# **Vulnerable Birth Mothers and Recurrent Care Proceedings**

## Estimating Recurrent Care Proceedings using CAFCASS Administrative Data A Technical Appendix

Dr Bachar Alrouh and Professor Karen Broadhurst Department of Sociology – Lancaster University

## Introduction

This technical Appendix explains the methodology that we have developed and implemented in order to restructure administrative data held by the Child and Family Court Advisory Service, to produce new datasets for research purposes. The objectives of the study were to derive a first estimate of recurrent care proceedings and understand the profiles of mothers and children, caught up in those proceedings.

The Children and Family Court Advisory and Support Service (CAFCASS) Case Management System (CMS) is a relational database (SQL Server Database), that can be managed and manipulated using SQL-based (Structured Query Language) Systems (e.g. Crystal Reports or Microsoft Business Intelligence Development Studio).

The CMS database is a national electronic system that allows managing and monitoring of the workload of family court proceedings across all the Cafcass teams in England. Cafcass uses CMS to produce monthly, quarterly and annual statistics on family law cases using family law applications as the unit of measurement. The CMS has been used since early 2007, but Cafcass has stopped entering new cases into the CMS and moved to an updated system (ECMS) in July 2014.

## **CMS Tables and Relationships**

The CMS has a huge number of tables that are used at different stages of the case management process. The first step was to identify, from the CMS complex structure, the main data tables and variables that can be used to answer the team's research questions.

The focus of the 'recurrent care' study was on S.31 proceedings under the Children Act 1989, the parties to those proceedings, and case outcomes. Therefore, we focused on 8 main data tables that can be used to identify these units of analysis. Figure 1 illustrates these data tables and their relationships.

Note: Table 1 explains the different shapes and relations that are used in the figures throughout this report.

Shape	Meaning		
Table N {Filter X}	The records in Table N are filtered according to the criteria specified in {Filter X}		
$\mathbb{N} \longrightarrow \mathbb{M}$	Each record from table N is linked to one or more records from table M		
N M	Each record from table N is linked to zero or more records from table M		
N M	Each record from table N is linked to only one record from table M		
N M	Each record from table N is linked to zero or one record from table M		



#### Figure 1: CAFCASS CMS Database Structure

#### 1. Case Table (tblCase)

This table includes all the cases (i.e. sets of proceedings) that the Cafcass case workers have been involved in. This table includes various case management variables. The main variables that were required for this project were:

- CaseID: A unique identifier code given by Cafcass for each case.
- LawTypeRef: Public or private Law.
- LocalAuthorityRef: The local authority of the Case.
- DateCaseCompleted: The date of completion for the whole case.

## 2. Application Table (tblCaseApplication)

This table that includes all the applications, where one or more applications can be linked to the same case. The variables we used from this table were:

• CaseApplicationID: A unique identifier code each application.

- CaseRef: A code that identifies the overall case.
- CaseApplicationTypeRef: The type of the application (e.g. Care, Supervision, Placement, etc...).
- CourtRef: The court where the application was lodged.
- DateOfApplication: Start date of the application.
- DateCompleted: End date of the application.

## 3. Person Table (tblPerson):

This table contains personal details of all the individuals who are part of the family court system. The variables we used from this table were:

- PersonID: A unique identifier code each person.
- DateOfBirth: The date of birth of the person.
- Gender: Male, Female or unidentified.
- EthnicityRef: The ethnicity of the person.
- ReligionRef: The religion of the person.
- FirstLanguageRef: The first language of the person.
- InterpreterReq: Whether an interpreter is required or not.
- Disability: Whether the person has a disability or not.
- AccessSupport: Whether the person has access support or not.

## 4. Relationship Table (tblPersonRelationship)

This table identifies the relationships between the individuals from the Person Table and the type of relationship. The variables we used from this table were:

- PersonRelationshipID: A unique identifier code for each relationship.
- ParentRef: The ID of the parent.
- ChildRef: The ID of the child.
- CRSRelationshipRef: The type of relationship (e.g. Mother/Father, Step, Adoptive, etc...).

## 5. Member Table (tblCaseApplicationMember)

This is a linking table that links an *Application* to a *Person*. A person can be linked to one or more applications, and an application can have one or more persons.

The variables we used from this table were:

- CaseApplicationMemberID: A unique identifier code for each member.
- CaseApplicationRef: The ID of the application.
- MemberRef: The ID of the person who is a member in the application specified in the previous variable.
- IsParty: whether the member is party to the proceedings.

## 6. Hearing Table (tblCaseHearing)

A table that contains all the hearings linked to the case. The variables we used from this table were:

- CaseHearingID: A unique identifier code for each hearing.
- CaseRef: A code that identifies the overall case.

## 7. Legal Output Table (tblLegalOuput)

A table that includes all the legal orders. The variables we used from this table were:

- LegalOutputID: A unique identifier code for each legal Output.
- CaseHearingRef: A code that identifies that hearing where this order was made.
- LegalOutputTypeRef: The type of the legal order (e.g. Care, Placement, Supervision, etc...)

• IsFinal: Whether this legal order is final or interim.

## 8. Resolution Table (tblMemberResolution)

This table links *Legal Outputs* to the *Application Members*. A member can have one or more legal orders in the as an outcome of the same application (e.g. care and placement, supervision and residence, etc...), and a legal order can be for one or more members (e.g. care order for 2 siblings in the same application).

## **Inclusion Criteria (Filtering the Records)**

Only records that met the following inclusion criteria were extracted from the CMS database:

- Case criteria:
  - Public law case.
  - Completed between 01/04/2007 and 31/03/2014.
  - At least 1 application in the case.
- Application criteria:
  - The application type is either S31. Care, or S.31 Supervision or Extension of Supervision Application.
  - Application started between 01/04/2007 and 31/03/2014.
  - At least 1 birth mother and 1 child in the application.
- Mother Criteria:
  - The person gender is *Female*.
  - The member type is Adult Respondent.
  - The member is *Party* to the application.
  - $\circ$   $\;$  The person is birth parent for at least 1 child in the application.
- Child Criteria:
  - The member type is *Child*.
  - The member is *Party* to the application.
- Father Criteria:
  - The person gender is *Male*.
  - The member type is *Adult Respondent*.
  - The member is *Party* to the application.
  - The person is birth or step parent for at least 1 child in the application.
- Legal Output Criteria:
  - The legal output type is one of the following:
    - Adoption Order (AO).
    - Placement Order (PO).
    - Care Order (CO).
    - Special Guardianship Order (SGO).
    - Residence Order (RO).
    - Supervision Order (SO).
    - Order of No Order (ONO).
  - The legal order is *Final*.

Figure 2 illustrates the structure of the query used to identify the records of interest and the inclusion criteria.



Figure 2: Restructured CMS Tables and Records Inclusion Criteria

## **The Research Relational Database**

The extracted records were exported into a Microsoft Access Database where the tables were restructured and simplified to focus on the mother as the main unit of measurement. The new structure was based on linking each mother to all her applications, and then linking children and partners (fathers/stepfathers) to the mother's applications (See Figure 3). Where a mother had an application that started on the same date, these were aggregated together (combined in one application). The size of the sample resulting from this process is described in Table 2.



#### Figure 3: Recurrent Care Project Data Structure (Microsoft Access Database)

Sample	Size		
Unique Mothers	43,635		
Mother's Applications	51,933		
Children in Mother's Applications*	88,977		
Of which, Unique Children	84,714		
Fathers in Mother's Applications*	40,247		
Of which, Unique Fathers	36,491		
Table 2 Decision of Comparison Comparison Comparison			

 Table 2: Recurrent Care Project - Sample Size

\* A child/father can appear in one or more applications

The legal outcomes were aggregated per child in a mother's application to create a new variable that represents the child's combination of legal outcomes when necessary (e.g. care and placement). All the different combinations were then categorised into one of the following 4 categories:

- Supervision/Order of No Order
- Residence/Special Guardianship
- Care

Adoption/Placement

<u>Note</u>: the initial analysis of the legal outcomes produced results that were not consistent with other studies. For example, the percentage of adoption and placement orders were far below expectations. Therefore, we extracted all additional subsequent applications (adoption, placement, special guardianship and residence applications) that were linked to the children included in our dataset, but were issued ('start date') after the end of the original cases. The outcomes of those additional applications were aggregated and linked to the child's original S.31 application, where applicable, and the legal orders of the child at the end of the S.31 application were adjusted to reflect any additional orders.

For example, a child was subject to a "Care Order" at the end of a S.31 care application. A new placement application started after the end of the care application and resulted in a "Placement Order". The legal order for the child was adjusted to "Care and Placement Orders" to reflect the outcomes of the linked applications.

## **The Research Dataset**

The final data restructuring concerned linking all the mother applications and ordering them chronologically as illustrated in Figure 4. This structure enabled data analysis to be carried out using the *mother* as a main unit of measurement.

Given the focus on s.31 proceedings, we had captured some cases where the application trajectory started with an extension of supervision order – these were excluded from the sample (n=94, 0.2%). As we describe in the linked published paper, reliable data was only available between 2007 and 2014, hence we were not necessarily capturing the start of mothers' legal 'careers'. The updated sample size is shown in Table 3:

Sample	Size	
Unique Mothers	43,541	
Mother's Applications	51,789	
Children in Mother's Applications*	88,694	
Of which, Unique Children	84,539	
Fathers in Mother's Applications*	40,136	
Of which, Unique Fathers	36,411	

Table 3: Recurrent Care Project – Updated Sample Size

\* A child/father can appear in one or more applications



#### Figure 4: Linked Episodes (SPSS dataset)

The youngest child in each application was identified as the index child in that application. The father of the youngest child in an application was identified as the index father in that application. Data Aggregation was then applied per mother for analysis purposes.

## **Data Quality Issues**

Administrative datasets do afford large sample sizes and this data is economic to export, however, the use of administrative data requires careful consideration of quality issues. Data entry errors, missing data, system upgrades and changes in the coding of the data over time must all be considered in appraising quality. This section explains how we have dealt with some of these issues.

#### **Data Inaccuracy**

Most of the data in the CMS were manually entered into the system by case workers through various interfaces and forms. This can give rise to data entry errors, particularly in relation to date variables illustrated in applications with end dates that were earlier than start dates or mothers and children with impossible dates of birth. We have converted illogical values for the dates into missing data where applicable.

Recording children's legal outcomes was another example of variables that contained illogical data. In a negligible percentage of cases, the combinations of legal orders did not always make sense (e.g. final placement and supervision orders for the same child in an application). Such cases were dealt with by prioritising the most likely legal order(s) combination.

#### **Changes in Data Coding Over Time**

The CMS has been used by Cafcass over a number of years and there were changes in how data values were coded (categorised). Old values were not always updated to reflect these changes, which led to some codes/categories being used in some years but ignored in others. This issue was very obvious in regard to the local authority variable. There were 266 different codes for English local authorities in the CMS, while in reality there are only 152 local authorities in England. This was because some local authorities have changed their names, others were merged together or even divided. Thus, some rationalisation of local authority data was required. Cases that included historic codes in our dataset were recoded using the current 152 current codes by matching, when possible, the old name with the new one, or recording it as missing when no possible match was available. We then categorised all

the 152 local authorities into 10 regions using the Department for Education *Outcomes for children looked after by local authorities* (<u>https://www.gov.uk/government/statistics/outcomes-for-children-looked-after-by-local-authorities</u>).

## **Missing Data**

The level of missing data for each variable was assessed, and variables with a large proportion of missing data (over 50%) were excluded from analysis. Ethnicity, religion, first language and disability variables had a huge amount of missing data (See Table 4) and therefore excluded from the analysis. As for the main variables that we included in the analysis, the date of birth was missing on 8% of the mothers, and around 1% of the fathers and children. The child's legal order data was missing on 9% of the concluded applications. Missing data on all the other main variables was negligible (less than 1%). Our analyses also included some calculated variables (e.g. ages, length of application, number of children/fathers in an application, etc.), which were derived from the main variables or from the relationships in the database. The level of missing data on these variables is clearly stated in the main article.

	Mothers	Fathers	Children
Sample Size	43541	36411	84539
Date of Birth	8%	1%	1%
Gender	Х	Х	0%
Ethnicity*	58%	65%	60%
Religion*	62%	69%	64%
First Language*	59%	66%	62%
Disability*	94%	97%	98%
	Mother's	Applicatior	15
Sample Size	51789		
Application Start Date	0%		
Application End Date	0%		
Application Type	0%		
Local Authority	0%		
	Child's Applications		
Sample Size	88694		
Legal Order Per Child Per Application	9%		

Table 4: Main Variables and Missing Data \* denotes excluded variables