass)	Cafcass *Since 2007*	Information on family court proceedings in which Cafcass are involved (private and public law) in England and persons involved. Overlaps with FamilyMan but contains different data items.	Appendix I
	FamilyMan	Ministry of Justice *Since 2003 but limited quality before 2011*	Information on all family court proceedings (private and public) and persons involved in England and Wales. Also contains domestic violence, financial settlements and divorce cases (not necessarily involving children).	Appendix II
Legal aid		Legal Aid Agency / Ministry of Justice* *Some from 2001*	Information on civil (which includes public and private family including mediation) and criminal legal aid cases, applications and claimants.	Appendix III
Children's Social Care	Children Looked After	Department for Education *Since 1992*	Information on all provision of out-of-home care in England, including recent care leavers.	Appendix IV
	Child in Need	*Since 2008*	All referrals to children's social care in England and cases opened, including whether the child had a child protection plan.	Appendix IV

Current significant limitations of these datasets include:

- Cafcass are not involved in all adoption cases and where children are adopted, they are assigned new identifiers. It is therefore not possible to follow them up post-adoption.
- There are limited data on private law cases.[10] Cafcass for example are only involved in private law matters up to the first hearing in approximately 70% of cases; they do not therefore record the final legal orders in these cases.
- The precise composition of the household is largely unknown. It is possible to link mothers and children in the Cafcass data but the father is unknown in a large proportion of cases.[6,20]
- Informal arrangements (i.e. agreements made where there is no state intervention) are not covered by any administrative dataset as there is no obligation for private parties to report their living arrangements.
- Even with these datasets linked together, they contain limited information about the problems that families who enter the care system face or the services that they are offered and take up. Additional locally-held local authority data are detailed by Holmes and others.[3]

* The Legal Aid Agency is an executive agency of the Ministry of Justice.

Figure 4: Overlap between Family Justice System datasets



3. '*The child's present and future life as a human being*':* a wider lens

Summary
The lives of children and families are complex with events playing out over time. Whether a particular child comes into contact with the Family Justice System is influenced by a host of factors not directly relevant to the court, such as health, social and economic circumstances. Decisions made in the Family Justice System also have long-lasting consequences. Here we consider data from wider services and research resources that can aid our understanding of how children come into the Family Justice System and the impact that decisions can have.

Recommendation
1. Data providers and the Family Justice Observatory should advocate for linkage, including considering how the Digital Economy Act 2017 and existing frameworks can enable cross-sectoral linkages by trusted third parties.

3.1. Trajectories through the system

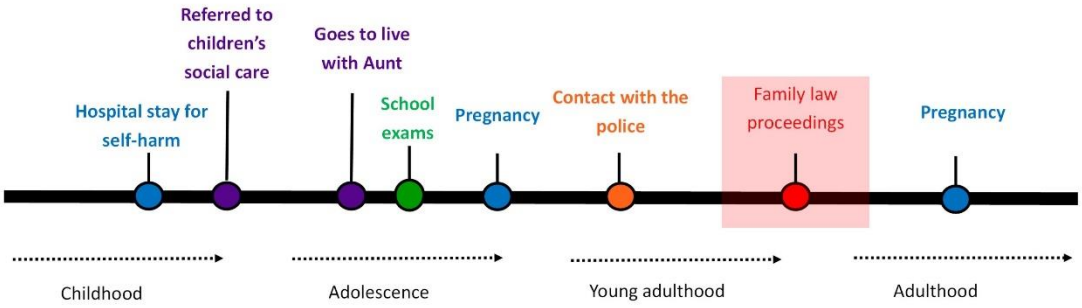
For many it is a long journey to the family court, with problems that start at birth (or even in utero) for the young people who eventually become parents themselves. Contact with the family courts can be understood as one event in the complex, inter-twined life trajectories of parents and children. Parents and children have contacts with a range of public services, including GPs, hospitals and mental health services, schools, social care services, housing and benefits agencies, the police and of course the family courts themselves (Figure 5). At best, each of these contacts is an opportunity for public services to intervene positively and support families and their children who are (or are at risk of) suffering family break-down, family conflict and neglect or high welfare need. At worst, these contacts might exacerbate problems within families and be an inefficient use of public funds.

By using data from these services, who gets into the system, when and what happens to them before, during and after contact with the FJS can be understood. For example, as noted in section 1.3, children undergoing private family law proceedings experience high levels of distress[8] that can persist for years after proceedings.[9] It is therefore essential to consider the short, medium and long-term trajectories of children into, through and out of the FJS and how decisions can affect these. This is particularly so given the court's duty to treat the child's welfare as a paramount consideration.† Linkage between

* In *In Re G (Education: Religious Upbringing)* [2012] EWCA Civ 1233 [26] the Court of Appeal, per Munby LJ, affirmed that the concept of welfare, which is the court's paramount consideration when determining any question concerning the upbringing of a child (Children Act 1989, s 1), 'extends to and embraces everything that relates to the child's development as a human being and to the child's present and future life as a human being'. His Lordship went further to state that the court must have consideration to 'the child's welfare now, throughout the remainder of the child's minority and into and through adulthood'.
† Children Act 1989, s 1.

datasets from different sectors, as well as linkage of individuals and family members within datasets over time, is necessary. This will require support and advocacy from data providers and the Family Justice Observatory.

Figure 5: The journey in and out of family courts



3.2. What data are available?

A range of administrative and research data sources are relevant. Administrative data sources include data from health care services, schools, the Department for Work and Pensions, Her Majesty's Revenue & Customs, the police and criminal courts. These datasets include children who may be at risk of adversity and coming into contact with the FJS. For example hospitalisation data can be used to identify children who are admitted to hospital with adversity-related injuries.[21] Similarly, data from general practice can be used to examine how concerns about neglect are recorded.[22,23] Locally-held data are also important. All local authorities collect data on children in their populations, including those who are in need or looked after. These include additional data items that do not form part of the national returns and are used to varying degrees for both strategic and operational purposes at local level. This issue is considered in more depth in a separate report by Holmes and others.[3]

Research data resources include birth cohorts and panel surveys. Birth cohorts collect rich data about children's social, economic, parental and household circumstances that provide essential context for understanding the impact of decisions made in the courts for different groups of children. These and panel surveys, which follow up groups of individuals, though not recruited at birth, use face-to-face, telephone and/or computer assisted interviewing techniques to collect data on the same participants at set time points. Researchers can also collect biological samples and physical measurements.

Table 2 gives an overview of these data sources. We have restricted our description to administrative datasets from health, education, employment and income (including social security benefits) and

crime and criminal justice, as these are most likely to include important risk factors and outcomes for children and families in contact with the FJS. We also include three birth cohorts and three panel surveys which are of particular relevance because of the ages of those included in their samples and the types of data collected. There are birth cohorts that recruited participants 40 or more years ago (the 1946,[24] 1958[25] and 1970[26] British birth cohorts) but the core datasets covered in this report are available only more recently which means that the participants in these cohorts will already be adults before they can be identified within the administrative data. For this reason we have de-prioritised these studies in this report.

Table 2: Overview of population-level datasets from wider services

	Data resource	Dates	Data provider	Brief description
Administrative data resources	**Health**			
	Hospital Episode Statistics (HES)	Since 1997 (admissions), 2003 (outpatients), 2007 (Accident & Emergency).	NHS Digital*	All admissions, outpatient appointments and Accident & Emergency visits at NHS hospitals in England and private hospital interactions funded by the NHS. Contains demographic, clinical and organisational information.[27]
	Clinical Records Interactive Search	Since 2007	South London & Maudsley NHS Foundation Trust	Mental health data from the South London & Maudsley NHS Foundation Trust linked with HES, ONS mortality data, Lambeth general practice records and the National Pupil Database.[28,29]
	The Clinical Practice Research Datalink (CPRD)	Since 1987	CPRD	GP records, organisational information about GP practices, and patient demographics on a sub-set (7%) of the population in England, Wales and Scotland who are registered with an NHS GP.[30,31]
	The Health Improvement Network	Since 2002	In Practice Systems and IMS Health	General practice records of about 6% of the population. Contains data on patient demographics, medical records and consultations.[32,33] There is some overlap with CPRS in terms of the population covered.

	Data resource	Dates	Data provider	Brief description
Administrative data resources	Maternity and Children's Data Set	In development	NHS Digital	Pregnancy booking, screening and tests, admissions, labour/foetus outcomes, postpartum details, baby screening, sexual health and infectious diseases (still under development).[34]
	Mental Health Services Data Set	Since April 2016	NHS Digital	Demographics, referrals to NHS mental health services, care planning, encounters with mental healthcare professionals, inpatient stays, diagnosis, interventions, outcome measures and discharges relevant to mental health, for children and young people. Not yet used for research purposes.[35]
	Personal Demographic Service	Since 2011	NHS Digital	A master index of patient records containing name, address, date of birth and NHS number for patients receiving treatment in an NHS setting in England, Wales and the Isle of Man. Does not contain any clinical information but can be used to facilitate linkage between datasets.[36,37]
	Health datasets can be used to understand healthcare utilisation among children and families in contact with the FJS. As these datasets also cover individuals not in care or otherwise in contact with the FJS, they can also be used to construct comparator groups. The health datasets can be linked to each other using identifiers such as names, addresses, date of birth and NHS number and they can be linked with the FJS datasets using the demographic identifiers. For example, Wijlaars et al[38] are linking HES to Cafcass data to examine patterns of healthcare utilisation of children who go on to experience care proceedings. The Personal Demographic Service can be used to facilitate linkage between FJS and health datasets, for example in the study by Wijlaars et al.[38]			
	Education			
	National Pupil Database	Since 2002	Department for Education	Attainment and progression at each key stage, eligibility for free school meals, special educational needs, absences and exclusions for all pupils in state schools in England. State school age is the term after the child turns 5 but data are also collected on early years provision, which includes some children as young as 2.[39]

	Data resource	Dates	Data provider	Brief description
Administrative data resources	The National Pupil Database is a vital source of educational data and can be used to examine achievement of children in and out of the FJS such as the study by Sebba et al[40] which examined educational outcomes of children in care, in need and in the general population.			
	Employment and income			
	Work and pensions longitudinal study	Since 2004	Department for Work and Pensions (DWP)	All claims for benefits and periods of employment (from HMRC tax records) for people who have ever claimed certain benefits in England. Does not include self-assessment tax records.[41,42]
	Her Majesty's Revenue and Customs (HMRC) data	Since 2003 (child benefit and tax credits). Until April 2013, earnings data in were only held on people earning above the tax threshold; from that date, all employees are included.	HMRC	Administrative data on child benefit, tax credits, income tax, national insurance contributions, earnings and others.[43-47]
	Data from the Department for Work and Pensions and HMRC could be used as a rich source of socioeconomic data on work and benefits receipt. HMRC data were used in a study by the Department for Education into graduate earnings[46] and in April 2017, the Department opened a consultation on the use of linked DWP, HMRC and education data to understand household income and education.[48]			
	Crime and criminal justice			
	Police National Computer	Since 1995	Ministry of Justice	All reprimands, warnings, cautions and convictions in Great Britain. Covers only recordable offences (those which attract a custodial sentence plus other defined offences). Has individuals' demographic data (including address and postcode) and the arrest/summons number. The database is subject to weeding rules meaning that an individual's full offending history may not be captured.[49,50]
	Police force crime records	Varies between areas	Local police forces	Offences recorded by police forces, geographical data, demographics of

	Data resource	Dates	Data provider	Brief description
				offenders and victims and charge outcomes. Each force has its own dataset and is its own data provider.[51] Quality of data may vary by police force.
Administrative data sources	LIBRA and CREST (criminal courts)	Since 1995 (crown court), 2008 (magistrates' courts)	Ministry of Justice	Data on activity of the criminal courts and defendants including names, dates of birth, gender, addresses, arrest/summons number and final outcome. Includes offences with no police involvement, e.g. by government departments or private prosecutions. Prior to 2008, police forces reported magistrates' courts proceedings.[52,53]
	Understanding which individuals come into contact with the criminal justice system could help identify groups at-risk of coming into contact with the FJS. Doing so could assist in putting in place support for families to avoid this and avoid re-offending and the consequences of offending.			
	Locally-held local authority data			
	Local authorities collect, hold and link data that are not submitted in the national returns as part of the children looked after or child in need datasets. Details vary by local authority but may include data such as social worker information and services received by children and families. See the separate report by Holmes et al[3] for more information. These data can be used to augment national data. For example, data on social worker changes were used alongside national data in the development of a Stability Index by the Children's Commissioner[54] and Bywaters et al obtained local authority data to examine deprivation and looked after children.[14]			
Research data resources	**Birth cohorts**			
	Avon Longitudinal Study of Parents and Children (ALSPAC, also known as 'Children of the 90s')	Participants recruited in 1991-2	University of Bristol	Environmental, social and genetic factors from >14,000 pregnant women, their children and partners over 20y. Participants recruited from the Bristol area of England.[55]

	Data resource	Dates	Data provider	Brief description
Research data resources	Born in Bradford	Participants recruited in 2007-10	Bradford Teaching Hospitals NHS Foundation Trust	Socio-economic characteristics, ethnicity and family trees, lifestyle factors, environmental risk factors and physical and mental health from 12,453 women with 13,776 pregnancies (recruited at ~28 weeks pregnant) and 3448 of their partners. Participants recruited from a single hospital in Bradford England.[56]
	Millennium Cohort Study	Participants recruited in 2000-2	Centre for Longitudinal Studies†	Parenting, childcare, school choice, child behaviour and cognitive development, child and parental health, parents' employment and education, income and poverty, housing, neighbourhood and residential mobility, social capital and ethnicity for a representative sample of children 19,519 born in the UK.[57]
	The cohort studies contain rich, detailed data about the lives and characteristics of children and families that are usually not available in administrative data. A number of studies have used ALSPAC, for example, to study various social and economic risk factors associated with child maltreatment[58-61] and the Millennium Cohort Study has been used to explore factors predictive of receiving social work support.[62] Participants in cohorts can be linked to administrative data. These studies are prone to participants' dropping out, an issue which is more likely to affect disadvantaged groups, including those in contact with the Family Justice System.			
	Panel surveys			
	Longitudinal Study of Young People in England (also known as 'Next Steps')	First cohort recruited in 2004 and followed up to 2010. Second recruited in 2013 and will be followed up until age 20	Department for Education†	Experiences of children and young people and e.g. views on local areas, community cohesion, social activities, risky-behaviours, crime/anti-social behaviours, health and future aspirations. Collected through yearly interviews with the same children recruited at school in England in year 9 (aged 13-14y).[63]
	Labour Force Survey	Since 1992 but data useable from 2002	Office for National Statistics†	Detailed employment circumstances collected by interview five times each quarter from randomly selected households identified from the Postcode Address File. There are 40,000 households (100,000 people) in the study at any one time, with 20% of the household sample being replaced each quarter.[64]

	Data resource	Dates	Data provider	Brief description
Research data resources	Understanding Society: the UK Household Longitudinal Survey	Since 2010 with some data going back further as it incorporates British Household Panel Survey which ran for 25y and finished in 2008	Institute of Social and Economic Research, University of Essex†	Information relating to family life, education, finance, employment, health and wellbeing collected annually from members of 40,000 selected households in the UK. Includes individuals aged 16 and older as respondents; children aged 10-15 complete a shorter youth questionnaire.[65]
	Panel surveys are similar to cohort studies in that they provide rich detail that is not captured in administrative data. The Next Steps survey, for example, has shown that social class, gender, ethnicity, step-family status and special education needs are all significant predictors of social service contact[66] and that those with contact with social services had poorer educational attainment.[67] The Labour Force Survey includes data on adults and children in households linked using a unique family identifier. It may therefore be possible to use it to link members of families together.			

* NHS Digital is a statutory body formally known as the Health and Social Care Information Centre.

† These datasets can be accessed via the UK Data Service.[68]

4. Using family justice data

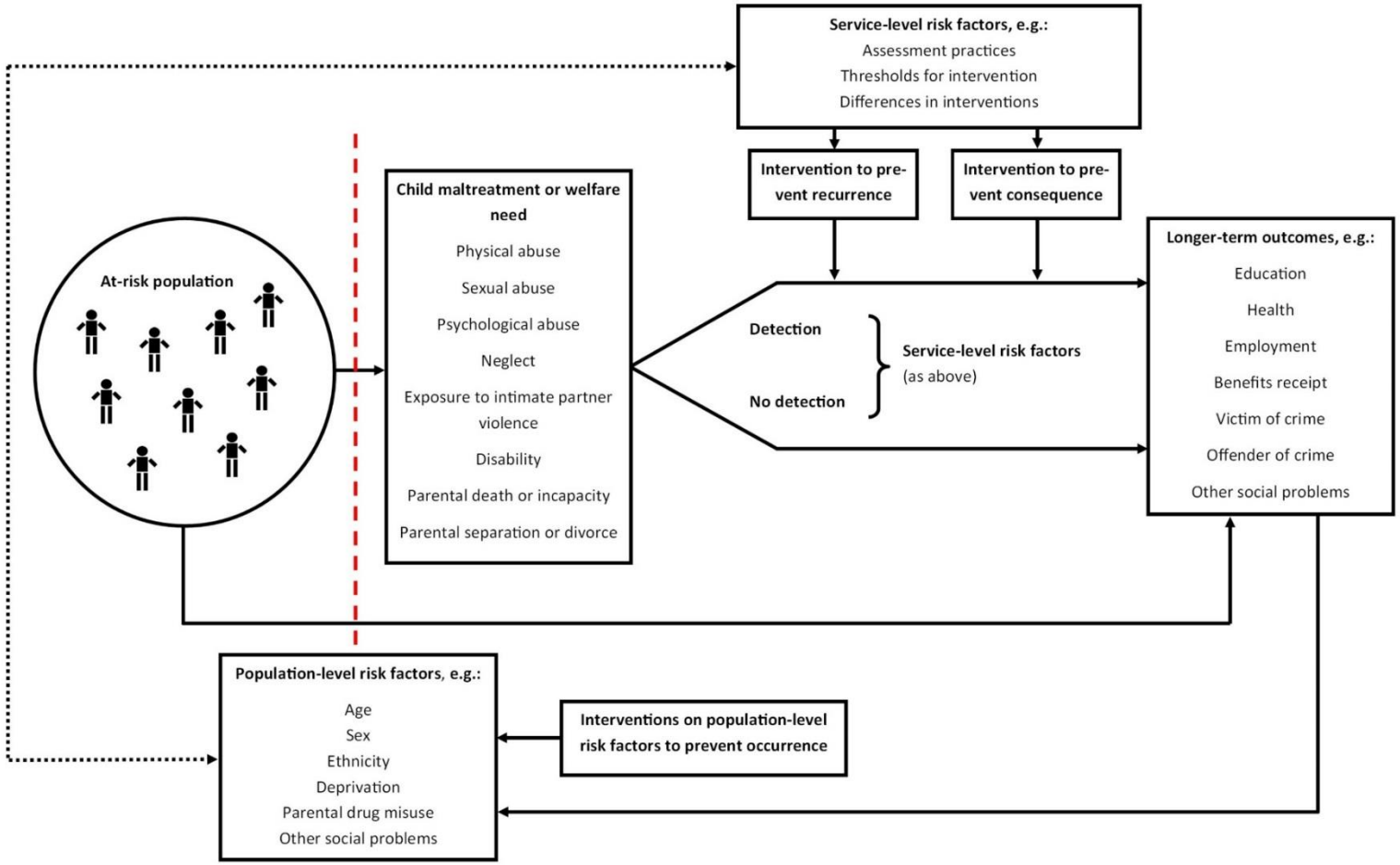
Summary
Previous chapters outlined the core family justice datasets and other data resources that can aid understanding of who comes into the system and the long-term consequences of decisions about who cares for children. In this chapter we explain how these data can be used to do this. Analysing variation is key but this must be done appropriately. The types of study—ranging in scope from less to more complex—are outlined. We also consider the strengths and limitations of using administrative data, ethical, legal and governance issues that the use of these data raises and how data can be linked. Finally, we discuss how data and research evidence can be translated into practice and some areas of contention about doing so that need to be resolved.

Recommendations
1. It is necessary to improve data providers' understanding of the needs of researchers and their capacity to evaluate and report their methods for data processing.
2. Research funders and the Family Justice Observatory should support capacity building and interdisciplinary collaboration between those with appropriate skills and knowledge.
3. An examination of how research evidence can be put into practice is required, particularly at the level of individual decisions, taking account of the fact that using research evidence is separate from using real-time identifiable client data.

4.1. A conceptual framework

Understanding how administrative data can be used necessitates the explication of a conceptual framework that links an at-risk population (i.e., all children) with exposure to maltreatment or high welfare need and longer-term outcomes. Such a framework is depicted in Figure 6. This framework emphasises that various risk factors will influence whether a child experiences maltreatment or welfare need, whether such is detected, whether any interventions are offered and, if so, what (e.g. support for parents, section 20 accommodation or an application for a care order in the courts). Clearly this is a complex framework, any aspect of which could be subject to research and further theoretical development. Administrative data make a contribution by providing information on those in the population who receive a given service. By linking up datasets, especially with those outwith the Family Justice System (FJS; see chapter 3), it is possible to draw a picture of the experience of the population (or a subset of it) and assess how outcomes vary along any particular axis of interest (e.g., geography, time or personal characteristics). Variation in outcomes can therefore be exploited using social epidemiological methods to understand the system, to highlight areas for further study and to identify possible targets for improvement.

Figure 6: A conceptual framework linking child maltreatment to longer-term consequences



Adapted from Macmillan and others.[69] The thick dashed line represents the fact population-level risk factors will determine whether a child experiences maltreatment or high welfare need. Those that do will either be detected or not and those that are detected may or may not be offered any number of interventions. Whether and what interventions are offered is determined by a set of service-level risk factors and these may interact in complex ways with the population-level risk factors (e.g. services responding differently to different groups of children). All children will experience longer-term outcomes that will also feedback into the population.

4.2. Determinants of variation in practice and outcomes

Variation can help identify how different practices result in different outcomes in similar populations. For example, the fact that the rate of children going into care in Kent in 2015/16 (45 per 10,000) is over three times as high as that in Essex (14 per 10,000) raises critical questions about what Kent is doing differently to Essex.[*] The reason for examining variation is not to construct league tables to examine 'performance'. This is usually problematic in all sectors for statistical reasons (explained below) and in family justice especially because we have limited information on what services *should* be doing. For example, we have limited knowledge as to why some local authorities have higher rates of care proceedings than others, i.e., whether high rates of care reflect higher levels of need or whether services are operating differently and thereby keeping children out of the care system. Even clear benchmarks, such as the 26 week time limit for care proceedings,[†] may be problematic as complex cases require longer and measuring performance on the basis of compliance with the time limit would be unfair to areas with higher rates of these cases.

Variation is a multifaceted construct.[72] The following all determine variation in outcomes: chance; the frequency, severity and type of maltreatment or welfare need ('case mix'); how maltreatment and welfare need are detected, assessed and intervened on ('policy and practice');[69,73] and data quality.[72] Differences in policy and practice are usually of interest as it is these which we usually wish to improve outcomes for children and families. However, the other factors—chance, case mix and data quality—mean that variation is not an indicator *per se* of quality. In order to draw conclusions about how services affect outcomes, a bald description of variation is insufficient and account must be taken of the other factors.

- Chance: There will always be a degree of variation which is due purely to chance, especially for smaller populations.[‡] There are statistical methods that researchers can use to quantify the extent to which variation is greater than expected by chance.
- Case mix: The frequency, severity and type of maltreatment or welfare need varies between different units under study (for example, local authorities). This is a difficult question that

[*] We calculated these rates from the published data on the number of children who started to be looked after by local authority in 2015/16[70] and population denominator data from the Office for National Statistics.[71]

[†] Children Act 1989, s 32(1)(a)(ii).

[‡] To illustrate: if a local authority only had 1,000 children living in it, and it took just 10 children into care, its rate would be 100 per 10,000 children. If it instead only took 5 children into care, its rate would be 50 per 10,000 children—a huge difference in the rates induced by a difference of just 5 children. Compare this to a local authority that has 15,000 children living in it. If it took 50 children into care, its rate would be 33 per 10,000. If it took 45 into care (also a difference of 5 children), the rate drops to 30 per 10,000. Thus, it can be seen that the rates for the larger authority are much more robust to small changes in absolute numbers whereas the smaller area's rates are much more volatile.

requires research as not all risk factors are known and methods for dealing with them are complex. The frequency, severity and type of maltreatment or need experienced will itself be determined by the population-level risk factors depicted in Figure 6.*

- Policy and practice: Differences in policies and practices around detection, assessment and intervention of child maltreatment and need also determine variation. These are the service-level risk factors in Figure 6. This highlights the need to understand the services that generate the data. It should also be emphasised that there may be complex interactions between the service-level and population-level risk factors in that some services may respond to different groups of the population in different ways.[14]
- Data quality: Finally (but not least importantly), any variation observed may be due to problems with data quality and/or variation in data quality, factors which themselves may be influenced the service-level risk factors. For example, the recording of 'need' (e.g. allegations of domestic abuse) in the Cafcass dataset is subject to time pressures on caseworkers, who are responsible for filling in this variable. Therefore, it may be omitted in busier times and this will affect apparent trends. Conducting qualitative research alongside the quantitative study can help uncover potential problems like these.

The purpose of taking these matters into account is to make fairer comparisons of the outcome between different groups. It is impossible using administrative data to know whether a particular factor *causes* some outcome† but this does not mean that such analyses are pointless. Examining variation is a starting point to further investigation. Well-designed and executed studies can provide reliable and valid evidence as to the phenomenon under study and can lead to further investigations and interventions to improve the system.[76]

Analyses that account for these factors abound in other fields, such as health. For example, Mayer and colleagues[77] demonstrate how case-mix-adjusted funnel plots can be used to account for change and case mix when evaluating variation among hospitals in 30-day mortality rates following radical cystectomy (removal of the bladder). They show how adjusting for case-mix factors (including patient gender, age and deprivation as well as process-of-care variables such as waiting times, number of beds available and the ratio of nurses to occupied beds) changes conclusions as to how well a hospital is

* For example in Kent there is a high proportion of unaccompanied asylum seeking children (UASC) who began to be looked after in 2015/16 (26 per 10,000 children living in Kent, which was the third highest rate of UASC going into care that year).[70] Other relevant factors include demographic factors such as the child's age, gender and ethnicity,[74,75] parental characteristics,[6] household socioeconomic circumstances,[61] neighbourhood or local authority deprivation[14] and how these factors interact.[14]

† This is because there will always be unmeasured influences that contribute to the variation.[72]

performing relative to others, especially when using a type of graph called a funnel plot.* In this particular case, they showed that no hospital could be confidently said to have a worse than average performance than other hospitals, despite the fact that there was huge variation in the crude mortality rates.

We stress that these methodologies require specialist epidemiological and statistical skills. Similarly, given the law's complexity and dynamism, research using family justice data requires collaboration with those who have knowledge and expertise of the system.[78] Research will therefore require close interdisciplinary collaboration and capacity building.[79,80]

4.3. Evaluating variation

It is possible to examine variation by conducting studies that employ increasingly sophisticated methods of analysis. These include:

- Counting service contacts
- Describing service contacts
- Following up children over time in individual datasets
- Following up children over time across multiple, linked datasets
- Analyses based on linked administrative datasets or linked family members over time
- Analyses based on administrative datasets linked to research data sources

Illustrative examples of studies employing these methods are given in Table 3.

Using linked datasets adds value to the use of single, unlinked datasets. Without linkage there is a risk that any inferences made will be incomplete and misleading. As can be seen from Table 3, using linked data can help overcome significant limitations associated with using one dataset alone. It is also crucial to note that data are needed not only on children and families in contact with the FJS, but those not in contact with it. This is because understanding risk factors for contact with the FJS, and outcomes after, a valid comparison group that can provide a 'baseline'. As the FJS datasets only capture data on children and families who come into contact with it, obtaining comparison groups is only possible through using linked data from other services.

* A funnel plot is a graph that plots as a dot, for each unit (e.g. hospitals or local authorities), the rate of the outcome (e.g. mortality rate or the rate of children entering care) against the number of individuals in the population of each unit. This firstly enables the researcher to visually account for random chance. By creating a series of funnel plots, each with statistical adjustments, the researchers can also visually assess how much variation is accounted for by case-mix factors. Examples are given in the article.[77]

Table 3: Examples of studies using family justice data

Nature of study	Study	Methods	Main findings	Limitations
Counting family court contacts using FamilyMan	Family Court Statistics Quarterly[81]	The Ministry of Justice routinely publishes data from FamilyMan on, for example, the number of public and private family law cases and the number of children involved in them.	In 2016, there were 18,952 public law cases (31,375 individual children involved) and 48,244 private law cases (72,836 individual children involved). These figures have risen every year since 2011.	Provides a limited snapshot of activity. Analysing crude numbers (i.e., not rates) does not take account of underlying demographic changes.
Describing legal outcomes using Cafcass data	Harwin et al[82]	Description of legal outcomes in public law cases using Cafcass data. Analysis of proportions of different orders by year with the aim of examining trends in special guardianship orders (SGOs) in particular. Denominator was all orders granted.	The proportion of cases ending with an SGO has risen over the years, as has the use of SGOs concurrently with supervision orders. More infants are subject to SGOs with time suggesting age is a risk factor.	Only able to provide limited information on child characteristics (age and gender) and only has information on court outcomes. The study was unable to provide detailed information on which children are given SGOs or longer-term consequences. Linking to other datasets such as health and education would help answer these questions.
Following up children over time in individual datasets using children looked after data	Mc Grath-Lone et al[74]	Using the children looked after data, analysis of the proportion of all children in England who entered care at least once. Denominator was all children born in certain year bands (e.g. 1992-1994).	3.3% of all children born between 1992 and 1994 entered care at least once before the age of 18; this proportion is rising with time. Children with black, mixed and other ethnicity more likely to enter care.	Limited data on services actually received before, during or after care. The child in need dataset, although itself limited, would identify local authority involvement since 2008/9. The consequences in terms of impact on other services that the rise in the proportion of children who enter care represents remains to be quantified through linkage to datasets from other services.

Nature of study	Study	Methods	Main findings	Limitations
Following up children over time across multiple, linked datasets using Cafcass data and Hospital Episode Statistics	Wijlaars et al[38]	Using the Hospital Episode Statistics linked to Cafcass, the authors will investigate healthcare utilisation in those who do and do not enter court-mandated care, as well as healthcare and maternity service use among mothers.	Study in progress. The results will inform whether children who enter care have higher healthcare use than those in the general population who do not. This study will therefore assess longer-term outcomes.	Not possible to follow up children who have been adopted as they are assigned new identifiers on adoption. Further work is required to determine whether and how adopted children can be followed up in administrative data.
Analyses based on linked datasets or linked family members over time using Cafcass data	Broadhurst et al[6]	Mothers and children were linked across time to calculate the proportion of mothers who had recurrent care proceedings (denominator being all those who had at least set of care proceeding).	24% of mothers returned to court within 7 years; younger mothers were more likely to do so. For most, episodes happened within a short space of time.	Used only Cafcass data and therefore misses children who entered care via section 20—these data would be in the children looked after dataset. It also misses mothers who had care proceedings before 2007 (as data were not available then) so it was not possible to fully capture trajectories for all mothers.
Analyses based on linked datasets or linked family members over time using children looked after data, child in need data and the National Pupil Database	Sebba et al[40]	Analysis of linked education, children looked after and children in need data. The population studies was all children at Key Stage 4 (age 15) in 2012/13.	Children not in need or care performed the best at Key Stage 4. Early care entry was associated with better results vs later entry. School and placement changes associated with poor outcomes.	There were no detailed data on foster or residential carers and no details of school and placement practices that may have contributed to outcomes. Such data are currently only collected at a local level.
Analyses based on administrative data linked to research data sources using a randomised-controlled trial and Hospital Episode Statistics	Robling et al[83]	A randomised-controlled trial to examine the efficacy of the Family Nurse Partnership. Outcomes included healthcare utilisation and pregnancy (from GP records and Hospital Episode Statistics).	The study found no benefit of enrolling in the Family Nurse Partnership in terms of these and other outcomes.	Outcomes were only measured in the short-term. More follow up (which is possible using the linked administrative data) is therefore required to quantify the long-term efficacy of the programme.

4.4. The strengths and limitations of using linked cross-sectoral administrative data

Administrative data, linked up across different sectors has a number of unique strengths:[84]

- **Scale**: These data cover either the entire population or a large and representative sample of it. The results are thus generalisable to the service that generated the data and the population served. Studies that collect data for the purposes of research often have small sample sizes, which can make it difficult to detect subtle but important differences. These studies also sometimes have problems relating to the types of people who are likely to take part in them. If a certain group is systematically under-represented in a study, then its results may be biased. Similarly, if they do take part in these studies, children who experience maltreatment or high welfare need are at particular risk of 'dropping out' of the study. Using population-level data can overcome these problems.
- **Real-world data**: The data reflect real-world practice and are directly meaningful as practice and research are using the same measures.
- **Trajectories over time**: Individuals can be linked over time (e.g. where there are multiple court appearances or episodes of care for the same child) and researchers can therefore follow up children and families over time and through different services, before and after court.
- **Knowing the whole population**: Using data on the whole population enables researchers to quantify rates as well as counts and can therefore quantify the proportion of the population affected. This gives a more reliable picture of practice as crude numbers will usually change with the underlying population notwithstanding any changes in practice. Calculating rates over time can also provide a cumulative estimate of the factor under study within the population (e.g., the proportion of children who ever enter care in their lifetimes).[75]
- **Comparison groups**: The comprehensive capture of the whole population over time makes it possible to design studies that include comparison groups. For example, it is possible to compare rates of care among local authorities and changes over time. It is also possible to examine whether outcomes vary by factors of interest such as levels of deprivation. Importantly, data on children and families not subject to the FJS can be examined which makes it possible to assess the association between FJS factors with short, medium and long-term outcomes such as education or health (see sections 4.1 to 4.3).
- **Efficiency**: These datasets already exist so researchers are free from the time and costs of primary data collection and there is no data collection burden on participants. There may however be costs associated with data access and storage (see section 4.5) and other costs related to the research project such as staff salaries.

The limitations of administrative data stem from the fact that they are collected for some operational reason (such as monitoring demand or for reimbursement) and are then made available to researchers for secondary analysis. Limitations include:[84]

- **Measuring contacts recorded by the service, not occurrence of events or experiences**: For example, self-report studies indicate that child maltreatment affects 4-20% of children each year but only a fraction of these are reported to children's social care services.[73,85] Access to services by those affected may vary by individual circumstances, between communities and over time. Analyses of administrative data need to take account of variation in service access (including waiting times for services), recognition and recording of maltreatment and welfare need and intervention.
- **Lack of detail about individual circumstances:** Administrative data contain quality information on aspects most relevant to the service (e.g., date and reason for court attendance) but have less information on broader circumstances such as a family's social and economic circumstances. They can also have limited detail on severity and complexity. The looked after children dataset, for example, only requires identification of one category of need (e.g., abuse or neglect, disability and so on).[86] Because of this, it is difficult to adjust for all possible risk factors and sometimes proxy indicators have to be used.
- **Data quality:** Administrative data may be subject to error introduced by misrecording (e.g. due to mistyping names and numbers) and missing data. If these errors occur in identifiers, this can adversely affect linkage and such errors can disproportionately affect certain groups such as unusual name structures. The processes in place to generate the data require further in-depth study to fully understand them, their context and any potential biases or inaccuracies caused by the ways in which data are collected, processed and stored.
- **Information about the data:** Any information about the data that is produced by the data provider may not be sufficient for researchers because it is, for example, produced to aid understanding of standard statistical outputs. Collaboration and in-depth study with the data providers could be conducted to produce this kind of information.
- **Changes in data recorded:** Researchers must be cognisant of changes to services and policy such as changes in the grading of exams or the availability of court orders.

4.5. Ethical, legal and governance issues

Using administrative data raises important ethical, legal and governance issues as all record-level data are potentially identifiable until anonymised.* As such, data users must ensure that the use of administrative data is ethical, lawful and safe. A number of protocols exist to ensure that proposed research is ethical, feasible and of public benefit, that researchers are adequately trained to process data securely, that data are held in systems that meet defined security specifications and that research outputs cannot potentially identify individuals. These aspects of data use are referred to as the 'Five Safes': Safe Projects, Safe People, Safe Data, Safe Settings and Safe Outputs.[88] These are explained in Table 4 and in a video on the Administrative Data Research Network (ADRN) website.[89]

Table 4: The Five Safes

Safe Projects	Is this use of data appropriate?	Legal, moral and ethical considerations. *Should* the data be used for this project?
Safe People	Can the researchers be trusted to use the data in an appropriate manner?	Data users and data providers should meet the appropriate standards of behaviour.
Safe Data	Is there a disclosure risk in the data themselves?	Whether data are identifiable informs the access environment, for example whether the data can be released into the public domain or should be stored and processed in a secure computing environment.
Safe Settings	Does the access facility limit unauthorised use?	Whether the setting is safe depends on the physical environment (such as physical security at a research data centre) and procedural safeguards (e.g. auditing).
Safe Outputs	Are the statistical results non-disclosive?	Is there a risk of identification in the published outputs? This could be directly (e.g. directly naming an individual) or secondarily (e.g. by deducing the identity of an individual from the output, possibly in combination with other information). Output statistical disclosure control are applied to mitigate risk of reidentification.

Adapted from Desai and others.[88]

* Anonymisation is 'the process of turning data into a form which does not identify individuals and where identification is not likely to take place.'[87] The Information Commissioner's Office has produce a code of practice that organisations can follow to ensure that anonymisation is effective.[87]

Most research requires independent ethical review, which is often carried out by a university research ethics committee. In addition, the data providers have their own data access procedures. Such procedures vary from provider to provider, but typically involve specification of the research purpose, the public benefit, how data will be used, which variables are required and details of ethics approval including consideration of data security and the appropriateness and qualifications of the researchers. The data provider must also have a legal power to share the data.[*] The process of approvals may require significant negotiation with the data providers and can take very significant amounts that must be planned for in advance.

Once access is approved, data will usually be transferred to the researchers securely. Data should be stored on a system that meets strict information governance requirements such as a safe haven—a system that offers a 'walled garden' approach to data access and storage and has strict access and statistical disclosure control protocols in place. Statistical disclosure control minimises the risk of outputs being disclosive meaning that individuals or organisations cannot be directly identified or identified by combining the published output with other information.

The foregoing has focused on safe and ethical use of data from the perspective of minimising the risk of harm should data be shared. It should also be borne in mind that there are risks associated with data not being shared for valid research purposes as less research means there is less evidence available for society to make decisions.[90] Rigid rules, lengthy application processes and costs[†] are all barriers to effective and timely research and so a balance between enabling quality research and protecting individuals' confidentiality is needed.[91,92]

4.6. Linkage methods

Linkage within and across datasets is challenging and time-consuming and can be done in two ways. Deterministic methods for linkage require an exact or almost-exact agreement on a specific set of identifiers such as date of birth, sex and postcode to 'match' an individual between datasets.[93] Deterministic methods are useful when records have unique identifiers which are accurate and

[*] Currently these are found in various statutes but this will be rationalised to some extent by the Digital Economy Act 2017, s 64. Once brought into force, the Act will give a general power to public authorities to share data for the purposes of research provided that a person's identity is not specified in the data and that it is not reasonably likely that a person's identity could be deduced (either from the data shared or in combination with other information)—in other words, that the data are anonymised. This power will not be available to bodies whose functions relate solely to health or solely to adult social care (or solely to both): s 73(2).

[†] At present, there are no financial costs imposed by the FJS data providers for access to their datasets. Costs in other domains, such as health, however, can be prohibitively expensive for researchers[91,92] and there are of course costs associated with researcher salaries and other overlays whatever data are used.

complete but are prone to 'missed matches' where records from the same individual fail to link because of inaccuracies or incompleteness of the identifiers. Probabilistic methods allow records from the same individual to be linked in the presence of recording errors and/or without the same unique identifier in both datasets.[93] There is a danger of creating 'false matches' between records if the probabilistic rules are too relaxed. Linkage error can directly affect the results by for example under- or over-estimating associations.[93]

Whichever method is used, high quality identifiers are needed. Table 5 shows some typical identifiers used for linking datasets such as names, sex and postcode and their availability across the core FJS datasets. It is important to check the quality and completeness of identifiers for any linkage project as this will directly affect the quality the linkage. Not all of the identifiers in Table 5 are complete: see the appendices for detail. The process of carrying out linkage can help identify problems with the data that can be fed back to the data provider, thereby improving data quality in the future.

Table 5: Availability of identifiers across the family justice datasets

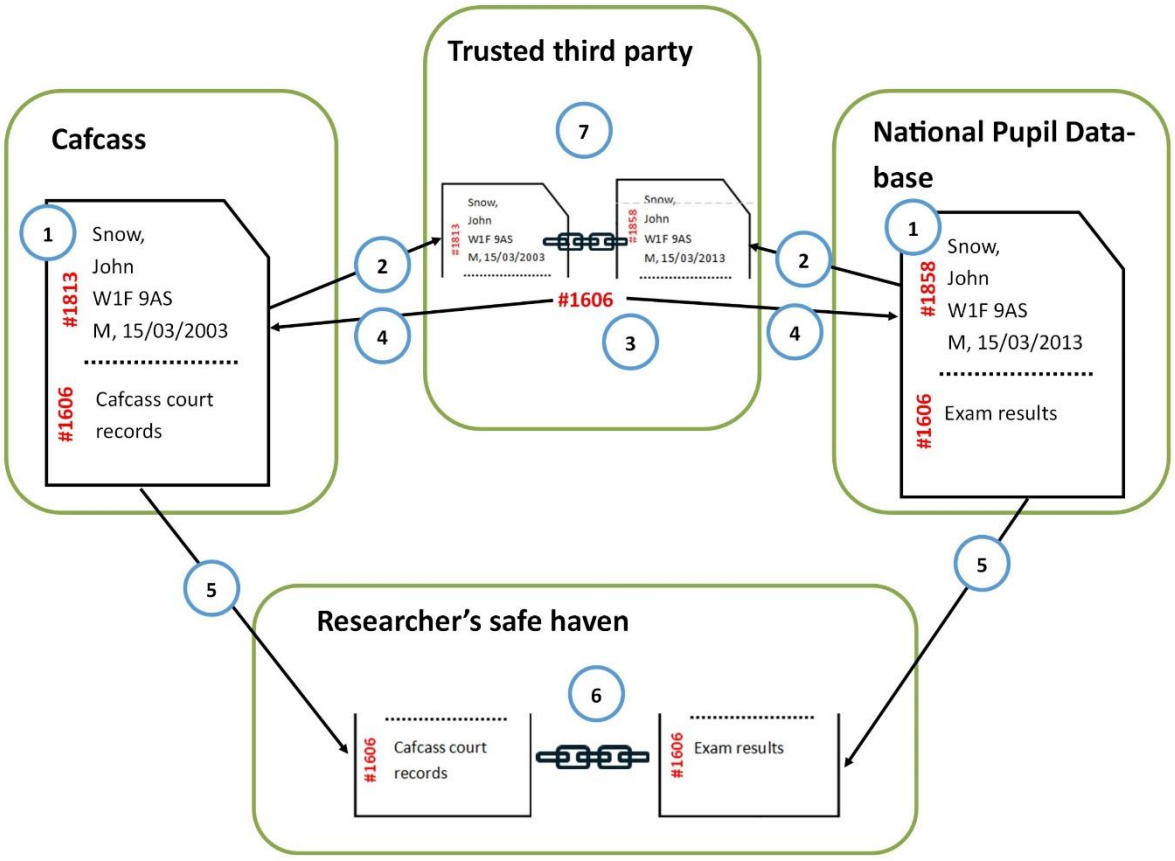
Identifier	Cafcass	FamilyMan	Legal aid	Children looked after	Child in need
Child's first name	✓	✓	✓	X	X
Child's last name	✓	✓	✓	X	X
Child's sex	✓	✓	✓	✓	✓
Child's date of birth	✓	✓	✓	✓	✓
Child's postcode	✓	✓*	✓*	✓*	X
Child's local authority	✓	✓	✓*	✓	✓
Child's ethnicity	✓*	X	✓	✓*	✓*
Mother's first name	✓	✓	✓	X	X
Mother's last name	✓	✓	✓	X	X
Mother's date of birth	✓	✓	✓	X	X
Mother's postcode	✓	✓	✓*	X	X
Unique Pupil Number	X	X	X	✓†	✓†
Mother-child link?	Yes – research[6]	Yes – untested	Yes – untested	No	No
Father-child link?	Yes – research planned[94]‡	Yes – untested	Yes – untested	No	No

* High proportion missing / invalid (may vary over time). † For children who have state entered school or pre-school only. ‡ Previous work found difficulties linking fathers and children within the Cafcass data.[6,20]

Methods for linking locally-held local authority data are outline in the local area case study report by Holmes and others.[3]

The need for identifiers means that requirements for safe and ethical processing are more stringent than for single datasets. To address this, linkage is performed by the data provider or a Trusted Third Party (TTP) whereas the analysis is performed by the researchers on a pseudonymised dataset (the principle of separation). A hypothetical example of how this can be done is shown in Figure 7.

Figure 7: Hypothetical linkage between Cafcass data and the National Pupil Database



Explanation

1. Within each dataset, each record is assigned an artificial ID (here the numbers #1813 in Cafcass and #1858 in The National Pupil Database).
2. The identifiers, including the artificial ID, are securely transferred to a TTP, such as the Office for National Statistics.
3. The TTP executes a linkage algorithm and assigns a linkage key to the records (here #1606). This key signifies that record #1813 in Cafcass belongs to the same person as record #1858 in the education data.
4. The linkage key and artificial ID are transferred back to the data providers and the data provider assigns the linkage key to the main records.
5. The Cafcass and exam data are transferred to the researchers, without any of the identifiers but with the linkage key.
6. The researchers link the main records using the linkage key.
7. Finally, the identifiers held by the TTP are destroyed.

An video demonstrating linkage can be viewed on the ADRN website.[95]

Data providers, linkers and researchers should ensure that sufficient information is reported to properly appraise the linkage and the evidence generated from the study. The Guidance for Information about Linking Data Sets, which covers steps from initial data provision, through to linkage and publication, should be referred to.[96] Researchers need such details so that they can take account of linkage error and other data quality issues, such as missing data, in the analyses to reduce biases. Linkage of data also helps to address errors. Hence more collaboration between data providers, linkers and researchers can improve data quality.

Given the time-consuming and therefore costly nature of carrying out linkage, the re-use of linked datasets is called for. To enable more efficient use of linked data, the ADRN has recently changed its policy of 'link for a specific purpose and then destroy' to link and retain by approved TTPs.[97]

4.7. Putting data into practice: three modes of action

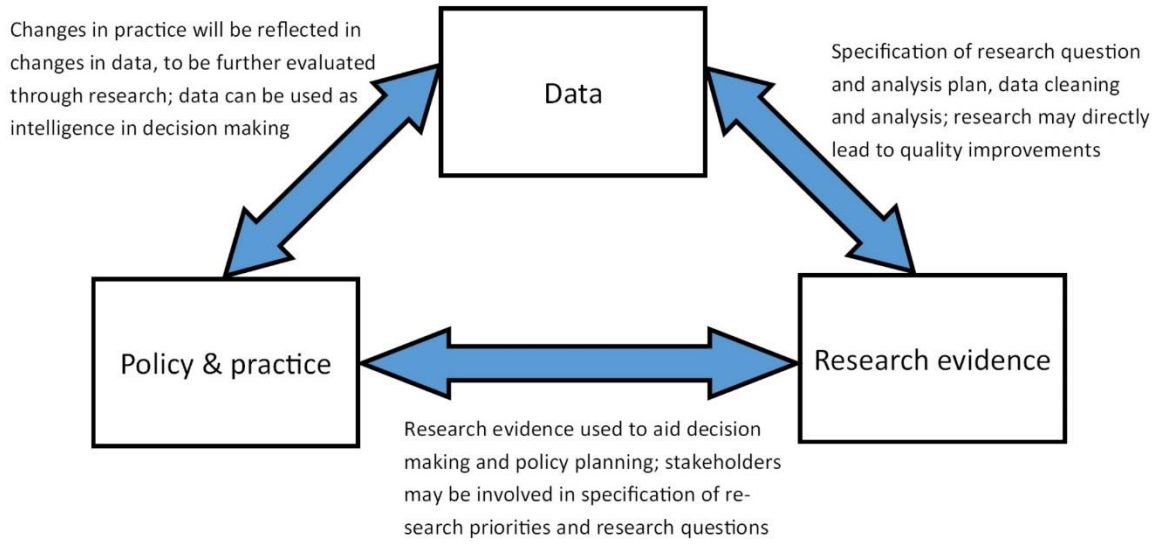
In this section we overview three modes of action by which data can be put into practice. Full development of these modes is beyond the scope of this report but an understanding of them is fundamental if effective use of data is to be made within the FJS. The modes are:

- Using research evidence from previous studies of populations (or groups) of children:
    - To guide policy development, legislation and service planning, including targeting certain groups
    - To inform decisions about who cares for children
- Using real time data from services to identify and approach specific children or families

Failing to distinguish between these different uses of research evidence and data could undermine public trust in the sharing of anonymised data for approved research purposes.

Here we also draw a distinction between data and research evidence. Data are raw, unprocessed pieces of information. Research evidence by contrast is the result of a carefully pre-defined analysis that is reported in sufficient detail that others can assess the validity and generalisability of the findings and, if necessary, replicate the study. Using data for research can lead to improvements in data collection and quality and using research evidence for policy and practice can generate new questions and uses of the data. This multidirectional framework is represented in Figure 8.

Figure 8: The interplay between data, research evidence and practice



4.7.1. An intermission: the place of evidence based on populations in legal scholarship

Before turning to the three modes of action, we pause to consider the position of population data and research evidence within the broader field of legal scholarship, of which McCrudden[78] provides a taxonomy. In the first place, he notes that four broad legal research agendas operate in Britain. These focus on (1) legal concepts and reasoning (e.g. the content, consistency and scope of legal rules); (2) the meaning and validity of law (e.g. jurisprudential questions around the nature of law); (3) ethical and political dimensions of policy delivered through legal mechanisms; and (4) the effect of law on human behaviour, attitudes and actions. Traditional legal scholarship resorts to doctrinal analysis that considers law to be a closed system of rules and, since the 1950s at least, to philosophical dimensions such as the relationship between law and morality. Legal research that incorporates more explicitly 'external' disciplines are divided by McCrudden[78] into socio-legal research, critical legal studies and law-and-economics. It is the socio-legal and law-and-economics approaches, the former of which is most closely associated with sociology, anthropology, psychology and other related disciplines, that are empirical in nature.[*]

Whereas the kinds of methods necessary to study the FJS using population data do not fit neatly into this taxonomy, the name of the particular discipline from which the evidence is generated is irrelevant. What matters is whether the research evidence is methodologically robust, valid and relevant to the

[*] The use of empirical methods in the study of law and legal processes is sometimes referred to in a broad sense as empirical legal studies/scholarship.

problem at hand. A plurality of different types of evidence, which includes that about the system itself but also about phenomena such as the impact of divorce on a child's mental well-being, is likely to be of benefit to the FJS. However, given the nature of population data, epidemiological and advanced statistical techniques are required. In particular, a life course social epidemiological framework is likely to be of most benefit as this emphasises the broader determinants of outcomes (the causes of outcomes as well as the 'causes of the causes')[98,99] and how these operate over life.[100] The need to view the FJS through this lens was highlighted in section 3.1. Such approaches must operate complementarily with traditional legal scholarship.[78,101] Research using family justice data should involve those with the appropriate scientific expertise as well as those with knowledge of the law and legal processes and those who understand policy and the services and practice that generate the data.

We now turn to the three modes by which data and evidence can be put into practice.

4.7.2. Research evidence to guide policy development, legislation and service planning

The use of research evidence from population-based studies for policy, planning and formulation of legislation is perhaps the least controversial aspect of putting data to use in practice, though precise details of implementation may be subject to fierce debate. For example, knowing how many families are likely to come into contact with the FJS in given areas in the next year is clearly useful—essential—for resource allocation and service planning. Using research to plan and evaluate novel interventions in order to determine whether they should be rolled out or decommissioned is also clearly of benefit as economic scarcity demands that the most cost-effective (not necessarily the cheapest) programmes and policies are in place. Research may also frame the policy debate at governmental and parliamentary levels. For example, Mc Grath-Lone and others[75] showed that 3.3% of all children born between 1992 and 1994 in England entered out of home care at least before the age of 18. This is broadly similar to Denmark (2.8% by age 18),[102] lower than in Manitoba, Canada (9.4% by age 12)[103] and the USA (5.9% by age 18)[104] and higher than in Western Australia (1.5% by age 12).[103] The use of this kind of evidence could help frame a policy debate by prompting critical policy questions as to what the 'right' level is for placement of children in out of home care, for whom, when and with what consequences.

4.7.3. Research evidence to inform decisions about who cares for children

It is indisputable that the ultimate goal of family proceedings is to make the right decision for the child. What constitutes the 'right decision' varies from case to case but will always be guided by section 1 of

the Children Act 1989—the principle that in all cases concerning the upbringing of a child, the child's welfare shall be the court's paramount consideration—as well as other legal provisions, in particular the article 8[*] rights of the family members. In determining such a question (whether in private or public family law), the court will hear evidence as to the truthfulness of claims of the parties and what is in the child's best interests. It is then charged with the task of making an appropriate order by assessing the welfare of the child, the proportionality of the care plan (if a care order case) and whether to make any order at all.[†]

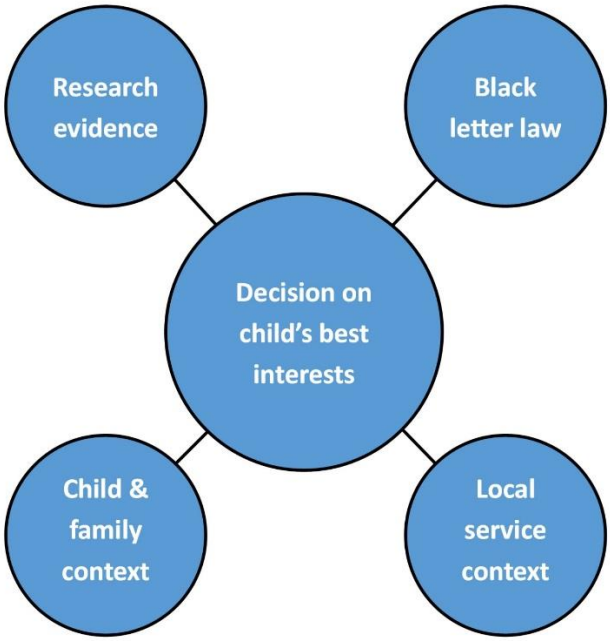
These questions are always subject to uncertainty, especially determinations as to what is in the child's interests not only now but in the future.[‡] This is a question that takes on especial prominence in cases such as adoption that will by their nature certainly affect the child for the rest of his or her life. We suggest that research evidence can be used to inform answers to these questions. Research evidence must not supplant judicial discretion and independence but it should align itself with it. We offer a tentative framework, based on decades of learning and experience in evidence-based medicine, where issues of uncertainty and prediction also arise,[105,106] which we refer to as evidence-informed family justice (EIFJ). This framework, depicted in Figure 9, posits that decisions are not made in a vacuum but instead in a complex and dynamic situation where facts are imperfectly known and predictions probabilistic. Our discussion here focuses on the judge as the decision maker but is not confined to the court. Other decision makers, such as local authorities, Cafcass officers and legal representatives or experts can and do make use of research in a similar fashion, though it is not yet clear who is best-placed to introduce research evidence into the court. It is likely that research evidence can be used by different actors at different stages of the child's trajectory through the system.

EIFJ describes a process of using research evidence as an aid to decision-making. Research evidence is one component of the decision which is also predicated on the law, the specific circumstances of the child and family and local services. For example, a question concerning who should look after a child will be determined with reference to, among other things, the law; the child's needs (e.g., disability, behaviour or mental health problems) as well as the values, preferences, ability and willingness of relatives to look after that child (child and family context); and the availability of non-related care

[*] European Convention on Human Rights, art 8. This establishes the right to private home and family life.
[†] Children Act 1989, s 1(5) states that the court should consider whether making an order is better than making no order at all. In such a case an 'order of no order' can be made.
[‡] *In Re G* [2012] EWCA Civ 1233. It should also be noted that the Children Act 1989, s 31(3A), requires the court, when making a care order, to consider the permanence provisions of the local authority care plan. This will be amended by the Children and Social Work Act 2017, s 8, to stipulate that the court must also consider the impact on the child of any harm, current *and future* needs and the way in which the plan would meet those needs.

settings (local service context). Research evidence could be used, if it were available, to determine which of these placement options might, in the long run, be best for this child based on the child's characteristics (such as age, gender and underlying chronic health conditions).

Figure 9: Evidence-informed family justice



Adjudication in the FJS is different to the process of diagnosis, prognostication and treatment in medicine. Adjudication is concerned with a court exercising power to determine the rights, statuses and liabilities of individuals vis-à-vis one another and the state in a binding fashion. Medical treatment does not hold the same legal and constitutional significance.[*] The EIFJ framework therefore requires further work to elucidate how it might operate in practice including resolution of legal and practical issues that we summarise here but which it is far beyond the scope of this report to tackle.

Legal issues

Two major legal issues that the use of research evidence in the FJS raises relate firstly to the status of research evidence within the law of evidence and secondly to whether the use of research evidence affects judicial independence.

[*] An exception being where there is a conflict between the medical professionals and the family as to what course of action is in the best interests of the child (e.g. *Great Ormond Street Hospital v Gard* [2017] EWHC 1909). Cases such as these raise important legal issues that, if no agreement is reached, will be resolved by the court. Research evidence will of course still be relevant in these disputes.

Rules around the admissibility of evidence are complex. The law of evidence 'regulates the process of proof of facts for the purposes of legal proceedings.'[107] It is concerned with reaching a decision that is right in law given the facts about which there is often challenge or uncertainty. The legal status of research evidence, which has been the subject of much debate in the Australian family legal system,[108-110] needs to be established. Issues to be resolved include whether research evidence is something of which the judge can allow to be considered (in which case conclusions from research are taken as established facts) or whether the argument based on research evidence must be proven on the balance of probabilities. If a court takes research evidence into account, can or should it do so without submissions from the parties (i.e., can a judge use research evidence, implicitly or explicitly, in reaching a decision and formulating his or her judgment)? If research evidence is used in the determination of the case, then in order to ensure a fair hearing the parties must have the opportunity to make their own submissions on the validity of that evidence. Here the adversarial nature of proceedings may adversely limit the impact of research evidence because of the opposing views of the parties as to its validity.

The status of research evidence within the law of evidence currently appears ambiguous. One judicial respondent in the FJO national stakeholder consultation[80] reported that the court would deal with research when presented as part of the submissions in the court papers. Research evidence can also be cited by expert witnesses. For example in *In Re H*[*] the psychologist's report, as quoted in Wall LJ's judgment, makes repeated references to research. There are also examples of research evidence being used as accepted doctrine. Another FJO stakeholder consultation[80] respondent stated that the Sturge and Glaser[111] report on contact and domestic violence can be relied upon and not challenged. This indicates it is held in particularly high esteem by the judiciary and may be due to its being commissioned for, and approved by, the Court of Appeal in *In Re L*.[†] Thus in *In Re H* the Court of Appeal criticised the trial judge for, among other things, failing adequately to deal with the recommendations in Sturge and Glaser.[111]

Treating research evidence as legal doctrine, however, is contrary to the notions of EIFJ, a central aspect of which is to use the best-available evidence—which may change over time as new evidence is reported. Using evidence as doctrine is in fact pseudoscientific.[109] No single study is ever determinative of an issue. Even systematic reviews of studies must be treated with caution as must of course guidance formulated on the basis of such reviews. Although the doctrine of judicial precedent is that like cases are treated alike, the question is what is in *this* child's interests as defined in law.

[*] *In Re H (A Child) (Contact: Domestic Violence* [2005] EWCA Civ 1404.
[†] *In Re L (A Child) (Contact: Domestic Violence); In Re V (A Child); In Re M (A Child); In Re H (Children)* [2001] Fam 260.

Each case will therefore fall to be determined on its own facts and the same study may have different value in similar but different cases. This requires a careful and critical consideration of the research evidence, its methods, its limitations, its generalisability and its implications even in matters that appear to be dictated by common sense.

The second legal issue is on judicial independence—that judges act independently of government and the parties before them and must decide cases according to the rule of law. These are fundamental constitutional principles that must be upheld to ensure the integrity of the law and legal processes.* However, we would reject the notion that the use of research evidence would undermine independence. Firstly, similar concerns have been raised in the context of judicial feedback but where such feedback exits, independence has not been considered compromised.[113] As Figure 9 makes clear, a process of EIFJ would use research evidence as one piece of the jig-saw in decision-making: research evidence alone never dictates a course of action. The discretion of the judge as arbiter of fact is maintained and all other relevant factors are still taken into account. Rather than undermine judicial independence and trust in the courts, the explication of how research evidence is used and ensuring that this is done so fairly and justly may in fact enhance the legitimacy of adjudication and, most importantly, the quality of decisions for the children subject to them.

Practical issues

Our model of EIFJ also raises practical issues that need to be resolved: principally questions of access to evidence; training and research literacy; court culture; trust in research; and a need to improve the evidence base. These issues were explored in the FJO stakeholder consultation.[80]

4.7.4. Using real time data to identify and approach specific children or families

Using real time data to identify and approach specific families is also controversial. For example, a local authority might use an algorithm to identify the risk to a child and those in need of safeguarding. It could then pre-emptively approach the child and family to offer preventive services. Alternatively, an authority might construct a data system (such as a screening tool) that does not rely on such algorithms but is used for similar pre-emptive purposes.

* The independence of the judiciary, the separation of powers and the rule of law are complex constitutional principles which it is far beyond the scope of this report to review (see, e.g., Tomkins).[112] It is sufficient to note here that the judiciary are considered autonomous and must not have their discretion unduly fettered.

Use of data for pre-emptive provision of services based on predicted risk of adverse outcomes can be beneficial but can also cause harm. There are concerns about the strength of the evidence underpinning the predictions and its applicability to the individual. Where algorithms are used there are concerns about transparency—understanding how and why a given person was identified can be hard to explain—and there are privacy concerns. As has been shown in the context of 'predictive policing',[114] unreliable algorithms can result in false positives, leading to people being approached who did not need services and thereby feeling accused and stigmatised. Unreliable algorithms can also result in false negatives meaning that people who could have benefited were missed. Privacy concerns revolve around the fact that identifiable data, including from other services, would be used without informed consent to target services. This may be an unwelcome intrusion, especially if not on the basis of statutory powers and without independent oversight. As those requiring local authority support are generally of poorer socioeconomic circumstances,[14] such interference might also be discriminatory. There may therefore be practical, ethical and legal objections to the use of data in this way.

5. Next steps

Summary

In this final chapter we draw together the recommendations that arose out of the previous chapters. We examine how each might be achieved and comment on their feasibility. Finally, we add a recommendation around the need for a community of researchers, drawing on a knowledge exchange event held in January 2017.

Recommendations

In previous chapters, we have outlined what is required for the effective use of family justice data. Here these recommendations are reproduced with detail, include the audience(s) for each, on the pages that follow.

1. Routine linkage is needed between the family justice data held by Cafcass, the Ministry of Justice, the Department for Education and local authorities as are mechanisms to authorise re-use of linked data for approved research projects.
2. To follow up children over the life course within datasets, identifiers should be consistently allocated and linked, including for adopted children.
3. To understand who benefits from services, each child record in national administrative data should contain information about the child and family and services offered and received.
4. The court and legal aid data held by the Ministry of Justice should be made available for approved research projects.
5. Data providers and the Family Justice Observatory should advocate for linkage, including considering how the Digital Economy Act 2017 and existing frameworks can enable cross-sectoral linkages by trusted third parties.
6. It is necessary to improve data providers' understanding of the needs of researchers and their capacity to evaluate and report their methods for data processing.
7. Research funders and the Family Justice Observatory should support capacity building and interdisciplinary collaboration between those with appropriate skills and knowledge.
8. An examination of how research evidence can be put into practice is required, particularly at the level of individual decisions, taking account of the fact that using research evidence is separate from using real-time identifiable client data.
9. The Family Justice Observatory should facilitate the establishment of a family justice research community.

Recommendation 1: Routine linkage is needed between the family justice data held by Cafcass, the Ministry of Justice, the Department for Education and local authorities as are mechanisms to authorise re-use of linked data for approved research projects.

Audience: Data providers (Cafcass, Ministry of Justice, Department for Education, local authorities).

Objective: As we have emphasised throughout this report, linkage is central to the most effective use of family justice data. To get a full understanding of the Family Justice System (FJS), the core family justice datasets—Cafcass, FamilyMan, children looked after (CLA) and children in need (CiN)—should be linked up. The necessity for doing so is obvious from Figure 4 in section 2.2, which shows the overlap between the datasets in terms of the populations covered by them. Data from the courts is only one piece of the jigsaw and what happens to children in children's social care is equally important. Linking these datasets would also enable cross-validating fields and filling in missing data, thereby significantly improving the quality of all of the datasets.

Immediate steps: A number of technical issues must be addressed to improve linkage accuracy and the utility of the data:

- Ensure that the family justice datasets have a common set of identifiers that are routinely collected. These could include, for example, full names, addresses, postcodes, date of birth, gender and ethnicity.
- Improving the quality and completeness of date of birth and postcode in FamilyMan. These are not always complete or recorded in the correct format meaning that linkages requiring these variables are likely to be incomplete.
- Improve the capture of client-level geographical data. In most of the datasets there are concerns about the quality of postcode data. This should be investigated in order to improve the quality of linkages.
- Within the CLA and CiN datasets, different local authorities use different identity numbers for the same children. A common and persistent identifier for each child should be used nationally. This could be a UPN assigned at birth rather than on entry to school.
- Efforts should be made to improve the recording of ethnicity, which is currently subject to high rates of missingness.
- Some of the linkages proposed in this report (such as between Cafcass, CLA and CiN) have not been tested. Methodological work should therefore be carried out and linkage accuracy should be evaluated and reported.
- Data providers should work with researchers to produce comprehensive data resource profiles with research-standard metadata. Doing so will further identify strengths and limitations of the datasets and will inform what improvements can be made.

Establishing linkages between the family justice datasets, and improving the quality of the identifiers as outlined above, should be achievable within the short term. The Ministry of Justice (MoJ) is already carrying out a linkage project internally (see Appendix II for details) but has relatively low match rates of between 60% and 75% and work is therefore required to improve this.

Particularly regarding linkage of family justice data to education data, children in the CLA and CiN datasets do not have a Unique Pupil Number (UPN) recorded unless one has been assigned once the child enters school. This includes only some pre-school children and it excludes those of any age who do not enter state school. It is therefore not possible to fully link children's episodes of need and care with education data in the National Pupil Database. Assigning a UPN at birth would potentially solve this.

Having independent researchers scrutinise and analyse the data not only means that the benefit of collecting data is more rapidly obtained as more research is conducted, it can also lead to identification of data quality issues that might not have been uncovered otherwise. Feedback loops between researchers and data providers can thereby improve the quality of the data for future use.

Finally, given that linking datasets requires significant effort and resources (section 4.6) the re-use of linked datasets by researchers should be enabled for efficiency, cost reduction and sharing of expertise.

Recommendation 2: To follow up children over the life course within datasets, identifiers should be consistently allocated and linked, including for adopted children.

Audience: Data providers (Cafcass, Ministry of Justice, Department for Education, local authorities).

Objective: The family justice datasets currently have significant limitations that render is difficult or impossible to follow up certain groups of children over time. These defects should be remedied if the full potential of the datasets is to be realised.

Immediate steps: Specifically the obstacles are:

- A linking variable between two records to show that they related to the same, adopted individual should be considered. An indicator that a child was adopted would be helpful (and such an indicator is to be collected in the CLA dataset from 2017/18 but this will not enable linkage to the full longitudinal record meaning that there will continue to be difficulty in following up adopted children over time and assessing their long-term outcomes).

- It is possible to link mothers and children in the Cafcass data but the father is unknown in a large proportion of cases. Siblings cannot be identified, either. In Cafcass and FamilyMan, family members can be linked using a relationship or role identifier.[20] The challenges of identifying family members goes beyond the merely technical and was identified as a key challenge at the Knowledge Exchange Event.[2]
- Ensuring consistent and well-recorded identifying variables as recommended in Recommendation 1 would also facilitate linkage of the same individual within a dataset over time.

Recommendation 3: To understand who benefits from services, each child record in national administrative data should contain information about the child and family and services offered and received.

Audience: Data providers (Department for Education, local authorities).

Objective: We noted in Tables 1 and 2 that even with the core FJS datasets linked together, they contain limited information about the problems that families who enter the care system face or the services that they receive. There is thus a paucity of information in the national data about these crucial aspects of the experience of children and families in the care system.

Immediate steps: Local authorities collect rich data about services offered and received which are detailed in a separate report.[3] Locally-held data and local area analyses could help inform the content of the national datasets, which in turn can provide information relevant to the local area and the data collected there. What data are returned to the Department for Education, and the process for doing so, should therefore be reviewed to ensure that relevant, robust information is collected. This work should build on the local area case study[3] and other work already undertaken in this area.[115-119]

Recommendation 4: The court and legal aid data held by the Ministry of Justice should be made available for approved research projects.

Audience: Data providers (Ministry of Justice).

Objective: The value of making data held by the MoJ available for external research use is self-evident. At present, these datasets are only available, at record-level, within the MoJ, which has limited capacity to carry out academic research. By making these datasets available to external researchers—following an appropriate process to ensure that doing so is ethical and legal (see section 4.5)—their potential can be realised more quickly.

Immediate steps: Templates already exist for the sharing of government-held data, for example Cafcass and Department for Education. A safe haven infrastructure also exists within the MoJ and externally, such as the Secure Anonymised Information Linkage databank,[120] that could be exploited meaning that ethical and safe data sharing could occur within the short term. The MoJ is at the time of writing (September 2017) exploring the use of its data lab.

Recommendation 5: Data providers and the Family Justice Observatory should advocate for linkage, including considering how the Digital Economy Act 2017 and existing frameworks can enable cross-sectoral linkages by trusted third parties.

Audience: All data providers and the Family Justice Observatory.

Objective: The lives of children and families are complex with events playing out over time. Whether a particular child comes into contact with the FJS is influenced by a host of factors that are not directly relevant to the court, such as health and economic circumstances. It is therefore crucial to understand these factors. Wider services, such as healthcare services, also collect administrative data and linkages to these datasets should be carried out. Such linkages could be carried out immediately using existing data infrastructure. The Nuffield Foundation-funded study by Wijlaars and others,[38] for example, is currently linking Cafcass data with the Hospital Episode Statistics and the Clinical Records Interactive Search.

Immediate steps: Linkage between FJS and others is in its infancy. Studies to appraise linkage should be carried out and reported. The process of linkage would also be greatly facilitated by technical improvements to the datasets as noted in Recommendations 1 and 2. Cross-sectoral linkages raise legal and ethical considerations noted in sections 4.5 and 4.6 and data providers should consider how the Digital Economy Act 2017, which will enable the sharing of anonymised but linked cross-sectoral data, and existing frameworks can enable them to link and share data for research purposes.

Recommendation 6: It is necessary to improve data providers' understanding of the needs of researchers and their capacity to evaluate and report their methods for data processing.

Audience: All data providers and researchers.

Objective: Complexity inherent in processing administrative data and in linkage processes means that the way in which data are processed and linkage is carried out may affect study results. Researchers can incorporate information on linkage error to address biases in results. It is therefore vital that researchers are able to appraise the methods used by data providers to process and link data.

Immediate steps: Those carrying out linkage should therefore work closely with researchers in this regard. Researchers could, for example, collaborate with data providers to design, test and evaluate linkage strategies to maximise the quality of linkage and hence the reliability of results. The Guidance for Information about Linking Datasets[96] provides information on what should be reported by all those involved in the linkage process.

Recommendation 7: Research funders and the Family Justice Observatory should support capacity building and interdisciplinary collaboration between those with appropriate skills and knowledge.

Audience: Research funders and the Family Justice Observatory.

Objective: Using administrative and population data in the FJS requires advanced epidemiological and statistical tools as well as a thorough understanding of the system itself and therefore research teams should have adequate knowledge and expertise of both. However, there are few researchers with legal skills and few lawyers with the relevant quantitative skills.[79,80] In addition, there should be in-depth studies into the data themselves. Such studies could form part of mixed-methods approaches that incorporate both quantitative and qualitative analyses to understand the phenomena under investigation as well as the processes that lead to data generation. This is important as studying how the data are recorded and stored can help reveal data quality issues and will inform interpretation of results.

Immediate steps: Research should involve collaboration between individuals with appropriate expertise such as lawyers, statisticians, epidemiologists, qualitative researchers, local area data analysts and social workers. This is discussed further under Recommendation 9, below. Funders can contribute to this goal by ensuring adequate representation of relevant professions on multidisciplinary research teams.

Recommendation 8: An examination of how research evidence can be put into practice is required, particularly at the level of individual decisions, taking account of the fact that using research evidence is separate from using real-time identifiable client data.

Audience: Services and researchers.

Objective: A programme of interdisciplinary research should be carried out to develop models for implementing research evidence in the FJS (see section 4.7). This will require consideration of the legal status of research evidence as evidence in the court room and how research evidence can be used fairly, legally and for the benefit of children. Practical issues around access to evidence, training and

research literacy, court culture, trust in research, and a need to improve the evidence base all also need to be addressed.

Immediate steps: Funding should be made available to study how research evidence can be put into practice. This work should in particular take account of the fact that using research evidence is a distinct process to using real-time identifiable client data by services.

Recommendation 9: The Family Justice Observatory should facilitate the establishment of a family justice research community.

Audience: The Family Justice Observatory.

Objective: Meeting the challenges of using population data to understand the FJS will require the establishment of a community of researchers and data providers, analysists and users. This was one of the key messages to come out of a Knowledge Exchange Event held in January 2017, attended by a range of stakeholders in the FJS and detailed in a separate report.[2] Building such a community will avoid duplication of effort, allow comparison of different approaches to data cleaning and analysis and ensure that research builds strategically and logically over time. In particular, the event identified the following as benefits of having a family justice research community:

- Being able to keep up-to-date with relevant research and data projects
- Sharing expertise on:
    - data quality and meaning, especially at the planning stage
    - permission pathways to access data and ethical issues
    - data cleaning and making data 'research-ready' including sharing relevant code, with appropriate acknowledgement
- Improving the meta-data (i.e. data about the datasets) available for researchers
- Making a co-ordinated case for wider access to existing linkage algorithms
- Establishing and using systems for the re-use of linked data
- Incorporating public and professional engagement in the research process
- More effective dissemination of research findings to relevant audiences

A family justice research community should be interdisciplinary and involve all stakeholders: researchers, practitioners, policy makers, data providers, service users, national and local authority analysts and the wider community. Given the complexities inherent in the law and the analysis of population data, research studies should be designed by individuals with appropriate legal and methodological expertise. This will require appropriate collaboration and capacity building among all

those studying the FJS and might include developing pathways for those with legal training to undertake research careers and ensuring that scientists have opportunities for immersion in the family court and social care settings and to undergo legal tuition.

Immediate steps: The Family Justice Observatory should be able to facilitate this by acting as a central, co-ordinating body for sharing findings and expertise, development of specialist data safe havens and carrying out public and professional engagement. By doing so, it can also contribute, in collaboration with researchers, data providers and funders, to meeting all of the above recommendations.

Conclusion

Population data have the capacity to be transformative of the FJS. The use of such data enables the specification of research questions that provide robust and relevant answers about the functioning of the FJS and the consequences for children and families who go through it. These data have only been used in a handful of studies to date by a very small community of active researchers. The use of scientific evidence in the FJS lags far behind that of other fields such as health, where data and research evidence have for decades formed part of routine clinical practice and the evaluation of new interventions and policies. Significant challenges therefore remain to be overcome but these are not insurmountable. There is a moral imperative to do so if we are to ensure that, as required by s 1(1) of the Children Act 1989, the best interests of the child are treated as paramount. A failure properly to meet these challenges would be not just a failure of the FJS but would be to fail the very children whose welfare we seek to protect and to promote.

References

1. Administrative Data Research Network. Glossary. https://adrn.ac.uk/public-engagement/pe-across-the-network/glossary/ (accessed 21 August 2017).

2. Woodman J, Broadhurst K, Gilbert R. (2017) To*wards a Family Justice Observatory - a scoping study: using population level data to understand the family justice system: report from a knowledge exchange event*. http://wp.lancs.ac.uk/observatory-scoping-study/ (accessed 13 July 2017).

3. Holmes L, Gillson D, Broadhurst K. (2017) *Strategic and operational use of data at the local level: findings from a case study* (in press, to be availale at http://wp.lancs.ac.uk/observatory-scoping-study/publications/.

4. Lancaster University. (2017) *Towards a family justice observatory: a scoping study*. http://wp.lancs.ac.uk/observatory-scoping-study/ (accessed 24 July 2017).

5. Family Justice Review Panel. (2011) *Family Justice Review: Interim Report*.

6. Broadhurst K, Alrouh B, Yeend E, et al. Connecting Events in Time to Identify a Hidden Population: Birth Mothers and Their Children in Recurrent Care Proceedings in England. *Br J Soc Work* 2015; **45**(8): 2241-60.

7. Bainham A. Private and public children law: an under-explored relationship. *Child & Fam LQ* 2013; **25**: 138.

8. Bream V, Buchanan A. Distress among children whose separated or divorced parents cannot agree arrangements for them. *Br J Soc Work* 2003; **33**(2): 227-38.

9. Trinder L, Kellet J. *The longer-term outcomes of in-court conciliation*. London, 2007.

10. Bryson C, Purdon S, Skipp A, et al. (2017) *Understanding the lives of separating and separated parents in the UK: what evidence do we need?* London: The Nuffield Foundation. http://nuffieldfoundation.org/data-understand-lives-separated-families (accessed 8 June 2017).

11. Cafcass. Care demand statistics. https://www.cafcass.gov.uk/leaflets-resources/organisational-material/care-and-private-law-demand-statistics/care-demand-statistics.aspx (accessed 24 July 2017).

12. Cafcass. Private law demand statistics. https://www.cafcass.gov.uk/leaflets-resources/organisational-material/care-and-private-law-demand-statistics/private-law-demand-statistics.aspx (accessed 24 July 2017).

13. Family Justice Review Panel. *Family Justice Review: Final Report*, 2011.

14. Bywaters P, Jones C, Sparks T. (2017) *Identifying and Understanding Inequalities in Child Welfare Intervention Rates: Comparative studies in four UK countries. Single country quantitative study*

report: England. http://www.nuffieldfoundation.org/inequalities-child-welfare-intervention-rates (accessed 19 September 2017).

15. Cafcass. About Cafcass. https://www.cafcass.gov.uk/about-cafcass.aspx (accessed 6 July 2017).

16. Ministry of Justice. Ministry of Justice. https://www.gov.uk/government/organisations/ministry-of-justice (accessed 6 July 2017).

17. HM Courts & Tribunals Service. HM Courts & Tribunals Service. https://www.gov.uk/government/organisations/hm-courts-and-tribunals-service (accessed 6 July 2017).

18. Ministry of Justice. Legal Aid Statistics in England and Wales. https://www.gov.uk/government/collections/legal-aid-statistics (accessed 1 August 2017).

19. Department for Education. Department for Education. https://www.gov.uk/government/organisations/department-for-education (accessed 6 July 2017).

20. Alrouh, B. and Broadhurst, K. (2015) *Using Cafcass administrative datasets to study recurrent care proceedings: Technical appendix*. http://wp.lancs.ac.uk/recurrent-care/ (accessed 19 September 2017).

21. Herbert A, Gilbert R, Gonzalez-Izquierdo A, Pitman A, Li L. 10-y Risks of Death and Emergency Re-admission in Adolescents Hospitalised with Violent, Drug- or Alcohol-Related, or Self-Inflicted Injury: A Population-Based Cohort Study. *PLoS medicine* 2015; **12**(12): e1001931.

22. McGovern A, Woodman J, Allister J, et al. A simple clinical coding strategy to improve recording of child maltreatment concerns: an audit study. *Inform Prim Care* 2015; **22**(1): 227–34.

23. Woodman J, Allister J, Rafi I, et al. Simple approaches to improve recording of concerns about child maltreatment in primary care records: developing a quality improvement intervention. *Br J Gen Pract* 2012; **62**(600): e478-e86(9).

24. MRC Lifelong Health and Ageing Unit at UCL. National Survey of Health and Development. http://www.nshd.mrc.ac.uk/nshd/ (accessed 21 June 2017).

25. Centre for Longitudinal Studies. The 1958 National Child Development Study. http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=724&sitesectiontitle=National+Child+Development+Study (accessed 21 June 2017).

26. Centre for Longitudinal Studies. The 1970 British Cohort Study. http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=795&sitesectiontitle=Welcome+to+the+1970+British+Cohort+Study (accessed 21 June 2017).

27. NHS Digital. Hospital Episode Statistics. http://content.digital.nhs.uk/hes (accessed 21 June 2017).

28. Maudsley Biomedical Research Centre. Clinical Records Interactive Search. http://www.maudsleybrc.nihr.ac.uk/facilities/clinical-record-interactive-search-cris/ (accessed 21 June 2017).

29. Downs J, Gilbert R, Hayes RD, Hotopf M, Ford T. Linking health and education data to plan and evaluate services for children. *Arch Dis Child* 2017; **102**(7): 599-602.

30. MHRA. Clinical Practice Research Datalink. https://www.cprd.com/intro.asp (accessed 21 June 2017).

31. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International journal of epidemiology* 2015; **44**(3): 827-36.

32. University College London. The Health Improvement Network. https://www.ucl.ac.uk/pcph/research-groups-themes/thin-pub/database (accessed 21 June 2017).

33. In Practice Systems Ltd. The Health Improvement Network. http://www.inps.co.uk/vision/health-improvement-network-thin (accessed 21 June 2017).

34. NHS Digital. Maternity and Children's Data Set. http://www.content.digital.nhs.uk/maternityandchildren (accessed 21 June 2017).

35. NHS Digital. Mental Health Services Data Set. http://content.digital.nhs.uk/mhsds (accessed 21 June 2017).

36. NHS Digital. Demographics. https://digital.nhs.uk/Demographics (accessed 4 April 2017).

37. Office for National Statistics. Personal Demographics Service data. https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/datasourceoverviews/personaldemographicsservicedata (accessed 7 August 2017).

38. Nuffield Foundation. Understanding the health needs of mothers and children involved in family court cases. http://www.nuffieldfoundation.org/understanding-health-needs-mothers-and-children-involved-family-court-cases (accessed 23 May 2017).

39. Department for Education. National pupil database. https://www.gov.uk/government/collections/national-pupil-database (accessed 21 June 2017).

40. Sebba J, Berridge D, Luke N, et al. *The Educational Progress of Looked After Children in England: Linking Care and Educational Data*. http://reescentre.education.ox.ac.uk/research/educational-progress-of-looked-after-children/ (accessed 19 September 2017).

41. Department for Work and Pensions. Work and Pensions Longitudinal Study. https://www.gov.uk/government/statistics/work-and-pensions-longitudinal-study (accessed 21 June 2017).

42. Administrative Data Research Network. Work and Pensions Longitudinal Study. https://www.adrn.ac.uk/get-data/catalogue/cataloguepage?sn=888025 (accessed 21 June 2017).

43. HM Revenue & Customs. HMRC Data Catalogue. https://www.gov.uk/government/publications/hmrc-data-catalogue (accessed 21 June 2017).

44. Administrative Data Research Network. Child Benefit. http://www.adrn.ac.uk/get-data/catalogue/cataloguepage?sn=888026 (accessed 21 June 2017).

45. Administrative Data Research Network. Tax Credits Claimants. http://www.adrn.ac.uk/get-data/catalogue/cataloguepage?sn=888027 (accessed 21 June 2017).

46. Department for Education. Graduate outcomes: longitudinal education outcomes (LEO) data. https://www.gov.uk/government/statistics/graduate-outcomes-longitudinal-education-outcomes-leo-data (accessed 11 July 2017).

47. Office for National Statistics. Income and benefits. https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/datasourceoverviews/incomeandbenefits (accessed 7 August 2017).

48. Department for Education (2017) *Analysing family circumstances and education*. https://www.gov.uk/government/consultations/analysing-family-circumstances-and-education (accessed 26 June 2017).

49. Francis B, Soothill K, Humphreys L. *Developing a reoffending measure using the Police National Computer database*. Lancaster: Lancaster University. http://www.research.lancs.ac.uk/portal/en/publications/development-of-a-reoffending-measure-using-the-police-national-computer-database(3af086da-bccd-436a-81dd-3a2d6939eb84).html (accessed 27 June 2017).

50. Francis B, Crosland P. (2002) *The Police National Computer and the Offenders Index: can they be combined for research purposes? Full Report*. http://eprints.lancs.ac.uk/50042/ (accessed 11 July 2017).

51. Administrative Data Research Network. Police Force Crime Records. http://www.adrn.ac.uk/get-data/catalogue/cataloguepage?sn=888053 (accessed 21 June 2017).

52. Ministry of Justice. Criminal justice statistics. https://www.gov.uk/government/collections/criminal-justice-statistics (accessed 21 June 2017).

53. Ministry of Justice. Criminal court statistics. https://www.gov.uk/government/collections/criminal-court-statistics (accessed 21 June 2017).

54. The Children's Commissioner (2017) *Stability Index*. London: Children's Commissioner for England. https://www.childrenscommissioner.gov.uk/publications/stability-index-initial-findings-and-technical-report (accessed 12 June 2017).

55. University of Bristol. Avon Longitudinal Study of Parents and Children. http://www.bristol.ac.uk/alspac/ (accessed 21 June 2017).

56. Born in Bradford. Born in Bradford Study. https://borninbradford.nhs.uk/ (accessed 21 June 2017).

57. Centre for Longitudinal Studies. Millenium Cohort Study. http://www.cls.ioe.ac.uk/page.aspx?sitesectionid=851 (accessed 21 June 2017).

58. Sidebotham P, Golding J. Child maltreatment in the "children of the nineties" a longitudinal study of parental risk factors. *Child Abuse & Neglect* 2001; **25**(9): 1177-200.

59. Sidebotham P, Heron J. Child maltreatment in the "children of the nineties": a cohort study of risk factors. *Child Abuse & Neglect* 2006; **30**(5): 497-522.

60. Sidebotham P, Heron J. Child maltreatment in the "children of the nineties:" the role of the child. *Child Abuse & Neglect* 2003; **27**(3): 337-52.

61. Sidebotham P, Heron J, Golding J. Child maltreatment in the "Children of the Nineties:" deprivation, class, and social networks in a UK sample. *Child Abuse & Neglect* 2002; **26**(12): 1243-59.

62. Zhang ML, Henderson M, Cheung SY, Scourfield J, Sharland E. Predicting the recipients of social work support, and its impact on emotional and behavioural problems in early childhood. *Child Fam Soc Work* 2017; **22**(2): 772-81.

63. Department for Education. LSYPE: Longitudinal Study of Young People in England. https://www.education.gov.uk/ilsype/workspaces/public/wiki/LSYPE (accessed 21 June 2017).

64. Office for National Statistics. Labour Force Survey. https://www.ons.gov.uk/surveys/informationforhouseholdsandindividuals/householdandindividualsurveys/labourforcesurveylfs (accessed 21 June 2017).

65. Institute for Social and Economic Research. Understanding Society. https://www.understandingsociety.ac.uk/ (accessed 21 June 2017).

66. Henderson M, Scourfield J, Cheung SY, Sharland E. Predictors of Social Service Contact among Teenagers in England. *The British Journal of Social Work* 2016; **46**(6): 1485-501.

67. Henderson M, Scourfield J, Cheung SY, Sharland E, Sloan L. The Effects of Social Service Contact on Teenagers in England. *Res Soc Work Pract* 2016; **26**(4): 386-98.

68. UK Data Service. https://www.ukdataservice.ac.uk/ (accessed 21 June 2017).

69. MacMillan HL, Wathen CN, Barlow J, Fergusson DM, Leventhal JM, Taussig HN. Interventions to prevent child maltreatment and associated impairment. *The Lancet* 2009; **373**(9659): 250-66.

70. Department for Education. Children looked after in England including adoption: 2015 to 2016. https://www.gov.uk/government/statistics/children-looked-after-in-england-including-adoption-2015-to-2016 (accessed 10 April 2017).

71. Office for National Statistics. Population estimates. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates (accessed 20 April 2017).

72. Lilford R, Mohammed MA, Spiegelhalter D, Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *The Lancet* 2004; **363**(9415): 1147-54.

73. Gilbert R, Kemp A, Thoburn J, et al. Recognising and responding to child maltreatment. *Lancet* 2009; **373**(9658): 167-80.

74. Mc Grath-Lone L, Dearden L, Harron K, Nasim B, Gilbert R. Factors associated with re-entry to out-of-home care among children in England. *Child Abuse & Neglect* 2017; **63**: 73-83.

75. Mc Grath-Lone L, Dearden L, Nasim B, Harron K, Gilbert R. Changes in first entry to out-of-home care from 1992 to 2012 among children in England. *Child Abuse & Neglect* 2016; **51**: 163-71.

76. Hennekens C, Buring J. *Epidemiology in Medicine*. Philadelphia, PA: Lippincott Williams & Wilkins; 1987.

77. Mayer EK, Bottle A, Aylin P, Darzi AW, Vale JA, Athanasiou T. What is the role of risk-adjusted funnel plots in the analysis of radical cystectomy volume-outcome relationships? *BJU International* 2011; **108**(6): 844-50.

78. McCrudden C. Legal research and the social sciences. *Law Quarterly Review* 2006; **122**(Oct): 632-50.

79. Genn H, Partington M, Wheeler S. (2006) Law in the real world: improving our understanding of how law works. http://www.nuffieldfoundation.org/nuffield-inquiry-empirical-legal-research-law-real-world (accessed 13 July 2017).

80. Broadhurst K, Robertson L, Mason C, Bowyer S, Wilkinson J. (2017) *Towards a Family Justice Observatory - a scoping study: main findings report of the national stakeholder consultation*. http://wp.lancs.ac.uk/observatory-scoping-study/ (accessed 1 August 2017).

81. Ministry of Justice. Family court statistics quarterly: October to December 2016. https://www.gov.uk/government/collections/family-court-statistics-quarterly (accessed 3 April 2017).

82. Harwin J, Alrouh B, Palmer M, Broadhurst KE, Swift S. (2015) *A national study of the usage of supervision orders and special guardianship over time (2007-2016): Briefing paper no 1: Special guardianship orders*. http://www.nuffieldfoundation.org/supervision-orders-and-special-guardianship (accessed 19 September 2017).

83. Robling M, Bekkers M-J, Bell K, et al. Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (Building Blocks): a pragmatic randomised controlled trial. *The Lancet*; **387**(10014): 146-55.

84. Administrative Data Liaison Service. Administrative data introduction. http://www.adls.ac.uk/adls-resources/guidance/introduction/ (accessed 24 August 2017).

85. Bentley H, O'Hagan O, Brown A, et al. (2017) How safe are our children? NSPCC. https://www.nspcc.org.uk/services-and-resources/research-and-resources/2014/how-safe-are-our-children-2014/ (accessed 24 August 2017).

86. Mc Grath-Lone L, Harron K, Dearden L, Nasim B, Gilbert R. Data Resource Profile: Children Looked After Return (CLA). *International journal of epidemiology* 2016; **45**(3): 716-7f.

87. Information Commissioner's Office. Anonymisation. https://ico.org.uk/for-organisations/guide-to-data-protection/anonymisation/ (accessed 17 July 2017).

88. Desai T, Ritchie F, Welpton R. (2016) Five safes: designing data access for research. Bristol: University of the West of England http://rsss.anu.edu.au/sites/default/files/Ritchie_5safes.pdf (accessed 27 June 2017).

89. Administrative Data Research Network. How does the Network protect your privacy? https://adrn.ac.uk/public-engagement/pe-across-the-network/protect-privacy/ (accessed 21 August 2017).

90. Ritchie F, Welpton R. (2011) Sharing risks, sharing benefits: data as a public good. In: Work session on statistical data confidentiality 2011; Eurostat,, Tarragona, Spain, 26 - 28 October 2011. http://eprints.uwe.ac.uk/22460 (accessed 17 July 2017).

91. Filippon J. Slow and costly access to anonymised patient data impedes academic research. *BMJ* 2015; **351**:h5087.

92. Gilbert R, Goldstein H, Hemingway H. The market in healthcare data. *BMJ* 2015; **351**:h5897.

93. Harron K. (2016) *Better knowledge, better society: an introduction to data linkage*. Colchester: ADRN. https://adrn.ac.uk/media/174200/data_linkage_katieharron_2016.pdf (accessed 19 September 2017).

94. Nuffield Foundation. Understanding the scale, pattern and dynamics of birth fathers' recurrent appearance in care proceedings. http://www.nuffieldfoundation.org/understanding-scale-pattern-and-dynamics-birth-fathers-recurrent-appearance-care-proceedings (accessed 1 September 2017). 2017.

95. Administrative Data Research Network. Protecting privacy. https://adrn.ac.uk/get-data/protecting-privacy/ (accessed 21 August 2017).

96. Gilbert R, Lafferty R, Hagger-Johnson G, et al. GUILD: GUidance for Information about Linking Data sets. *J Public Health* 2017: 1-8.

97. ADRN. Data Reuse for Research Purposes Policy (ADRN034). https://www.adrn.ac.uk/get-data/documentation/ (accessed 24 August 2017).

98. Marmot M, Atkinson T, Bell J, et al. (2010) Fair *society, healthy lives: the Marmot review: strategic review of health inequalities in England post 2010*. http://www.instituteofhealthequity.org/resources-reports/fair-society-healthy-lives-the-marmot-review (accessed 21 July 2017).

99. Kaufman J. Social epidemiology. In: Rothman K, Greenland S, Lash T, eds. *Modern Epidemiology*, 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.

100. Kuh D, Ben-Shlomo Y, Lynch J, Hallqvist J, Power C. Life course epidemiology. *JECH* 2003; **57**(10): 778-83.

101. Eisenberg T. Why do empirical legal scholarship. *San Dieg L Rev* 2004; **41**: 1741-6.

102. Ubbesen MB, Gilbert R, Thoburn J. Cumulative incidence of entry into out-of-home care: Changes over time in Denmark and England. *Child Abuse & Neglect* 2015; **42**: 63-71.

103. O'Donnell M, Maclean M, Sims S, Brownell M, Ekuma O, Gilbert R. Entering out-of-home care during childhood: Cumulative incidence study in Canada and Australia. *Child Abuse & Neglect* 2016; **59**: 78-87.

104. Wildeman C, Emanuel N. Cumulative Risks of Foster Care Placement by Age 18 for U.S. Children, 2000–2011. *PloS ONE* 2014; **9**(3): e92785.

105. The JAMA Network. (2014) Evidence-based medicine: an oral history. http://ebm.jamanetwork.com/ (accessed 20 July 2017).

106. Smith R, Rennie D. Evidence based medicine—an oral history. *BMJ* 2014; **348**:g371.

107. Dennis I. *The Law of Evidence*, 5th edn. London: Sweet & Maxwell; 2013.

108. Rathus Z. The role of social science in Australian family law: collaborator, usurper or infiltrator? *Family Court Review* 2014; **52**(1): 69-89.

109. Rathus Z. Social science or 'lego-science'? Presumptions, politics, parenting and the new family law. *Queensland University of Technology Law and Justice Journal* 2010; **10**(2): 164-90.

110. Rathus Z. A call for clarity in the use of social science research in family law decision-making. *Australian Journal of Family Law* 2012; **26**: 81-115.

111. Sturge C, Glaser D. Contact and domestic violence - the experts' court report. *Fam Law* 2000; **30**(9): 615.

112. Tomkins A. *Public Law*. Oxford: OUP; 2003.

113. Masson J. (2015) Developing Judgment: The role of Feedback for Judges in the Family Court. Bristol: School of Law, University of Bristol. http://research-information.bristol.ac.uk/files/45213553/FINAL_REPORT_pub.pdf (accessed 16 August 2017).

114. Lum K, Isaac W. To predict and serve? *Significance* 2016; **13**(5): 14-9.

115. Gatehouse M, Ward H, Holmes L. (2008) Deve*loping Definitions of Local Authority Services and Guidance for the New Children in Need Census*. London: Department of Children, Schools and Families.

116. Holmes L, McDermid S. (2012) *Understanding Costs and Outcomes in Child Welfare Services*. London: Jessica Kingsley Publishers.

117. Ward H, Holmes L, Dyson P. (2008) *The Costs and Consequences of Child Welfare Interventions: Mapping Children in Need Services*. Loughborough: Centre for Child and Family Research, Loughborough University.

118. Ward H, Holmes L, Soper J. (2008) *Costs and Consequences of Placing Children in Care*. London: Jessica Kingsley Publishers.

119. McDermid S. The nature and availability of child level data on children in need for use by Children's Services practitioners and managers. *Research, Policy and Planning* 2008; **26**(3): 183-92.

120. SAIL Databank. SAIL Databank. https://saildatabank.com/ (accessed 17 July 2017).

121. Cafcass. Policies. https://www.cafcass.gov.uk/leaflets-resources/policies-and-templates-for-associate-fcas/policies.aspx (accessed 31 July 2017).

122. Cafcass. Research at Cafcass. https://www.cafcass.gov.uk/contact-us/research.aspx (accessed 31 July 2017).

123. Gov. DBS checks (previously CRB checks). https://www.gov.uk/disclosure-barring-service-check/overview (accessed 31 July 2017).

124. University of Bristol (2017) Establishing outcomes of care proceedings for children before and after care proceedings reform. http://gtr.rcuk.ac.uk/projects?ref=ES/M008541/1 (accessed 4 April 2017).

125. Rodger H, Cappleman S, Dale R. (2015) Cafcass study of parental order applications made in 2013/14. https://www.cafcass.gov.uk/contact-us/research.aspx (accessed 11 May 2017).

126. Trinder L, Hunt J, Macleod A, Pearce J, Woodward H. (2013) Enforcing contact orders: problem-solving or punishment? Exeter: Exeter Law School. https://ore.exeter.ac.uk/repository/handle/10871/16765 (accessed 3 April 2017).

127. Cafcass (2012) Analysis of congruence between Cafcass recommendations and the final court order

https://www.cafcass.gov.uk/media/126292/cafcass_analysis_of_congruence_study_final_101212.pdf (accessed 26 June 2017).

128. Ministry of Justice. Family Justice Statistics Quarterly. https://www.gov.uk/government/collections/family-court-statistics-quarterly (accessed 22 June 2017).

129. Ministry of Justice. User guide to legal aid statistics in England and Wales. https://www.gov.uk/government/statistics/a-guide-to-legal-aid-statistics-in-england-and-wales (accessed 1 August 2017).

130. Department for Education. Children Looked After by Local Authorities in England: Guide to the SSDA903 Collection 1 April 2017 to 31 March 2018. https://www.gov.uk/government/publications/children-looked-after-return-2017-to-2018-guide (accessed 23 May 2017).

131. Department for Education. Children in need census 2017 to 2018: Guide for local authorities – version 1.1. https://www.gov.uk/government/publications/children-in-need-census-2017-to-2018-guide (accessed 23 May 2017).

132. Department for Education. Children in need census 2017 to 2018: specification https://www.gov.uk/government/publications/children-in-need-census-2017-to-2018-specification (accessed 23 May 2017).

133. Department for Education. Characteristics of children in need: 2015 to 2016. https://www.gov.uk/government/statistics/characteristics-of-children-in-need-2015-to-2016 (accessed 18 September 2017).

134. Department for Education. CLA return 2016 to 2017: technical specification https://www.gov.uk/government/publications/children-looked-after-return-2016-to-2017-technical-specification (accessed 23 May 2017).

135. Department for Communities and Local Government. *English indices of deprivation 2015*. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015 (accessed 11 July 2017).

136. Clark D, King A, Sharpe K, et al. Linking routinely collected social work, education and health data to enable monitoring of the health and health care of school-aged children in state care ('looked after children') in Scotland: a national demonstration project. *Public Health* 2017; **150**: 101-11.

137. Bewley H, George A, Rienzo C, Portes J. *The Impact of the Troubled Families Programme: Findings from the Analysis of National Administrative Data*.

https://www.niesr.ac.uk/publications/national-evaluation-troubled-families-programme-national-impact-study-report (accessed 19 September 2017).

138. Wade J, Simclair I, Stuttard J, Simmonds J. (2014) *Investigating Special Guardianship: experience, challenges and outcomes*. Research report available at https://www.gov.uk/government/publications/investigating-special-guardianship (accessed 21 March 2017).

139. Selwyn J, Wijedasa D, Meakings S. *Beyond the Adoption Order: challenges, interventions and adoptions*. London, 2014.

140. Department for Education (2013) *Data pack: improving permanence for looked after children*. https://www.gov.uk/government/publications/improving-permanence-for-looked-after-children-data-pack (accessed 1 June 2017).

141. Schofield G, Thoburn J, Howell D, Dickens J. The Search For Stability and Permanence: Modelling the Pathways of Long-stay Looked After Children. *The British Journal of Social Work* 2007; **37**(4): 619-42.

142. Dickens J, Howell D, Thoburn J, Schofield G. Children Starting to be Looked After by Local Authorities in England: An Analysis of Inter-authority Variation and Case-centred Decision Making. *The British Journal of Social Work* 2007; **37**(4): 597-617.

Appendix I – Cafcass data profile

Children and Family Court Advisory and Support Service (Cafcass) social workers are appointed in family court cases to safeguard and promote the welfare of children. In public law cases a 'Children's Guardian' is appointed to represent the child independently of the local authority and the parents. In private law cases, a 'Family Court Adviser' provides safeguarding information to the court for the first hearing. If welfare issues are identified, Cafcass may be ordered to carry out further work after the first hearing, such as providing a welfare report, and in some cases representing the child as a Children's Guardian. Cafcass' Electronic Case Management System (ECMS), formerly the Case Management System (CMS), contains information on cases, applications and legal outcomes as well as individuals involved in cases.

A1.1. Overview

Data provider	The Children and Family Court Advisory and Support Service (Cafcass)
Population covered	All family court cases where Cafcass is involved. This includes public law cases (all section 31 care order applications, placement order applications, and some adoption applications) and private law cases involving children (all section 8 applications, some enforcement applications and some special guardianship applications) except financial matters (estimated to be about 10% of all private law cases e.g. disputes over child support payments). In around 70% of its private law cases, Cafcass is not involved beyond the first hearing. The dataset does not include any children entering care via non-judicial routes (section 20).
Size	• 2016/17: 14,596 care order applications[11] • 2016/17: 40,580 private law cases[12] • Between April 2007 and March 2014, there were 43,635 unique mothers and 84,714 unique children linked with public law applications[20]
Data overview	Demographic data (names, gender, date of birth, ethnicity [only recently made mandatory] and postcode history), case factors (e.g. allegations of domestic abuse), relationships among parties, hearing and application dates, proceeding outcomes (e.g. legal orders granted – only final legal orders from July 2014 onwards). Data are available at individual, application, case and court level. Not all of the above data are mandatory to input or available across the whole timeframe.
Structure	Two distinct databases: the 'old' Case Management System (CMS) and the 'new' Electronic CMS (ECMS). The ECMS was introduced in July 2014 and has a different table structure. There are three main units of analysis levels within the system: cases, applications and persons. Data are stored as a relational database. See Alrouh and Broadhurst[20] for detail of the structure.
Years covered	April 2007 onwards.
Data entry	Most data are entered by administrative staff and some by the practitioners. Administrators create a case record on receipt of an application. Cafcass

	practitioners and administrators enter data and upload documents while the case is ongoing and on case closure. Some practitioners have authority to close their own cases and some are subject to Service Manager oversight. In public law cases, some data are provided by the local authority so those fields may be more complete. Some missing fields are flagged up in the ECMS; however, a case can be closed so long as a legal output has been entered. Service Managers should periodically assess the quality of practitioners' closure practices. A Case Recording and Retention Policy is available. Some information on case recording is also in the Operating Framework. Both of these are available from the Cafcass website.[121]
Permission pathway, ethics & costs	Access to Cafcass data is governed by an internal Cafcass process, details of which are available on Cafcass' website.[122] The process is represented in Figure 11, below, which is based on the Cafcass Research Governance Framework. The research governance application form is available to download from the website. It includes space for researchers' details, sources of data, aims and objectives, methodology, data storage and information security arrangements, timetable and an impact assessment. Requests for access to data requires submission of a current, clear Disclosure and Barring Service check.[123] Ethical approval is required and can be obtained from the researchers' own institution or, for those at institutions without a research ethics committee, from the Cafcass Research Governance Committee. There are currently no costs for access other than reimbursed costs for computer equipment required to access the database (Cafcass require access via their secure computer system).
URL	https://www.cafcass.gov.uk/contact-us/research.aspx (accessed 1 August 2017).
Linkage	Not routinely done with any other data sources. Family members across time have been linked (Broadhurst et al)[6,20] and linkage between Cafcass and hospital episodes statistics is planned by Wijlaars and others.[38] Cafcass data have been linked as part of a data share project with the Ministry of Justice and Department for Education but this is not currently available to researchers: see Appendix II for more details.

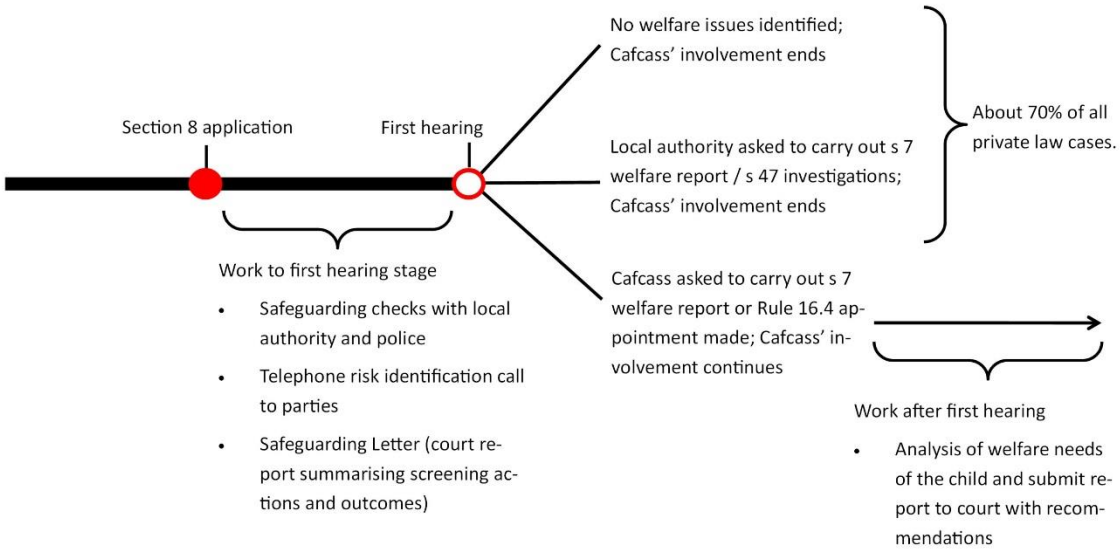
A1.2. Data completeness and accuracy

A1.2.1. Population

The Cafcass database only includes cases where Cafcass have been involved. This is public law cases (all care order applications, placement order applications and some adoption applications) and private law cases involving children (all section 8 applications, some enforcement applications and some special guardianship applications) except financial matters (estimated to be about 10% of all private law cases e.g. disputes over child support payments).[81] The information held on private law cases is more limited than public law cases as Cafcass' involvement in private law cases ends at the first hearing in about 70% of cases. Longitudinal data for individual cases (including legal outcome) is only available therefore for the remaining 30% of cases (see Figure 10). There are cases where Cafcass has

been asked to provide a welfare report or where a Children's Guardian has been appointed.* It should be noted that of the 70% in which Cafcass are not involved, there still may be welfare issues being investigated by the local authority; if the authority subsequently applies for a section 31 order, this will be recorded in the Cafcass data as a new public law case and can be linked using the identities of the children or family members involved.[6,20]

Figure 10: Simplified representation of a private law case and Cafcass' involvement



Adoption is poorly represented in the Cafcass data as the organisation is not normally involved after the care order has been granted. Similarly, Cafcass holds no data on children becoming looked after via a section 20 agreement with parents, which is about 60% of children who become looked after each year (such children are captured in the children looked after dataset, detailed in Appendix IV).[70]

A1.2.2. Validation checks and missing data

In order to describe the proportion of missing data on key variables, we analysed data on all cases commenced between 01/04/2007 and 31/03/2017. The numbers of cases and people on cases in public and private family law are given in Tables 6 and 7, respectively. These tables also show the

* Rule 16.4 appointments are cases where a guardian is appointed to represent the child. These are typically cases involving children subject to entrenched and emotionally harmful parental conflict. They are a small minority.

number and proportion missing on key variables. Note that these are for all public and private law cases: proportions may differ by type of case (e.g. application for care order). Further, we made no attempt to complete missing data by linking up individual records and filling in blanks from previous or future cases and we made no attempt to validate data that were not missing. These figures should therefore be taken as a preliminary estimate of missingness only.

Table 6: Missing data in Cafcass public law cases by year case started

	2007/8	2008/9	2009/10	2010/11	2011/12	2012/13	2013/14	2014/15	2015/16	2016/17
Cases										
Number	10492	10386	12085	11930	13068	13653	13903	14627	17071	19032
Missing end date	0 (0.0%)	0 (0.0%)	1 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.0%)	2 (0.0%)	17 (0.1%)	196 (1.1%)	4113 (21.6%)*
Missing Local authority	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.0%)	3 (0.0%)	0 (0.0%)	0 (0.0%)	137 (0.9%)	229 (1.3%)	240 (1.3%)
No application recorded	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.0%)	0 (0.0%)	1 (0.0%)
No child recorded	0 (0.0%)	1 (0.0%)	0 (0.0%)	1 (0.0%)	0 (0.0%)	0 (0.0%)	9 (0.1%)	25 (0.2%)	35 (0.2%)	54 (0.3%)
No respondent recorded	127 (1.2%)	88 (0.8%)	156 (1.3%)	691 (5.8%)	639 (4.9%)	531 (3.9%)	567 (4.1%)	608 (4.2%)	685 (4.0%)	609 (3.2%)
Child										
Number	16586	16574	19937	19962	21292	22371	22773	24175	27826	31037
Missing forename	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	4 (0.0%)	0 (0.0%)	1 (0.0%)
Missing surname	1 (0.0%)	0 (0.0%)	1 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.0%)
Missing gender	7 (0.0%)	0 (0.0%)	1 (0.0%)	2 (0.0%)	6 (0.0%)	4 (0.0%)	5 (0.0%)	38 (0.2%)	44 (0.2%)	44 (0.1%)
Missing date of birth	2 (0.0%)	2 (0.0%)	6 (0.0%)	19 (0.1%)	14 (0.1%)	6 (0.0%)	4 (0.0%)	82 (0.3%)	40 (0.1%)	39 (0.1%)
Missing ethnicity†	4285 (25.8%)	5020 (30.3%)	10661 (53.5%)	11728 (58.8%)	12665 (59.5%)	15525 (69.4%)	18070 (79.3%)	18280 (75.6%)	7143 (25.7%)	4293 (13.8%)
Missing postcode	7444 (44.9%)	6902 (41.6%)	8076 (40.5%)	7498 (37.6%)	8228 (38.6%)	8504 (38.0%)	8500 (37.3%)	9684 (40.1%)	11710 (42.1%)	12331 (39.7%)
Adults involved										
Number	22615	23018	26286	20252	22367	23539	23945	25606	29876	33711
Missing forename	1 (0.0%)	2 (0.0%)	5 (0.0%)	2 (0.0%)	4 (0.0%)	1 (0.0%)	1 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Missing surname	2 (0.0%)	1 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.0%)
Missing gender	139 (0.6%)	105 (0.5%)	135 (0.5%)	129 (0.6%)	153 (0.7%)	138 (0.6%)	147 (0.6%)	227 (0.6%)	262 (0.9%)	370 (1.1%)
Missing date of birth	6418 (28.4%)	5783 (25.1%)	6038 (23.0%)	2885 (14.2%)	2925 (13.1%)	2841 (12.1%)	2647 (11.1%)	2769 (10.8%)	3168 (10.6%)	3386 (10.0%)
Missing ethnicity†	8721 (38.6%)	9421 (40.9%)	15391 (58.6%)	11569 (57.1%)	13208 (59.1%)	15373 (65.3%)	17316 (72.3%)	17639 (68.9%)	8101 (27.1%)	5428 (16.1%)
Missing postcode	5177 (22.9%)	3854 (16.7%)	4041 (15.4%)	2653 (13.1%)	2781 (12.4%)	2804 (11.9%)	2572 (10.7%)	3036 (11.9%)	3878 (13.0%)	4706 (14.0%)

There were 65 cases with missing start date. * On-going cases have an empty end date. † Ethnicity became a mandatory field in July 2015.

Table 7: Missing data in Cafcass private law cases by year case started

	2007/8	2008/9	2009/10	2010/11	2011/12	2012/13	2013/14	2014/15	2015/16	2016/17
Cases										
Number	38932	40856	45790	43370	41663	45886	46355	33095	38632	41787
Missing end date	0 (0.0%)	4 (0.0%)	1 (0.0%)	7 (0.0%)	8 (0.0%)	23 (0.1%)	64 (0.1%)	212 (0.6%)	863 (2.2%)	5420 (13.0%)
Missing Local authority*	29142 (74.9%)	28642 (70.1%)	31530 (68.9%)	37655 (86.8%)	11343 (27.2%)	19633 (42.8%)	17992 (38.8%)	16063 (48.5%)	8393 (21.7%)	5746 (13.8%)
No application recorded	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.0%)	0 (0.0%)	0 (0.0%)
No child recorded	2 (0.0%)	6 (0.0%)	5 (0.0%)	2 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.0%)	69 (0.2%)	74 (0.2%)	197 (0.5%)
No respondent recorded	784 (2.0%)	827 (2.0%)	747 (1.6%)	234 (0.5%)	273 (0.7%)	275 (0.6%)	826 (1.8%)	1283 (3.9%)	2070 (5.4%)	2317 (5.5%)
No applicant recorded	457 (1.2%)	297 (0.7%)	182 (0.4%)	44 (0.1%)	28 (0.1%)	44 (0.1%)	624 (1.3%)	342 (1.0%)	156 (0.4%)	300 (0.7%)
Child										
Number	60168	62762	70267	65914	62868	68857	69235	49963	57979	62615
Missing forename	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (0.0%)	0 (0.0%)	0 (0.0%)
Missing surname	0 (0.0%)	0 (0.0%)	1 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Missing gender	102 (0.2%)	43 (0.1%)	7 (0.0%)	9 (0.0%)	1 (0.0%)	3 (0.0%)	13 (0.0%)	66 (0.1%)	90 (0.2%)	134 (0.2%)
Missing date of birth	3 (0.0%)	1 (0.0%)	18 (0.0%)	62 (0.1%)	90 (0.1%)	50 (0.1%)	54 (0.1%)	112 (0.2%)	117 (0.2%)	101 (0.2%)
Missing ethnicity†	32136 (53.4%)	35895 (57.2%)	51804 (73.7%)	52141 (79.1%)	50382 (80.1%)	57443 (83.4%)	57517 (83.1%)	38617 (77.3%)	26047 (44.9%)	17612 (28.1%)
Missing postcode	8009 (13.3%)	3161 (5.0%)	2279 (3.2%)	2909 (4.4%)	2932 (4.7%)	2451 (3.6%)	1996 (2.9%)	1497 (3.0%)	1976 (3.4%)	1966 (3.1%)

Continued overleaf

Table 7 continued

	2007/8	2008/9	2009/10	2010/11	2011/12	2012/13	2013/14	2014/15	2015/16	2016/17
Respondent										
Number	41907	43241	48691	46594	45040	49462	49346	34135	39240	42208
Missing forename	3 (0.0%)	1 (0.0%)	0 (0.0%)	2 (0.0%)	1 (0.0%)	2 (0.0%)	0 (0.0%)	1 (0.0%)	2 (0.0%)	3 (0.0%)
Missing surname	1 (0.0%)	1 (0.0%)	2 (0.0%)	4 (0.0%)	3 (0.0%)	5 (0.0%)	1 (0.0%)	1 (0.0%)	0 (0.0%)	1 (0.0%)
Missing gender	830 (2.0%)	441 (1.0%)	348 (0.7%)	279 (0.6%)	99 (0.2%)	74 (0.1%)	37 (0.1%)	201 (0.6%)	253 (0.6%)	285 (0.7%)
Missing date of birth	7270 (17.7%)	4696 (10.9%)	3514 (7.2%)	2768 (5.9%)	2030 (4.5%)	1549 (3.1%)	1251 (2.5%)	1195 (3.5%)	1283 (3.3%)	1304 (3.1%)
Missing ethnicity†	22510 (54.8%)	24908 (57.6%)	35535 (73.0%)	36222 (77.7%)	35276 (78.3%)	40037 (80.9%)	39826 (80.7%)	25585 (75.0%)	14308 (36.5%)	9722 (23.0%)
Missing postcode	5471 (13.3%)	2919 (6.8%)	2448 (5.0%)	2665 (5.7%)	2049 (4.5%)	1573 (3.2%)	1305 (2.6%)	1020 (3.0%)	1192 (3.0%)	1159 (2.7%)
Applicant										
Number	40180	42777	48081	45428	43820	48264	48374	34461	40501	43461
Missing forename	2 (0.0%)	1 (0.0%)	0 (0.0%)	2 (0.0%)	5 (0.0%)	0 (0.0%)	1 (0.0%)	1 (0.0%)	0 (0.0%)	2 (0.0%)
Missing surname	1 (0.0%)	0 (0.0%)	1 (0.0%)	0 (0.0%)	0 (0.0%)	3 (0.0%)	1 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Missing gender	753 (1.9%)	377 (0.9%)	230 (0.5%)	256 (0.6%)	104 (0.2%)	72 (0.1%)	52 (0.1%)	186 (0.5%)	186 (0.5%)	291 (0.7%)
Missing date of birth	1999 (5.0%)	812 (1.9%)	242 (0.5%)	193 (0.4%)	256 (0.6%)	197 (0.4%)	168 (0.3%)	236 (0.7%)	261 (0.6%)	259 (0.6%)
Missing ethnicity†	21601 (53.8%)	24264 (56.7%)	34980 (72.8%)	35440 (78.0%)	34420 (78.5%)	39343 (81.5%)	39302 (81.2%)	25934 (75.3%)	13831 (34.1%)	8181 (18.8%)
Missing postcode	2357 (5.9%)	1054 (2.5%)	716 (1.5%)	518 (1.1%)	565 (1.3%)	461 (1.0%)	441 (0.9%)	593 (1.7%)	815 (2.0%)	869 (2.0%)

There were 39 cases with missing start date. * Where local authority is missing, the local authority of residence could be inferred from the child's or carer's postcode. † Ethnicity became a mandatory field in July 2015.

Identifying fathers in Cafcass (and other administrative datasets) is difficult. In the study by Broadhurst and others,[6,20] the researchers were only able to identify fathers in 73% of cases. Identifying fathers is the subject of a current research study.[94] Further, no data are available on where the child was placed in care, only the legal order. Researchers must either infer the most likely placement outcome by legal order[6] or link with the children looked after dataset held by the Department for Education (Appendix IV).

It is likely that initial demographic data and court data at the date the case is opened on Cafcass' system are complete. For private law cases, incoming application forms are scanned centrally and the PDF copies are saved in the appropriate ECMS case folder. This data entry is completed by administrative staff at the start of the case. Practitioners and administrators are then responsible for remaining data entry while the case is ongoing and on case closure. Public law cases are opened on the system by local teams but Cafcass is currently piloting a central data input system similar to that used in private law cases. In addition, Cafcass carries out level 1 police checks in most private law cases; as these use addresses and the parties' identities, address and identity data are likely to be complete.

The ECMS does not allow a case to be closed without information on legal order being entered. However data user should be aware of the following:

- In the ECMS, this should be a final legal order but a non-final order (such as an interim care order) could be entered in error. In the old CMS, orders were attached to hearings and interim orders could validly be entered.
- In private law in particular, there will be a large proportion of cases with an unknown legal order because Cafcass' involvement ends, and the case is closed on the system, before the final order is made.
- The system does not check that the final order entered logically follows from the application. There could therefore be impossible combinations of applications/orders.

Finally, some data items may have limited completeness due to practitioner workload demands. An example of this is the 'case factor' variable, which captures information such as whether there are allegations of domestic abuse. Practitioners are responsible for entering these data but might omit to do so where there is limited administrative capacity. The data may be recorded within the case plan and case file, but may not be transferred into the relevant reporting fields; work is underway to explore whether the case plan can be integrated into ECMS to improve reportability. It might be possible to infer the presence of domestic abuse from data on domestic violence remedies in FamilyMan or from legal aid data, though this has not been tested.

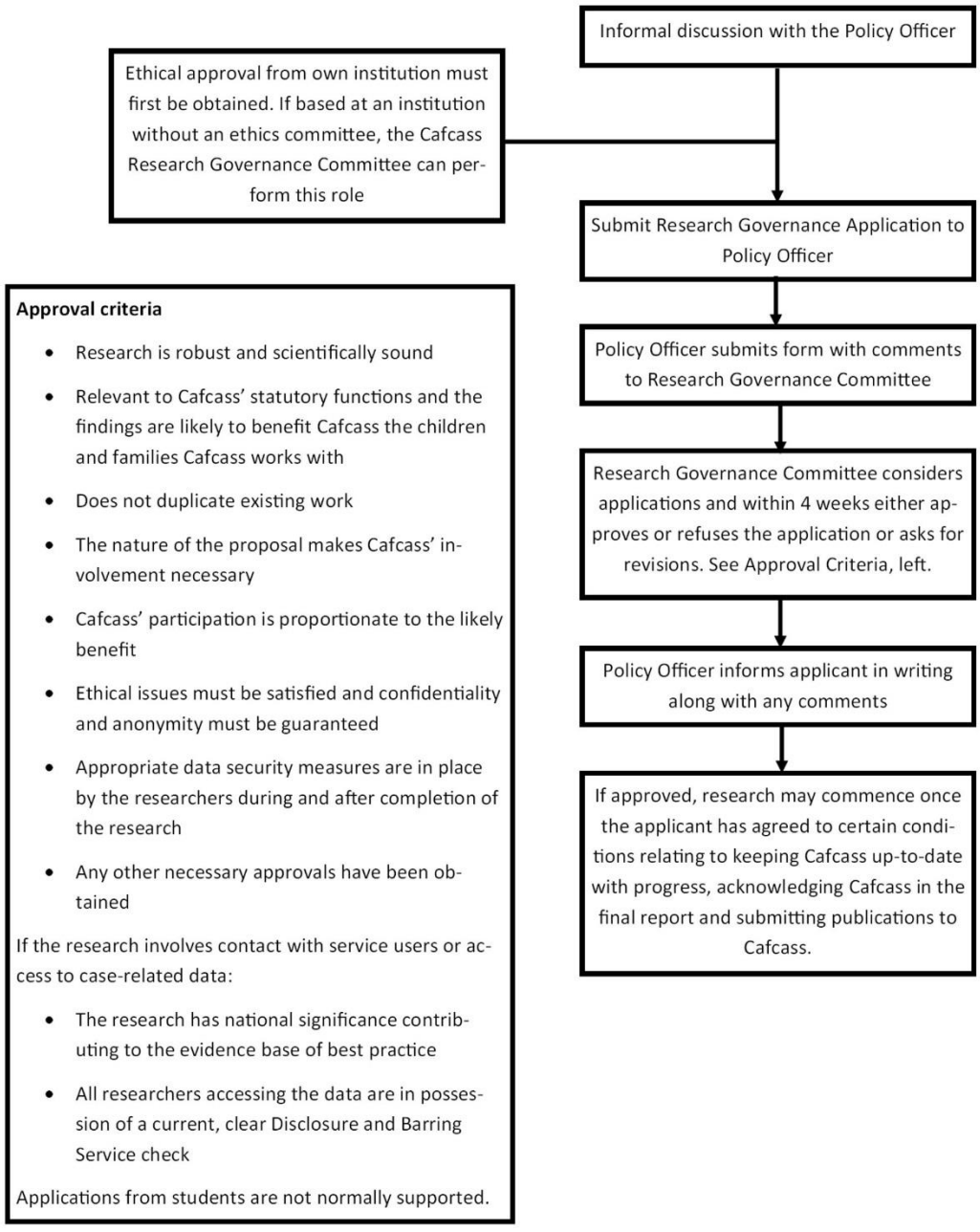
A1.2.3. Changes over time

Researchers must be mindful of changes in coding over time when analysing and interpreting their results:

- Cafcass data are only available from April 2007 but children and their families may have been in contact with Cafcass and the courts before then. For some cases, there will therefore be left-censoring (events prior to 2007 are hidden from the dataset) and this may affect incidence estimates. However, as time progresses, this will become less of an issue.[6,20]
- Significant changes to the Family Justice System in April 2014 (such as the introduction of child arrangement orders and the abolition of residence and placement orders) led to changes in the types of orders available and their coding, which may explain trends over time (i.e. changes are an artefact of the way data is recorded).
- Changes to the names and boundaries of local authorities must also be considered and may need post hoc rationalisation for analysis.[20]
- Ethnicity has always been available in the CMS and ECMS but only became a mandatory field in July 2015. This has reduced the proportion of missing ethnicity values significantly (Tables 6 and 7).
- The introduction of the ECMS resulted in variable names and table structure changes. For example, in the CMS, legal outputs were attached to hearings and interim orders as well as final orders were recorded; in the ECMS, legal outputs are attached to children and only the final legal order is documented. Further, open cases were migrated to the ECMS when it was introduced but they were left on the ECMS system. There is therefore some duplication across the two systems (identity variables are consistent across the two systems), but migrated cases are easily identifiable as they include a migration date. When working with data that span both the CMS and ECMS, researchers must be careful to ensure continuity between the two datasets.

A1.3. Permission pathway

Figure 11: Cafcass research approval process



Source: Based on information provided in the Cafcass Research Governance Framework.[122]

A1.4. Linkage

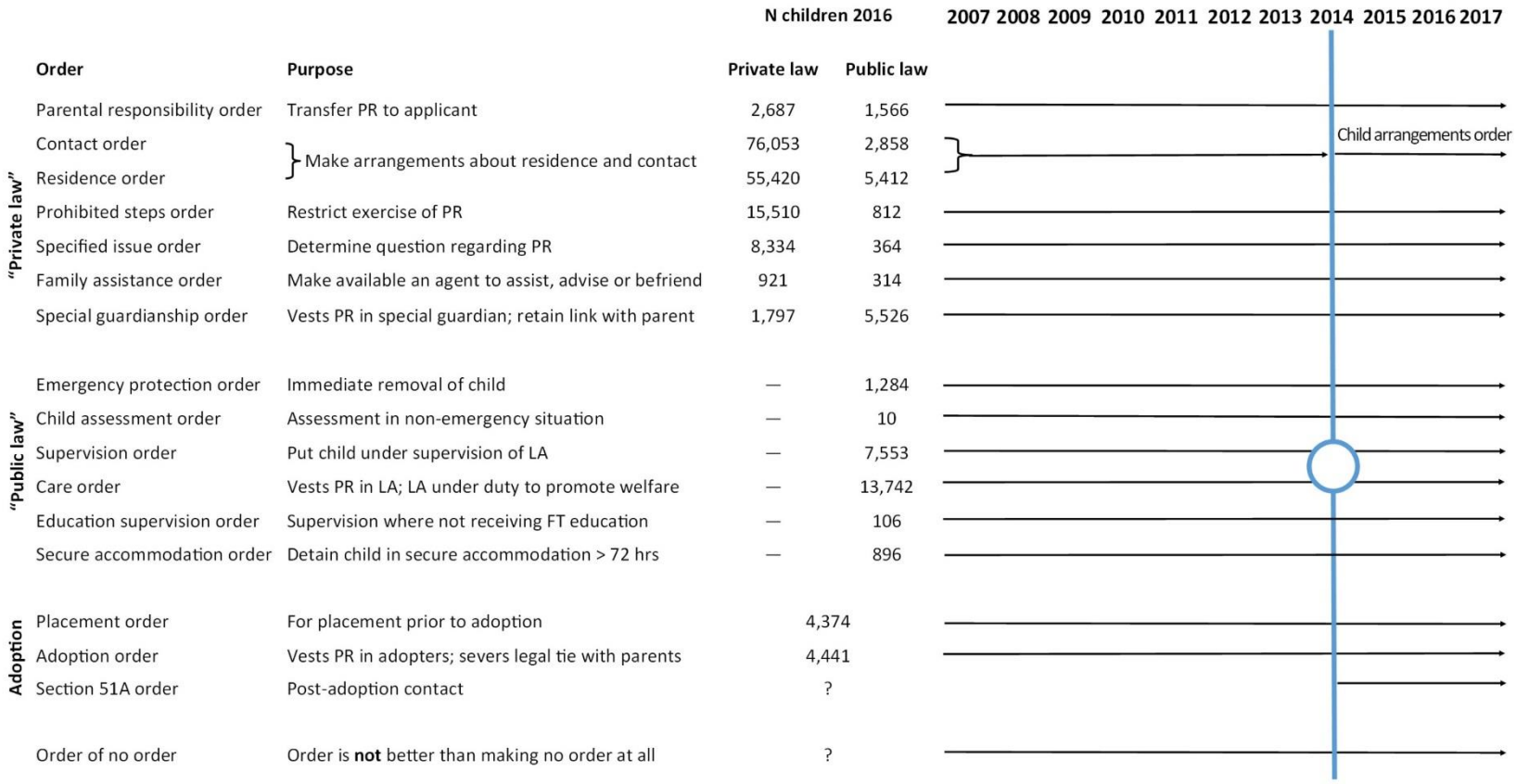
There are currently no routine linkages underway though a number of different linkages with Cafcass data are possible and in progress. The database includes a unique identifier for each person (whether child, parent or otherwise) and therefore events in the same individuals can be linked over time. Family members can also be linked together using the relationship identifier. This was exploited by Broadhurst and colleagues[6] in their study on recurrent care proceedings where they were able to link mothers to different children and show how likely it is that a mother will return to court following index care proceedings. Given the availability of identifiers such as name, date of birth and gender, it is also possible to link other administrative datasets. An example of this is the study underway by Wijlaars and colleagues[38] to investigate the health needs of children and mothers who go through care proceedings. This study will involve linkage between Cafcass data and the Hospital Episode Statistics via the Patient Demographic Service.[36] Cafcass data are also included in a data share project with the Ministry of Justice, which is detailed in Appendix II.

A1.5. Legal complexity and implications for research

An understanding of the legal orders and the legal system is essential for researchers to properly specify research questions and check the validity of the data. For example, a care order cannot be granted on an application for a section 8 order. However, it is possible to grant special guardianship orders or section 8 orders on applications for care orders. A child may be subject to multiple different orders at different time points. If a child is apparently subject to incompatible orders at the same time, this may be an error in the date field as well as an error in the way the order was recorded.

Data users should also be mindful of legal changes over time. Child arrangements orders, which replaced both contact and residence orders, were introduced on 22 April 2014 as were section 51A orders (similar to child arrangements orders but specifically for use for post-adoption contact). On that date the single family court was also introduced as well as the 26 week deadline in care cases. These factors may influence outcomes but will also change coding practices. Figure 12 gives an overview of the main orders over time as well as the number of children subject to each order in public and private law proceedings in 2016.

Figure 12: An overview of family law court orders



This diagram shows the main orders available to the court in determining family cases and their availability since 2007 (the year from which Cafcass data are available) as well as the number of children involved in these orders in 2016. No figures on section 51A orders or orders of no order are available on-line. Adoption and placement orders are not disaggregated by public or private law. PR: parental responsibility; LA local authority; FT full-time.

Source: Ministry of Justice Family Justice Statistics Quarterly (https://www.gov.uk/government/statistics/family-court-statistics-quarterly-october-to-december-2016, accessed 26 July 2017).

A1.6. Literature review of studies using the Cafcass data resource

Table 8: Summary of studies using record-level Cafcass data

Authors	Year	Overarching aim	Methods	Sample	Main findings	Data quality issues commented on	Linkage
Brandon et al[94]	In progress (2017-19)	To understand the role of fathers in recurrent care proceedings.	a) Quantitative analysis of Cafcass data on fathers in care proceedings and recurrent proceedings. b) A survey of nationally representative sample of fathers in care proceeding. c) In-depth qualitative interviews with fathers involved with care cases.	Care/supervision applications from 01/04/2007 to 31/03/2017.	Study in progress. At the time of writing (October 2017), an application to Cafcass for data access is pending.	There may be incomplete coverage of fathers[6,20] and limited information on the fathers (hence the survey and in-depth interviews).	Individual fathers within Cafcass data longitudinally.
Wijlaars et al[38]	In progress (2017-19)	To link Cafcass with Hospital Episode Statistics and Clinical Records Interactive Search data to understand health needs of children and mothers going through the care system.	a) Children will be followed for the first year of life and healthcare utilisation among those in case / not in care will be compared. b) Mums to be followed to identify outcomes after proceedings prior red flags. c) Examine pregnancy data to look at pregnancy spacing in repeat proceedings.	All children and mothers in Cafcass public law proceedings since 2007.	Study in progress.	Study currently underway. Will not be able to follow children after adoption as new identifiers are assigned to adopted children therefore children will be followed up only until the first year of life.	Linkage to Hospital Episode Statistics and Clinical Records Interactive Search.
Masson et al[124]	In progress (2017)	To explore the outcomes of care proceedings and whether these	A random sample of care cases from six local authorities is being examined before and after the introduction of the	Random sample of care cases from 6 local authorities.	Study on-going. Interim findings show decline in use of magistrates, possibly fewer hearings, variation in the	None commented on.	Cafcass, DfE and local authority data in the local authorities chosen.

Authors	Year	Overarching aim	Methods	Sample	Main findings	Data quality issues commented on	Linkage
		have changed before and after the Public Law Outline (PLO).	PLO. Data will be drawn from Cafcass, DfE and local authority files.		proportion of cases ending at the IRH, more kin assessments, fewer POs, more SOs and SGOs with SOs, more COs.		
Harwin et al[82]	2015	To describe trends in the use of special guardianship orders	Description of legal outcomes in child's public law cases.	All section 31 and placement cases between 01/04/2007 and 31/03/2015. Indexed on the child's case.	The proportion of cases ending with a special guardianship order has risen over the years, especially among children under 1. The use of special guardianship orders concurrently with supervision orders has also increased.	Limited data on child profile beyond age and gender.	None.
Broadhurst et al[6]	2015	To describe women's repeat involvement with care proceedings.	Mothers linked across time in dataset. Took first event between 2007-2011 as index event ('first' removal of child). Estimated a) % of women with subsequent removal 2007-2014 and b) variation by age at first removal and c) time between court proceedings	Care/supervision applications from 01/04/2007 to 31/03/2014.	24% of women return to court within 7 years – excludes voluntary out-of-home care placements. The proportion was higher for women who were younger (16-19y) at index event. For most repeat clients, court episodes happen within a short space of time (median interval 17 months), typically following birth of new child.	Some data entry errors in relation to end dates. Changes in coding over time (e.g., local authorities changing names). Missing data: some variables with large amounts of missing data excluded from analysis. 9% missing legal order data.	Individual mothers within Cafcass data longitudinally.
Rodger et al[125] (internal Cafcass research)	2015	To explore characteristics of parental orders.	Review of parental order (which transfer legal parentage from surrogate to commissioning parents) cases identified from the Cafcass data.	All parental order cases (smaller random sample selected for in-depth review).	189 parental orders made in 2013/14, 75% in relation to one child, 24% for 2 and 1% for 3. 79% for one male and	Some duplicate cases and a small number of cases incorrectly labelled as parental order applications.	None.

Authors	Year	Overarching aim	Methods	Sample	Main findings	Data quality issues commented on	Linkage
					female applicants and 21% for two male applicants.		
Trinder et al[126]	2013	To investigate why contact order cases return to court for enforcement.	Case review to profile enforcement cases (e.g. demographic details and reason for bringing an enforcement action); to ascertain the approach of the court; and to detail the outcomes of interventions.	1. All applications in the Cafcass case management system and electronic case files for enforcement orders in March and April 2012 (n = 205). 2. All other cases from November 2011 to October 2012 where the outcome was unpaid work (n = 10).	Most actions brought by non-resident fathers usually in connection with a contact order that had broken down or was only partially complied with. Most common approach (about 50% of cases) was co-parenting support; least common were punitive sanctions (<10% of cases). Very few cases resulted in unpaid work (more commonly used as a threat).	Does not include cases where contact orders were not being complied with and a party applied to vary it or cases where no enforcement application was made.	None.
Internal Cafcass research (anonymous)[127]	2012	To explore the extent to which Cafcass recommendation on s 7 (welfare) reports and court decisions on them are congruent.	Review of private law cases involving a request by the court for a s 7 reports, identified from the CMS.	170 randomly selected cases closed during June 2012 where is a s 7 report was recorded as having been requested by the court.	In 75% of cases was the recommendation of the Cafcass officer and the court's final decision congruent.	35 cases had 'significant' missing information (no detail given except that some included where no final legal order was in the electronic case file).	None.

Appendix II – FamilyMan data profile

The family courts make a major intervention in the lives of these families by helping to resolve significant disputes or crises through the application of the law. In this report, we are concerned with matters relating to the care of children. The FamilyMan database, held by the Ministry of Justice, which is the government department responsible for Her Majesty's Courts & Tribunals Service, captures information about family court cases.

A2.1. Overview

Data provider	The Ministry of Justice (MoJ)
Population covered	All cases going through the family courts (includes cases not covered in this report such as divorce petitions, domestic violence and financial remedies). Previously covered the county courts and Family Proceedings Courts that shared premises with county courts.
Size	• In 2016: 18,952 public law Children Act cases (31,375 individual children involved)[81] • In 2016: 48,244 private law Children Act cases (72,836 individual children involved)[81]
Data overview	Contains data on numbers of cases, applications and children involved in public and private law cases as well as court events and legal outcomes (interim and final orders). Names, gender, date of birth and postcode are available as is information on whether there is any harm alleged, whether the parties have representation and whether mediation has occurred.
Structure	FamilyMan is a relational database with separate tables for people and events.
Years covered	Data are available back to the late 1990s but much of this is retrospective. Prospective data are available from 2003 but prior to 2007 data for Family Proceedings Courts were weighted estimates based on a subset of courts. By December 2010, an administrative data system upgrade was completed and rolled out nationally. Therefore the highest quality data are from 2011 onwards.
Data entry	Data are entered by local court staff for case management purposes. They are transferred to the MoJ where a data team collates them into FamilyMan, which is updated once per month.
Permission pathway, ethics & costs	Record-level FamilyMan data have so far not been made available for research use.
URL	https://www.gov.uk/government/collections/family-court-statistics-quarterly (accessed 1 August 2017).
Linkage	FamilyMan data are part of the data share project with Cafcass and the Department for Education detailed below.

A2.2. Data completeness and accuracy

A2.2.1. Population

FamilyMan covers all cases going through the family courts. This includes cases not involving any children and not involving Cafcass. Such cases are outwith the scope of the present report. FamilyMan therefore covers the same cases as Cafcass but is broader. The overlap with, and linkage to, Cafcass data is detailed below.

A2.2.2. Validation checks and missing data

Date of birth and postcode are not always available or accurate meaning that linkage using these variables may be incomplete. The system accepts dates entered in an incorrect format (e.g. MM/DD/YYYY instead of DD/MM/YYYY) and so cleaning of date fields may be required. The postcode field is incomplete for some individuals. Some fields, such as gender, have drop-down boxes on data entry meaning that these fields are likely to be complete (though not necessarily 100% accurate). Some fields are free-text and the court records are stored electronically. This information contained in these is therefore theoretically available but may require significant amounts of preparatory work to make it research-ready; whether this is feasible will depend on the individual research project.

A2.2.3. Changes over time

Prospective data are available from 2003 but are of limited completeness or quality. Prior to 2007, data were weighted estimates based on a subset of Family Proceedings Courts. This is because different courts had different IT systems. A new administrative data system was rolled out fully by December 2010 meaning that the highest quality and complete data are available from 2011.

A2.4. Permission pathway

FamilyMan data are not currently available for external research use. The Ministry of Justice is currently exploring ways of sharing data for research use.

A2.5. Overlap with Cafcass data

FamilyMan covers all family court cases including those not involving children such as divorce and domestic violence. Therefore the population covered by Cafcass is a subset of that of FamilyMan. Cafcass however contains different data items to FamilyMan (e.g., Cafcass can record diversity data) and the Cafcass system has greater capacity for updates to its interface and data items captured. Cafcass is accessible to researchers with an established access pathway and research governance framework.

One major variable captured in FamilyMan and no longer captured in Cafcass is whether any interim orders are made, and which.[*] In the old CMS (prior to July 2014), legal outputs were attached to hearings and therefore interim orders could be recorded. In the new ECMS (from July 2014), they are attached to applications and only final orders are documented. Interim orders however continue to be captured in FamilyMan.

A2.6. Linkage

Individuals within FamilyMan, such as children or their parents, are identified using unique identifiers. It is therefore possible to link up children's separate applications and cases (applications and cases are also assigned unique identifiers) longitudinally and it is possible to link family members using a field that indicates relationship to child. Identifiers for both children and parents in FamilyMan also include names, date of birth, postcode and gender, though date of birth and postcode are not always complete.

FamilyMan is part of a data share project, conducted by the Ministry of Justice. This project linked Cafcass, FamilyMan, children looked after and education data held by the DfE. Linkage between Cafcass and FamilyMan was performed deterministically by the MoJ using full names, date of birth, gender and postcode. An initial match rate of about 60% was achieved but work has been conducted and this has improved to between 70% and 75% by using additional matching rules and ensuring comparability between the Cafcass and FamilyMan denominators. Matching between DfE data and FamilyMan, performed by the DfE, was successful for about 70% of children. Analyses on the linked

[*] An interim order is an order made during the course of proceedings before the final order is made. For example, an interim care order can be granted under the Children Act 1989, s 38, if the court has reasonable grounds for believing that the threshold criteria are met but more time is required (such as to further investigate the facts) before a final order is made.

data are on-going within the MoJ and work is planned to more thoroughly assess the quality of linkages.

A2.7. Literature review of studies using the FamilyMan data resource

FamilyMan record-level data are not readily available to researchers and we are not aware of any instances of their uses in academic research. Data are used by the MoJ internally and summaries are published on-line as the Family Courts Statistics Quarterly publications.[128]

Appendix III – Legal aid data profile

Legal aid, administered by the Legal Aid Agency, an executive agency of the Ministry of Justice, helps individuals fund legal advice and/or representation.

A3.1. Overview

Data provider	The Legal Aid Agency (an executive agency of the Ministry of Justice)
Population covered	Applications and certificates granted for criminal and civil legal aid (family and non-family) in England and Wales.
Size	• In 2016/17, there were 96,558 grants for civil legal help and controlled legal representation (CLR)* and 106,962 grants for civil representation. This is down from 761,583 grants for legal help in 2000/01 and 157,723 grants for civil representation in 2006/07. • Of the 162,995 grants for legal help and CLR in 2016/17, 35,143 (22%) were for family matters. Of the 106,962 grants for civil representation, 92,564 (87%) were for family. • 96,558 civil representation certificates were completed in 2016/17, of which 82,477 (85%) were family.[18]
Data overview	Demographics of claimants (names, gender, date of birth, ethnicity, address, postcode), applications for legal aid, area of law, outcome, expenditure. Includes mediation information and assessments and mediation sessions and information on family legal aid certificates granted through the domestic violence route. A data index is available (see URL, below).
Structure	A number of different data sources and databases comprise the legal aid statistics. These include the Corporate Information System, Contracted Work and Administration Database, Client and Cost Management System and standalone databases. The user guide (URL below) provides information on quality issues.
Years covered	Varies according to reforms. Some data are available from 2000/01 whereas others are only more recent.
Data entry	Some data are input by Legal Aid Agency caseworkers from paper and electronic files. Increasingly in more recent years, data are entered directly on-line by the legal aid providers (e.g., the solicitors or law centres providing legally aided advice) and this is mandatory in some areas. There are automatic and manual validation processes in place, including auditing by the National Audit Office.
Permission pathway, ethics & costs	Record-level legal aid data have so far not been made available for external research use.
URL	https://www.gov.uk/government/collections/legal-aid-statistics (accessed 1 August 2017).
Linkage	Caseworkers use some data linkage to verify data on individual cases but there are currently no routine, complete linkages underway.

* Controlled legal representation includes some immigration and mental health (i.e. non-family) cases.

A3.2. Data completeness and accuracy

A3.2.1. Population

Legal aid, which is administered by the Legal Aid Agency (LAA),[*] is funding available so that individuals can seek legal advice or representation for free or at a subsidised cost. It is available in public family law proceedings[†] regardless of financial circumstances (i.e., it is not means tested). In private family law, legal aid was available on a means-tested basis until 1 April 2013, after which it is only available where there is evidence of domestic abuse or a risk of such abuse. Legal aid is also available on a means-tested basis for mediation.

Civil legal aid, as available in family matters, can be broken down into three schemes: family help, civil representation and mediation. In summary, legal help includes advice and assistance and civil representation includes representation at court. Funding is also available for exceptional cases (cases not normally covered by legal aid) and for appeals, both of which are rare compared to standard work.

Applications for legal help and mediation are decided by the service providers (i.e., the lawyers actually providing the advice) who hold legal aid contracts within which each provider is allocated a number of 'matter starts' that are for them to administer. The LAA uses data submitted by the service providers for remuneration. It therefore holds data on successful applications for legal help and mediation but no data on unsuccessful applications. The LAA, by contrast, processes applications for civil representation and therefore holds information on successful and unsuccessful applications for this scheme.

Applications for legal aid can be made by adults and children. In public family law, all members of the family might be legally aided and will each have applications in the database. In private family law, it will usually just be the parents (who are the parties to the case), though children can be joined as parties in cases where it is required to do so to protect their welfare.

A3.2.2. Validation checks and missing data

Civil representation claims prior to October 2013 were entered by LAA caseworkers from paper and electronic files from the service providers. Since then, more applications were made on-line and, since

[*] The Legal Aid Agency came into existence on 1 April 2013 with the Legal Aid, Sentencing and Punishment of Offenders Act 2012. Prior to this, the Legal Services Commission (established 1 April 2000) carried out these functions, which itself was preceded by the Legal Aid Board.

[†] Referred to in the legal aid context as special Children Act proceedings, these include applications for emergency protection orders, child assessment orders, care orders and supervisions orders.

April 2016, this is mandatory. Most information is therefore entered directly into the system by the providers. Automated validation checks are applied and caseworkers manually check files to determine whether legal aid should be granted. An application must be on the system for legal aid to be granted and so the database is considered comprehensive.[129]

Data on legal help are submitted directly and electronically by the service providers. As with civil representations, a grant of legal help must be on the system to pay the provider; this is therefore likely comprehensive.[129] For mediation, data were entered by LAA staff from papers received from the mediation providers until February 2015 from which date a wholly electronic system is used. Again, the system is used for remuneration and reconciling provider payments. Data are subject to a quarterly revision process meaning that the latest quarter's figures may change.[129]

The LAA has a number of assurance processes in place including: annual duplicate removal; review of a sample of files to ensure that work claimed for was actually carried out; checking invoices against the provider's reported work; reconciliation of payments made against claims; and auditing overseen by the National Audit Office.[129]

The LAA publishes aggregate data on characteristics of legal aid applicants by scheme and law category (public or private).[18] These include the number of certificates where data are missing. The proportion of certificates with missing characteristics data in 2016/17 is shown in Table 9.

Table 9: Proportion of closed cases with missing characteristics data in legal aid dataset, 2016/17

Scheme	Variable	Private law	Public law
Civil representation (family law cases)	Ethnicity	24%	27%
	Disability	25%	35%
	Gender	1.8%	0.5%
	Age	~0%	~0%
Family legal help*	Ethnicity	21%	
	Disability	24%	
	Gender	0.7%	
	Age	~0%	
Mediation involving children	Ethnicity	4.4%	-
	Disability	4.8%	-
	Gender	~0%	-
	Age	0.5%	-

* Data on legal help are not available broken down by law type.

Finally, a valid postcode is missing in about 35% of civil representation cases. It may be possible to infer the postcode using the rest of the address combined with information about the service provider (whose postcodes are complete) but this has not so far been tested.

A3.2.3. Changes over time

Significant changes to the legal aid system, particularly as regards private family law, will mean that the types of cases covered will differ over time and this may affect how data are coded. Prior to the coming into force of the Legal Aid, Sentencing and Punishment of Offenders Act 2012 on 1 April 2013, legal aid was available in both public and private family law. In the latter case, eligibility was subject to a means test. Since 1 April 2013, legal aid has only been available (subject to a means test) in private family law if there is evidence of (or evidence of a risk of) domestic violence or child abuse. Legal aid is also available for family mediation. It may be possible to use legal aid data as a crude marker of socioeconomic position in cases prior to 1 April 2013 and after then as a marker of domestic violence (though both of these possibilities are currently untested).

Prior to December 2014, the type of domestic abuse in private family law was not correctly recorded in some cases. The recording of this was rectified from then. This means that data on the type of abuse in private family law cases prior to December 2014 are not reliable. However, overall figures (whether there was domestic abuse or not) is considered robust, as is the disaggregation from December 2014.[129]

Finally, there have been changes to the databases themselves which means that care must be taken to ensure comparability of variables between the old and new systems.

A3.3. Permission pathway

Legal aid data are not currently available for external research use. The Ministry of Justice is currently exploring ways of sharing data for research use.

A3.4. Linkage

Some data linkage is used by caseworkers for the purposes of verifying details on individual cases but there are currently no routine linkages with the whole dataset.[129] Within civil representation matters, data are available on applications, certificates granted and cases closed. There are demographic data of the clients: name, address history including postcode, gender, ethnicity and date of birth. For legal help, data are available on matters started and completed and identifiers include name, date of birth,

gender, ethnicity and postcode. The same identifiers are available for mediation, which includes data on mediation assessments, starts and outcomes.*

The possibility of linkage is being explored within the LAA. The quality of client postcodes is poor which may hinder linkage.[129] Similarly, there is a high proportion of missingness on ethnicity (Table 9) meaning that this variable will be of limited utility in linkage algorithms. It is possible to link individuals using an assigned unique identity number (the civil representation system is designed to track individuals, with cases linked to an individual applicant).

A3.5. Literature review of studies using the legal aid data resource

To date, record-level data have not been used in academic research. The data have been mainly used for routine statistics and internally.[18]

* Assessment refers to the initial meeting to determine whether mediation is feasible. If so, mediation is started and will result in an outcome of full agreement, partial agreement or no agreement.

Appendix IV – Children looked after and child in need data profile

The local authority is central in the lives of children looked after and children in need. In the former cases, the local authority has responsibility to look after children where there are no parents or those with parental responsibility are unable to do. It also has jurisdiction to bring care proceedings in the family court. In the latter cases, the local authority is responsible for the provision of services to maintain the welfare of children in its area. Data generated from these services is sent to the Department for Education where they constitute the children looked after (CLA) and children in need (CiN) datasets.

A4.1. Overview

	Children looked after (CLA) dataset	**Children in need (CiN) dataset**
Data provider	The Department for Education (DfE)	
Population covered	All looked after children and recent care leavers. Does not cover children in private fostering arrangements.[130]	All children in need including children on child protection plans. It also includes data on children referred to children's social care where no further action was taken.[131,132]
Size	• On 31 March 2016 there were 70,440 looked after children.[70] • 32,050 children started to be looked after in 2015/16.[70] • 3.3% of children born between 1992 and 1994 have an episode of care by the time they are 18 years old and this rate is rising with more recent birth cohorts.[75]	• In 2015/16 there were 621,470 referrals (just over 5% of all children).[133] • There were 401,600 children starting an episode of need.[133] • There were 394,400 children in need on 31 March 2016.[133]
Published data resource profile	Mc Grath-Lone and colleagues.[86]	None.
Data overview	Unique Pupil Number where assigned (since 2005/6), child characteristics, local authority, date entered care, reason for being in care, legal status (e.g. care order or s 20), placement type, provider changes and date and reason care ended. For children looked after for at least 20 working days, the date and child's participation in the statutory review are recorded. For children in continuous care for	Unique Pupil Number where assigned, child characteristics, local authority, referral date and dates of assessments / section 47 enquiries, dates of child protection conference and plan, referral source (since 2013/14) and referral reason. Also includes data on factors which are a concern for the present case, e.g. substance

	Children looked after (CLA) dataset	Children in need (CiN) dataset
	more than 12 months Strengths and Difficulties Questionnaire data and other outcomes are also available (since 2009). Information on recent care leavers (since 2002). Data on children placed for adoption are also collected.	misuse, mental health, learning or physical disability, abuse/neglect, trafficking, missing, gangs, self-harm and antisocial behaviour.
Structure	Each distinct period of care for each child is an episode. "Periods of care" are constituted by one or more continuous episodes and may include one or more placements and/or one or more legal statuses. A period ends when a child ceases to be looked after, or if the type of care changes from respite to non-respite (or vice versa), even if these episodes are continuous.	The dataset has four modules: child identifiers, child characteristics, children in need and child protection plans.[131] Data can be analysed on a child or case basis.
Years covered	1992 onwards. Between 1998 and 2003 individual-level data were only collected on children with a birth data divisible by 3, giving a one-third sample for those years. Aggregate data were collected for all other children (the CLA100 return). Not all data items have been collected throughout all years.	2008/9 onwards. Not all data items have been collected throughout all years.
Data entry and storage policies	Data are submitted to the DfE by local authorities. Guides on CLA[130] and CiN[131] datasets for local authorities contain details on data security, uploading, amendments, validation and definitions for each data item. Technical specifications are also available.[132,134]	
Permission pathway, ethics & costs	DfE data are assigned to one of four tiers of sensitivity (tier 1 being the most sensitive). All CLA and CiN data are tier 1. Researchers must apply on an electronic application form and their application will be considered by the Data Management Advisory Panel. Access to each data item must be justified. Details are available in the User Guide (URL below). Currently, the DfE does not charge for access to this dataset or for linkage with external datasets.	
URL	For the User Guide and meta data tables: https://www.gov.uk/government/publications/national-pupil-database-user-guide-and-supporting-information (accessed 1 August 2017). For data access applications: https://www.gov.uk/guidance/national-pupil-database-apply-for-a-data-extract (accessed 1 August 2017).	
Linkage	CLA data are routinely linked by the DfE to the National Pupil Database data from 2006 and to CiN data from 2009. All three have been linked for research.[40] CLA data have been linked as part of a data share project with the Ministry of Justice and Cafcass but this is not currently available to researchers: see Appendix II, section A2.6, for more details.	

A4.2. Data completeness and accuracy

A4.2.1. Population

CLA dataset: The population covered by the CLA dataset is restricted to children who are looked after, whether through court-mandated means or section 20 accommodation, and therefore includes a broader population of looked after children than Cafcass. It does not include children subject to private fostering arrangements or children receiving local authority support under section 17 of the Children Act, who are children in need (these are contained in the children in need census, see below). For asylum seeking children, a distinction is drawn between accompanied asylum seeking children (who are treated as all other children and who may or may not become looked after) and unaccompanied asylum seeking children (USAC). USAC are further distinguished by whether they are accommodated by the local authority and looked after (in which case they are included) and those who are being given financial support under section 17 for accommodation (in which case they are not). Details are available in the local authority guide.[130]

CiN dataset: Data on all children assessed by the local authority as being in need under section 17 of the Children Act are held in this dataset. This includes children looked after and children on child protection plans. The dataset also includes data on referrals to children's social care but where no further action was taken. Although the Act defines a child as any person under the age of 18, local authorities are asked to return data on individuals aged 18 or over if they are receiving care and accommodation or post-care support from children's services.[131] The return should include all children even where provision for them has been made by adult social care teams. Children on the local authority's disabled children register are only included in the CiN returns if the presence of the child's name on the register is a trigger for any action by children's social care services. If not, for example if the family receives a newsletter only, then the child should not be included in the data.[131]

All looked after children are by statutory definition children in need. Therefore, the CLA population is a subset of the CiN population and children in the former dataset should all appear in the latter dataset. Sebba and others,[40] however, noted in their analysis that there was a small discrepancy between the two datasets with some children appearing in the CLA dataset but not the CiN. They report that numbers were too small to affect inferences.

A4.2.2. Validation checks and missing data

When data are uploaded by local authorities for the CLA dataset, the DfE system imposes a number of logic checks; these are detailed in full in 115 pages of the technical specification[134] and include

checks for missing data, unlikely or impossible combinations of legal status and placement; unlikely or impossible sequences of dates; and information which contradicts that already held. For example, the system cannot record blank genders or ethnicities (though the latter can be recorded as unknown) and dates must be in a valid format (DD/MM/YYYY). Logic checks ensure, for example, that where a new episode has started, any previous episode must have finished and the end date recorded. Records that fail validation checks are flagged for action by the local authority. Local authorities are also able to amend records from 2003/4 (the DfE can amend records from before then on the request of the local authority). Checks are also performed on the CiN dataset and errors and queries are flagged for the attention and action of the local authority.[132]

Children who enter school are captured in the National Pupil Database (NPD) and are assigned a Unique Personal Number (UPN). If a child with a UPN enters care, their UPN is also recorded in the CLA return (since 2005/6) and in the CiN dataset. However, because about 38% of children who enter care each year are under 5,[70] a significant proportion of children in the CLA dataset will not have a UPN.

Data are not collected on parents or other family members meaning only child-level analyses are possible unless linked to other datasets. It is also not possible to identify sibling groups.

The DfE data collection system does not allow missing values but certain conventions are used where information is not known:[131,134]

- In the CLA dataset, gender must be recorded as either male or female. In the children in need data, a gender value of 0 can be recorded for unborn children or where the gender is unknown and a code for indeterminate gender where a gender has not yet been assigned.*
- If date of birth is not known then local authorities are instructed to provide an estimate. In the CLA return, local authorities are specifically asked to use the 15th of the estimated month and year.
- Ethnicity should be reported by the child or, where this is not possible, the parent or carer. Local authorities are permitted to enter codes for information not yet obtained or refused.
- Children without a UPN are assigned a special code indicating the reason for not having one.

* This can occur, for example, where a child is born with a disorder of sexual development that makes it impossible to assign a gender at birth.

A4.2.3. Changes over time

CLA dataset: Data collection commenced in 1992 and has been on-going since.[86] However, between 1998 and 2003 complete record-level data were collected on children with a date of birth divisible by 3 only (aggregate data were collected for other children). This means that only one third of the population is covered between those dates and any analysis requiring longitudinal follow up over that period will have to be restricted to that third.[86]

In addition, not all variables have been available for all years. Most child characteristics and episode information (e.g. legal status, local authority, start and end dates of each episode) are available from 1992 (subject to the one-third restriction noted above). A notable exception is ethnicity which has only been collected since 2001. Individual-level adoption and care leaver information are available from 2004 and individual-level outcomes data for children in care for continuous periods of 12 months or more are available from 2009.[86]

Finally, in the 2017/18 data collection, a new variable relating to whether the child has returned to care after or during a previously permanent placement is being collected.[130] The local authority is required to indicate whether the previous arrangement was adoption, special guardianship, residence order or child arrangements order or unknown. This might make it possible in the future to examine the trajectories of adopted children who are otherwise assigned new identities though whether this is truly feasible is yet to be determined.

CiN dataset: The CiN census was first implemented in 2000 and repeated in 2001, 2003 and 2005, after which it was suspended and re-introduced in 2008.[115] This report only reviews the census from 2008. Most variables in the CiN dataset have been collected continuously since 2008/9 when data collection began. Referral date is only recorded from 2012/13 and referral source from 2013/14. Number of previous child protection plans and the start and end dates of the plans were originally collected in 2008/9, discontinued, and collected again from 2012/13.

In both datasets, data users should be mindful of changes to local authority structure over time (e.g. the amalgamation of local authorities). Analyses taking into account local authority may therefore need rationalisation before analysis commences.

A4.3. Permission pathway

The pathway for data access for the CLA and CiN uses an application form which is common to the NPD. Users should consult the User Guide.[39] The application will be considered by the Education Data

Division. Data extracts are available in four tiers depending on sensitivity, with tier 1 being the most sensitive. Requests for access for tier 1 data will be triaged and dealt with by the Data Management Advisory Panel (senior members of the department). All CLA and CiN data are considered tier 1. Access to each tier 1 data item must be individually justified. The Education Data Division / Data Management Advisory Panel will ensure that information governance standards are met and that the processing of data is lawful, secure and in line with department and governmental standards. If approved, the researchers must sign a licence agreement and provide individual declarations. Access to CLA data also requires a valid Disclosure and Barring Service certificate.

A4.4. Linkage

The datasets contain a child ID which means that children can be linked longitudinally within the datasets. However, a new child ID is assigned if a child moves local authority. This is likely a very small fraction but it is not known how many children move whilst not being looked after and subsequently enter care in another local authority. Linkage across local authorities requires the UPN, which is only recorded for children who have entered school or pre-school. New identifiers are also assigned when a child is adopted making it difficult or impossible to follow these children reliably.[*]

The CLA and CiN datasets are routinely linked with each other and with the NPD, using the UPN where this has been assigned, by the DfE. The NPD and CLA dataset also formed part of the Ministry of Justice data share project detailed in Appendix II, section A2.6.

Data on home and placement postcode are collected in the CLA dataset and are used to derive distance between home and placement, local authority and placement location (within or outside local authority).[130] However, there are concerns about the quality of these data. This means that children cannot be linked to a lower-layer super output area (LSOA) and then to the Index of Multiple of Deprivation.[135] Nor are data collected on individual-level socioeconomic circumstances. The LSOA and postcode are, however, captured in the NPD so any child with a UPN can be linked to an LSOA and postcode to obtain an area-based measure of deprivation.[†] The DfE ran a consultation between April and July 2017 on linking educational outcomes with parental income data from the Department for Work and Pensions and Her Majesty's Revenue & Customs. At the time of writing (October 2017), the

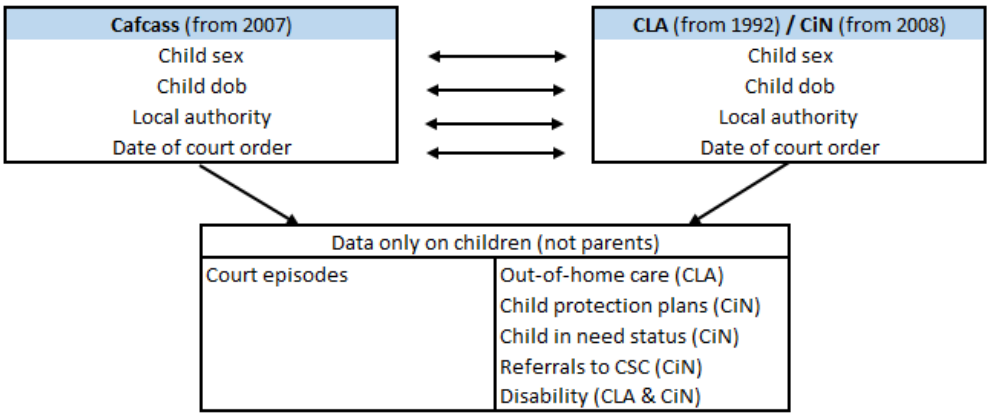
[*] As noted above, from the 2017/18 data collection, it might be possible to track looked after children who have been previously adopted. This would rely on accurate date of birth and gender identifiers and it is not yet clear whether this truly is feasible.

[†] An exemplar project using Scottish data on looked after children, education and health has been conducted and was found to have good linkage rates by using the Scottish Candidate Number (the Scottish equivalent of the Unique Pupil Number).[136]

DfE are analysing the responses.[48] This may provide an alternative measure of socioeconomic position that may also be linkable to the CLA and CiN datasets.

New linkages may also be possible. Key variables in the CLA and CiN data are the child's sex, date of birth and local authority. This combination uniquely identifies 87% of the children in the CLA dataset (about 97% who also have a care order in Cafcass, if using date of court order to match as well). Potential linkage, which has so far not been tested, between Cafcass (which contains postcode and, by extension, local authority), CLA and CiN is represented in Figure 13.

Figure 13: Potential linkage between Cafcass, CLA and CiN data



CSC: Children's social care

It is not known whether linkage using sex, date of birth and local authority would be sufficient when linking to the Personal Demographic Service and Hospital Episode Statistics. This is an area requiring further research.

A4.5. Data items not returned

All local authorities collect additional data items which are not returned nationally. These data could be used to augment analyses of national data, could be used in place of national data[14] (though sample sizes will be reduced and if authorities are not selected at random the results may be biased) or could be considered for inclusion in future returns. See the report by Holmes and others[3] for more information.

A4.6. Literature review of studies using the CLA and CiN data resources

Table 10: Summary of studies using record-level CLA and CiN data

Authors	Year	Overarching aim	Methods	Sample	Main findings	Data quality issues commented on	Linkage
The Children's Commissioner[54]	2017	An exploratory first-step in the development of a Stability Index that can be used for monitoring and improvement stability of placements.	Using CLA data, education data and data on social worker changes supplied by 22 local authorities, a stability index was created and analysed.	All children looked after in the year 2015/16; all children in the 2015/16 Autumn and Summer School Censuses; and all children in care in the 22 local authorities providing data on social work changes.	An estimated 71% of children experienced a change in the placement, school or social worker over a 12 month period. 45% experienced a change in one measure, 19% in two measures and 5% in all three measures.	Some changes in placement may be positive and planned and the data do not disclose this. There may be unmeasured confounders of which there are no data in the datasets.	CLA linked with education data and data supplied by 22 local authorities.
Mc Grath-Lone et al[74]	2017	To describe rates of re-entry to care and risk factors.	Rates of re-entry to out of home care for children exiting care between 2007 and 2012. Regression modelling used to identify risk factors. Used CLA data only.	Data on all episodes of care (CLA data) since 1992 for children with birthday divisible by 3 (due to data collection restrictions).	There were decreases in re-entry (23% in 2007 to 14% in 2012) and changes in type of exit from out of home care (e.g, greater use of special guardianship orders in all age groups). Age, ethnicity, reason for care, previous history of care, placement length, changes, being in voluntary care and reason for exit were all associated with re-entry to care.	Cannot examine re-entry of adopted children as they are given new identifiers when adopted. No detail on whether planned or unplanned exit from care and limited detail on other characteristics (e.g. health and family composition). Not possible to evaluate consequences for the child of re-entry.	Individual children's episodes of care.
Bewley et al[137]	2016	To estimate the impact of participation in the Troubled Families	Using data from the Troubled Families Programme linked to national administrative data, the authors used statistical methods	25% of families who participated in the programme and a comparison group of	There was evidence of a slight decrease in the proportion of children in care after 12 months in those who were in	Not all local authorities supplied Troubled Families data (we are aware that other local	Troubled Families data were linked to children's social care,

Authors	Year	Overarching aim	Methods	Sample	Main findings	Data quality issues commented on	Linkage
		Programme on outcomes such as whether the child was in need or in care 12 months after starting the programme.	to estimate the impact that participation in the programme had on various outcomes.	families who fell below the eligibility criteria of the programme.	the programme but no consistent evidence that the programme had an impact on the proportion of children in need. There was no evidence of other systemic impacts of the programme on other outcomes.	analyses and linkages are being conducted but results are not widely available). Many individuals did not link to the national data; such non-linkages may be correct but this was not certain.	education and other datasets.
Mc Grath-Lone et al[86]	2016	CLA data resource profile	N/A	N/A	Contains information on the CLA dataset's coverage, variables, structure and other factors.	Restriction of data between 1998 and 2003 to 1/3 sample; local ID is assigned by each local authority; new identity for adopted children; no detail on support received; names are not collected.	N/A
Mc Grath-Lone et al[75]	2016	To describe the cumulative incidence of entry into care, how this varies over time and by ethnicity.	Cumulative incidence of entry into care calculated for children in age bands and by broad ethnic group. Data on placement types and duration described. Used CLA data only.	Data on all episodes of care (CLA data) since 1992 for children with birthday divisible by 3 (due to data collection restrictions).	3.3% of all children born between 1992 and 1994 entered care at least once before the age of 18; this proportion is rising with time. Rates of entry into care varied significantly by ethnicity: black and mixed raced children being more likely to enter care than white and Asian children.	1/3 sample. Limited data on type of care received.	Individual children's episodes linked in order to calculate cumulative incidence.

Authors	Year	Overarching aim	Methods	Sample	Main findings	Data quality issues commented on	Linkage
Sebba et al[40]	2015	To identify key care and educational factors that are associated with the progress of children in care.	Analysis of linked education, CLA and CiN data, supplemented by in-depth interviews with 26 young people who have been in care for at least 12 months.	All children in the National Pupil Database, CLA database and CiN database.	Children not in need or care performed the best at Key Stage 4. Early entry into care was associated with better performance at Key Stage 4 compared with later entry into care and children in need. School and placement changes are associated with poor outcomes.	No data on foster or residential carers, details of school and placement practices and instability. Missing data from some schools and local authorities, especially on the Strengths & Difficulties Questionnaire.	National Pupil Database, CLA and CiN datasets.
Ubbesen et al[102]	2015	To quantify the risk of entering care by age and how this varies by time and between Denmark and England.	Using CLA data, age-specific cumulative incidence of ever entering care during childhood in Denmark and eight local authorities in England was calculated.	All children entering care in eight local authorities and all of Denmark from 1992 and 2008.	Incidence decreased over time in Denmark but increased in England.	English analysis limited to 1/3 sample. The Department for Education only provided data from eight local authorities which they deemed representative.	None.
Wade et al[138]	2014	To describe the characteristics of guardians children subject to special guardianship orders, long-term outcomes and key issues in policy and practice.	Analysis of the CLA dataset as well as a national survey of local authorities, a survey of special guardians and in-depth interviews with children and other stakeholders.	Children who exited care via a special guardianship order between 2006 and 2011 (n=5,936).	From the analysis of the CLA dataset: Of the 5,936 orders, 55% were for children less than 5, a quarter for those aged 5-9 and the rest for children 10+. The gender split was equal. Three quarters of the children were white (lower than, for example, those who are adopted).	The authors note that the data say little about the experiences of children.	None.
Selwyn et al[139]	2014	To establish the rate of adoption disruptions post-	A denominator of all adoptions was obtained from the CLA dataset between 2000 and 2011.	All adoptions between 2000 and 2011.	From the analysis of the CLA dataset: 565 adoptions out of 37,335 (1.5%) were known to	Because adopted children are assigned new identifiers in the	None.

Authors	Year	Overarching aim	Methods	Sample	Main findings	Data quality issues commented on	Linkage
		order and explore the characteristics and risk factors of disruptions.	To determine the rate of disruptions, the authors conducted a survey of local authority and voluntary adoption agency managers to identify post-order disruptions. The study also featured a survey of families and interviews.		have been disrupted. Identified risk factors for disruption included older age at placement and longer time between the order being made and placement.	CLA dataset, it is not possible calculate disruption rates using that dataset alone, hence the survey. There is likely underreporting of disruptions as these were based on knowledge of local teams, which may be incomplete.	
Department for Education[140]	2013	A data pack to provide detail about placements and placement stability.	Descriptive analysis of the number of placements and possible risk factors using the CLA data.	All looked after children in 2011/2012.	Of all looked after children, 67% had one placement, 22% had two placements, and 11% had three or more placements in the year ending 31 March 2012. Children with more placements are less likely to achieve five good GCSEs than children with a single placement.	None commented on.	CLA data linked with education data in National Pupil Database.
Schofield et al[141]	2007	To describe continuity and disruption to placements in children in long-stay care.	Analysis of CLA data on children who had been in care for 4 years or more ('long-stay'), supplemented by analysis of published aggregate data and a survey completed by social worker.	1,002 looked after long-stay children as of March 2001 in 24 local authorities.	Half of the sample were in local authority foster placements and the rest were in other types of placement (e.g., placed with parents, family or friends, residential care or placed for adoption). There was variation in the proportion of children who had remained in one	Analysis limited to 1/3 sample. Only 324 questionnaires (on 1,002 children) were returned. Questionnaires were necessary to capture more detailed data on social work practice. The authors comment that the outcome measure	None.

Authors	Year	Overarching aim	Methods	Sample	Main findings	Data quality issues commented on	Linkage
					placement for two years from about 33% to 90% by local authority.	used says little about the experiences of children.	
Dickens et al[142]	2007	To describe variation in rates of children going into care and possible causes.	Analysis of CLA data supplemented by analysis of published aggregate data and a survey completed by social workers.	Children in 24 local authorities who started to be looked after between October 2000 and March 2001.	There was variation in the rates of children going into care (in the sample authorities, 11 to 41 per 10,000 children). This appeared to be associated with deprivation, turnover and management control though these associations are likely confounded.	Analysis limited to 1/3 sample.	None.

CiN child in need; CLA children looked after.

Appendix V – Useful resources

Below are some useful resources relevant to the use of administrative and research data resources. This list is non-exhaustive and omits references to materials requiring specialist subject matter knowledge (such as on substantive legal rules or statistical methodology).

Administrative Data Research Network (ADRN)

The ADRN is a network of UK universities, government, national statistical authorities, funders and research centres with centres in England, Wales, Scotland and Northern Ireland. It delivers a service providing secure and lawful access to de-identified linked administrative data for research. The ADRN website (https://adrn.ac.uk/) contains a host of information (some of which is referenced below) on the use of administrative data and the services the ADRN offers.

ADRN Frequently Asked Questions

https://adrn.ac.uk/public-engagement/pe-across-the-network/faq/

More information about the ADRN and the use of administrative data.

ADRN Glossary

https://adrn.ac.uk/public-engagement/pe-across-the-network/glossary/

Definitions of key terms relevant to the use of administrative data.

ADRN YouTube channel

https://www.youtube.com/user/ADRNUK

A selection of videos about administrative data and the ADRN.

Information Commissioner's Office on anonymisation

https://ico.org.uk/for-organisations/guide-to-data-protection/anonymisation/

Information published by the Information Commissioner's Office on anonymisation and its official Code of Practice.

Guidance for Information about Linking Datasets (GUILD)

Published in the Journal of Public Health[96] and accessible via https://academic.oup.com/jpubhealth/article/doi/10.1093/pubmed/fdx037/3091693/GUILD-GUidance-for-Information-about-Linking-Data

Guidance about the information that needs to be made available about the data linkage process, by data providers, data linkers, analysts and the researchers who write reports.

National Pupil Database (NPD) User Group

http://www.bristol.ac.uk/cmpo/npd-user-group/

The NPD User Group supports users of the NPD and is hosted by the Centre for Market and Public Organisation, University of Bristol. It was formerly known as the PLASC/NPD Users' Group (PLUG).

Secure Anonymised Information Linkage (SAIL) Databank

https://saildatabank.com/

A safe haven enabling data linkage and analysis of administrative data.

UK Data Archive

http://www.data-archive.ac.uk/

A repository of research data sets such as the Millennium Cohort Study and the Longitudinal Study of Young People in England as well as much more.