Towards a Family Justice Observatory
A scoping study

Using population level data to understand the family justice system
Report from a knowledge exchange event

Authors: Jenny Woodman[1], Karen Broadhurst[2] and Ruth Gilbert[3]

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

Acknowledgements: We would like to thank the delegates who presented, attended, discussed and commented on the report. Delegates are listed in Appendix 2.

Please cite this as: Woodman, J., Broadhurst, K., Gilbert, R. (2017) *Using population-level data to understand the family justice system: Report from a knowledge exchange event*. Available at: http://wp.lancs.ac.uk/observatory-scoping-study/

Background and purpose of event:

This report encapsulates key messages arising from an all-day event, held on Monday 30th January at the Farr Institute of Health Informatics Research, London. The event was part of a wider scoping study which will propose an organisational model ('an observatory') for improving the generation and use of research evidence within the family justice system, funded by The Nuffield Foundation, led by Professor Karen Broadhurst at the University of Lancaster. Existing population level data is one of a number of evidence resources being considered to produce a step-change in the quantity and type of research available to inform policy and practice within the family justice system. Population level data is a powerful source of evidence but also presents analytic, governance and ethical challenges. This event brought together members of this research community, including representatives from data providers in order to:

- identify opportunities and challenges for research relevant to the family justice system within existing population-level data
- identify ways forward and key points to consider when trying to increase the amount of good quality research using population-level data in the field
- gather ideas about how an observatory might support members of the research community who are using population-level data to answer questions relevant to the family justice system

The day consisted of a series of short talks by academics who presented their research using population-level data, with a focus on linking datasets and the opportunities and challenges of the data and the linkage. The programme of talks is included as **Appendix 1**. Each talk was followed by questions and discussion and there was an open discussion at the end of the day, focusing on the issues raised during the day and the role that an observatory could play for researchers.

Delegates:

The event brought together academics using administrative, survey or birth cohort data (population-level data) in research relevant to the family justice system as well as representatives from The Ministry of Justice, NHS Digital and Cafcass (Children and Family Court Advisory and Support Service). We included speakers and delegates from across disciplines, all with expertise in collecting or using population-level data to understand trajectories of vulnerable populations in and out the family courts, out-of-home-care, the criminal justice system, the education system and employment. Speakers and delegates were identified via their research, which was already known to members of the large and cross-disciplinary research team. See **Appendix 2** for list of delegates.

Our rationale for the broad and cross-disciplinary approach was a view of the family courts as one event in the complex life trajectories of parents and children: we are interested in using population-level data to understand and improve the family justice system, including the way it intersects with other public services.

This report provides a summary of the key points from the day, concentrating on one overarching message that came out of discussion: building a family justice community by drawing together diverse and disparate researchers, data providers and service planners and providers who are working in a broad range of disciplines and services relevant to the family justice system. This report is a synthesis of key themes from the day and incorporates feedback from delegates on a draft copy.

1. Key message: share learning and expertise within a broad family justice community and make coordinated case for improved data and access

Need to establish a 'family justice' community of researchers, data providers and data users in order to avoid duplication of effort, allow comparison of different approaches to data cleaning and analysis and ensure that research builds strategically and logically over time.

There was one overarching theme of the discussion and debate: an appetite and need to build a working community and network of researchers, local data users and data providers all with an interest in using population-level data to a) understand and improve the family justice system and/or b) understand the groups of children and families who have contact with the family courts or are at high risk of contact and c) identify opportunities for improvement in service provision and policy for these groups across all public sectors. Delegates envisioned this as a broad cross-disciplinary group to include those working in research and policy for children and families who do not necessarily go through family courts, such as children who enter care through non-legal routes or children and families living with adversity (only some of whom will have contact with the family courts). This broad focus will contextualise the working of the family courts and provide important comparator groups. Establishing this broad type of 'family justice' community might address some of the following challenges, as identified by delegates:

- **Keeping up-to-date with relevant research and data projects**
  Research which is relevant to the family justice system and which uses population-level data is being carried out across the country in a diverse set of teams and disciplines. Relevant data improvement and/or linkage projects are also on-going within a broad range of government departments, local authority settings and organisations. The breadth and dispersion of this work makes it difficult for researchers and funders to keep abreast of current knowledge and on-going research and projects. Some of the delegates were not aware of each other's work before attending the seminar. Lack of a common electronic database for published research in FJS was identified as a factor. Keeping abreast of current knowledge ensures that work is not duplicated and that new projects build on those which have already been undertaken.

- **Understanding data quality and meaning, especially at the research planning stage**
  Researchers could share learning about data quality within population-level datasets (% of complete data in a variable, % of meaningful data and changes in data quality over time) as well as changes in the structure of the data and/or definition of variables. For example, there have been key changes to Children Looked After dataset (SSDA 903) over the last few years which require some between-year mapping. This learning could be published as responses or adjuncts to data resource profiles, where these existed. Important commentary on the meaning of specific data items could be included in data profiles or appendices to papers and reports. For example, to use the Children Looked After dataset researchers need to have an understanding of what it *means* that a child leaves care with a special guardianship order to a foster carer. Sharing expertise on quality and meaning of variables would help other researchers plan studies and also facilitate a coordinated effort by researchers to work with data providers to improve the quality of key variables and suggest new data items. Data, data providers and governance (see point below) will be different across the four UK countries.

- **Negotiating permission pathways to access data**
  There was agreement that getting permissions to use the data from e.g. Department for Education or NHS Digital was a complex and time-consuming matter. Getting the right data within the right timescale and at the right price was a major barrier to successful research using population-level data. The process could take months or years and cost tens of thousands of pounds (e.g. data extracts from NHS digital, Hospital Episodes Statistics).One research team in the room only discovered that they needed a separate data application for specific variables within datasets held by Department for Education once they had received their dataset (e.g. pupil referral unit data or Strengths and Difficulties Questionnaire (SDQ) scores). The researchers then had to apply again and wait for these variables. These researchers would be willing to share such learning with others in the field. It might also be possible to share learning relating to ethical issues around using administrative data for research, such as opt-in/opt-out mechanism for consent.

- **Making administrative data‘ research-ready’ (cleaning data and deriving variables)**

Making administrative datasets ‘research-ready’ is time-consuming. This involves cleaning the data and/or restructuring and then group and classify data points into meaningful variables ready for analysis. For example, researchers using the Cafcass data described how they spent considerable effort restructuring the data to map the previous and current data collection system onto one another before starting analysis. Approaches to establishing a start and end point for care proceedings within the Children Looked After data was given as another example of data preparation, needing a expert level of knowledge about legal routes in and out of care. This expertise, learning and code (e.g. SPSS, STATA or R syntax) could be shared. Researchers should properly acknowledge and attribute the large amount of work that has gone into the code they reuse.

- **Understanding process of linking datasets and the implications for results**
  Understanding the linking process is crucial for interpreting the results but much of this linkage is currently done in a black-box within government departments. A research community could make a coordinated case for wider access to existing linkage algorithms from government departments and research teams. Such sharing would allow researchers to compare the implications of different approaches to linkage, illustrated with specific research questions. There may be a way that an observatory can facilitate this sharing by, for example, acting as a repository or as the hub of a research network.

- **Linking family members within and across datasets (including data on dads)**
  There was agreement that we need to understand trajectories of children AND their parents in order to understand the point at which family members have contact with the family courts and other services. However, it is currently not easy to identify family units within datasets. It is possible to link mothers and children through birth records within hospital data (Hospital episode statistics) and also within the Cafcass data. Data collection does not always reflect the importance of fatherhood to men and children. For example, the Children Looked After data contains information on when a young women in care becomes a parent but not equivalent data for young men in care. There was widespread agreement in the room that better data about fathers (more complete,

more detail) was crucial and that researchers could usefully make a case for this and investigate ways of reliably linking fathers (biological and social) to mother and child pairs.

- **Re-use of linked datasets**
  Linking datasets is a highly skilled and time-consuming process but current permissions to link datasets are given only for a specific research study. A key example is the Troubled Families Evaluation which linked administrative data from children's social care, education, the police and employment and pensions to evaluate the impact of the Troubled Families Programme on family members. This is a very rich linked national dataset which is currently unavailable for re-use outside of the Department for Communities and Local Government. When government departments link data for their own purposes, there is currently no obligation or incentive for them to make arrangements so that the wider research community might be able to re-use this linked data. Such arrangements would need to be made at the planning stage of linkage projects so that it is covered in data sharing agreements.

- **Establishing a denominator (comparison) population**
  Researchers need a denominator population against which to compare children in out of home care and families who come into contact with the family courts. This allows researchers to describe how characteristics in these groups differ from that of the general population or differently defined 'at risk' groups. Establishing a denominator or comparator population is very difficult. For example, the National Pupil Database (NPD) contains data on school aged children but we know that a quarter of children who experience out of home care will do so only before their fifth birthday (i.e. before they start school).[1] This problem can be partially addressed by using the 'early years' data within the NPD, which contains educational data for all 2, 3 and 4 year olds in settings that receive direct government funding.

- **Public and professional engagement**
  Public engagement has been a key factor in the success of cohort studies such as the Avon Longitudinal Study of Parents and Children (ALSPAC). A community of family justice

researchers could tackle the important issue of translating research-findings for lay readers and professionals and promote evidence-based policy, particularly targeting messages at different users of research.

2. The role of an observatory for researchers

- One possible function of an observatory might be to create a virtual family justice research community and lead or coordinate on the issues described above: the sharing of on-going and finished projects, research findings, expertise and learning (including through searchable databases), making the case for research access to linkage processes and re-use of linked national datasets. It will be important to consider how best to properly acknowledge academic ‘behind the scenes’ work that is shared.

- The observatory might support the development of specialist data safe havens across the country, with (hopefully) permissions and infrastructure for in-house linkage of administrative datasets. Attention will have to be paid to differences between the four UK countries in terms of public services, data providers and the law.

- An observatory might take on a role of public and professional engagement or support researchers in engaging professionals and the public.

- The observatory needs to be complementary to but distinct from the ‘what works?’ centres.

There were several concrete suggestions that could be taken forward as a starting point, within or outside an observatory:

- Funders within this field could investigate how to improve the indexing and searchability of all research literature relevant to the family justice system, with automated indexing. There are likely to be important lessons to be learned from the socio-legal community in America.

- Create a group to make the case for anonymized re-use of existing linked national data, for example those linked by government departments (eg Troubled Families Dataset).

- Create user groups for researcher using a) Cafcass and b) the Children Looked After and Child in Need data, based around the model of the Department for Education data users group (PLUG) which focuses largely on NPD users. PLUG has an annual meeting which is key to its success. Such a group could include researchers, data providers and possibly data users (policy-makers and practitioners).

- Researchers write up and publish their descriptive results concerning data quality and linkage for use by other researchers
- Data providers, such as Cafcass, provide a list of data applications on their website so researchers can keep up-to-date with on-going studies.

APPENDIX 1: Programme of Presentations

Each talk lasted 15 minutes (20 minutes for paired speakers), followed by 10 minutes of discussion

Speaker	Title
Judy Sebba Nikki Luke	The Educational Progress of Looked After Children in England: Linking Care and Educational Data
Louise Mc Grath-Lone	Factors associated with re-entry to care: analysis of administrative social care data
Helen Baldwin Linda Cusworth	Home or care? Examining child outcomes through the linkage of administrative, cohort and primary data
Andy Boyd	Linking ALSPAC to diverse administrative data: initial findings and the quest for complete coverage
Helen Gray	Evaluating the impact of the Troubled Families Programme using linked national administrative datasets and local programme data
Stuart Bedston Bachar Alrouch	Studying S31 care proceedings and their families: reshaping the Cafcass case management system
Linda Wijlaars	Opportunities and challenges of using health data for family justice research: examples from hospital and primary care data
Liz Trinder	Private family law cases: Private law children research and the tyranny of paper files and legacy systems

APPENDIX 2: List of speakers and delegates, ordered alphabetically by affiliation and then surname

Affiliation	Name	Area of expertise	Most relevant population-level datasets used*
Academic			
Bristol, School of Social and Community Medicine	Andy Boyd	Data linkage, Information Security, Research Governance, Research Infrastructure	ALSPAC linked to diverse administrative data including from health, social care, police.
	Alison Teyhan	Epidemiology. Health and educational outcomes	ALSPAC linked to CLA and CIN
Bristol, Law School	Ludivine Garside	Socio-legal studies, economics, social medicine.	CLA and CIN, Cafcass, primary care data
	Judith Masson	Socio-legal studies. Family law and child law. Child welfare.	Cafcass and local authority social care data
Cardiff, School of Social Sciences	Jonathon Scourfeld	Social work, child protection practice, social inequality	CLA, CIN, local authority social care data, cohort and panel studies including ALSPAC
Coventry, The Centre for Technology Enabled Health Research	Paul Bywaters	UK child welfare systems, child protection, Looked After Children, social inequalities, data linkage	Local authority social care data, CLA, CIN, Index of Multiple Deprivation
Exeter, School of Law	Liz Trinder	Socio-legal studies. Private family law cases.	Cafcass (at local level)
Institute for Employment Studies	Helen Gray	Policy evaluation. Employment and welfare to work, training and skills, criminal justice and social policy interventions.	NPD, PNC, WPLS (linked)
Lancaster, Sociology	Bachar Alrouch	Quantitative social science. Public family law. Information systems.	Cafcass
	Karen Broadhurst	Social Work. Family justice system. Inequalities.	Cafcass
	Stuart Bedston	Quantitative social science. Family justice system.	Cafcass, HES

Continued overleaf

Key to datasets on page 14

Affiliation	Name	Area of expertise	Most relevant population-level datasets used*
Lancaster, Mathematics and Statistics	Brian Francis	Social statistics. Quantitative criminology.	Crime survey
London School of Hygiene and Tropical Medicine, Department of Health Services Research and Policy	Katie Harron	Statistics. Linked electronic healthcare data for child and maternal research	Primary care data, HES and other health data
Loughborough, Centre for Child and Family Research	Lisa Holmes	Child welfare and social work	Local authority social care data
	Harriet Ward	Child welfare, social work and policy	Local authority social care data, CLA, CIN, section 251 expenditure data
Oxford, Rees Centre for Adoption and Education	Nikki Luke	Social welfare. Fostering, education, mental health.	CLA, NPD
	Judy Sebba	Social work. Fostering, Looked After Children. Adoption.	CLA,NPD
York, Social Policy and Social Work	Helen Baldwin	Applied health and social research. Child health and welfare, criminal justice and substance misuse.	CLA, CIN, BiB
	Linda Cusworth *also affiliated to Lancaster University*	Quantitative social research. Child wellbeing and child outcomes, particularly mental health and educational attainment, for children in and on the margins of care	CLA, CIN, BiB, NPD

Continued overleaf

Key to datasets on page 14

Affiliation	Name	Area of expertise	Most relevant population-level datasets used*
University College London, Administrative Data Research Centre	Ruth Blackburn	Public health. Health and social care services. Infectious disease control	Primary care data, HES,
	James Doidge	Epidemiology and Public Health. Evidence-based public policy, including child health and social determinants of health. Data linkage.	NCP, data from Ministry of Justice, prisons and probation services, HES, NPD
	Ruth Gilbert	Clinical epidemiology. Child health and welfare. Linked administrative data	HES, primary care data, Cafcass, NPD, CLA, CIN data from Ministry of Justice and Police.
	Matthew Jay	Social epidemiology, chronic pain, social welfare and EU law	HES
	Louise Mc Grath-Lone	Public health. Child welfare.	CLA, NPD
	Linda Wijlaars	Epidemiology. Child health. Health services. Family justice.	Primary care data, HES, Cafcass
	Jenny Woodman	Public health and service provision for vulnerable families. Combining administrative and qualitative data	Primary care data, CIN
Non-academic			
Cafcass*	Jigna Patel	Data content, quality and access for family court data held by Cafcass	Cafcass
Ministry of Justice	Ross Black	Intelligence. Data linkage.	MoJ data
NHS Digital	Dickie Langley Garry Coleman	Data governance, permissions and access for data held by the Department of Health.	Administrative health data
Nuffield Foundation (funders)	Teresa Williams	Improving the generation of research evidence to inform family justice system and decision making	

Office of the Children's Commissioner	Emily Emmott	Evidence informed policy	Children and families, child development, children's social care using Local authority youth offending team data, Local authority social care data, CLA, NPD, MSC, Crime survey, NCDS
	Leon Feinstein	Director of Evidence at OCC	

See overleaf for key to datasets

* KEY TO DATASETS, Appendix 2
ALSPAC: Avon Longitudinal Study of Parents and Children. Birth Cohort study in Bristol area, held by Bristol University
BiB: Born in Bradford
Cafcass: Data from the family courts in England, held by Children and Family Court Advisory Support Service (Cafcass)
CIN: Child in Need data. Social care data from England, held by Department for Education
CLA: Children Looked After. Social care data from England, held by Department for Education
HES: Hospital Episode Statistics. Data from hospital admissions in England. Held by NHS Digital
NCDS: National Child Development Study
NPD: National Pupil Database. School data from England, held by the Department for Education
MCS: Millenium Cohort Study
Crime survey: Survey in England Wales about crimes, included those that go unreported.
PNC: Police National Computer. Data on cautions, arrests and convictions in England and Wales, held by Police
Primary care data: data from GP consultations in the UK
WPLS: Work and Pensions Longitudinal Study. Data on employment and income from tax records in England, held by Department for Work and Pensions.

References

1. Mc Grath-Lone L, Dearden L, Harron K, et al. Factors associated with re-entry to out-of-home care among children in England. Child abuse & neglect 2017;**63**:73-83.