

# Identifying and Understanding Inequalities in Child Welfare Intervention Rates: Quantitative Methodology Technical Paper

Paul Bywaters, Geraldine Brady, Chantel Jones, Tim Sparks, Claire McCartan, Lisa Bunting, Jade Hooper, Marina Shapira, Brigid Daniel, Martin Elliott, Jonathan Scourfield

August 2017

## Contents

### Introduction

- Purpose
- Forms of Intervention

### Data Collection

- National Children's Services Data Collections
- Census date
- Content
- Neighbourhoods
- Deprivation
- Population
- Local Authorities
- Sampling

### Ethics and Access

- Ethical approval
- Access

### Cleaning

#### Final Sample

- England
- Northern Ireland
- Scotland
- Wales
- Spread of deprivation within each nation
- Spread of Deprivation across the UK
- Ethnicity

#### Aligning country data

- Child abuse and neglect categories
- Placement types
- Legal Categories

### Analysis

- Data Files
- Syntax – cross country consistency
- Outputs and rate calculations
- Statistical tests
- Adjusted rates
- Cross country comparisons
- Individual country analysis
- Detailed ethnicity analysis
- Religious background

## References

[Appendix 1: Categories of maltreatment in official data in the UK](#)

[Appendix 2: LAC Legal Status Coding](#)

## Introduction

This paper outlines and explains the quantitative data methods employed in the project entitled ‘Identifying and Understanding Inequalities in Child Welfare Intervention Rates: comparative studies in four UK countries’ funded by the Nuffield Foundation, 2015-17 ([www.covntry.ac.uk/cwip](http://www.covntry.ac.uk/cwip)). It is designed to give a detailed overview of the processes involved in collecting, cleaning and analysing with the quantitative data across the four UK countries.

## Purpose

The over-arching purpose of the quantitative elements of the project was to be able to compare state child welfare intervention rates and patterns *between* the four UK countries: England, Northern Ireland, Scotland and Wales, and between areas *within* the countries. The project was designed to enhance knowledge of the role of family socio-economic circumstances as a key factor in children’s chances of involvement with social care services. But it was also interested to examine the interaction of deprivation with other factors: aspects of identity (gender, age and ethnicity); the types of abuse and neglect that are reflected in child protection interventions; and placement types and legal status of ‘looked after children’.

This was done by integrating information from three main kinds of administrative data sets:

- Child welfare statistics, collected by local authorities (or Health and Social Care Trusts in Northern Ireland) and reported to national governments, on children who have contacts with children’s social care services
- Child population data, reported in the periodic UK-wide census and updated in the form of mid-year estimates by the Office of National Statistics
- Deprivation data, reported in each of the four UK countries in the form of an Index of Deprivation at a variety of levels of geography from small neighbourhoods upwards.

In this report, following this general introduction, we outline the main decisions taken at each step in the processing of the quantitative data in order to reconcile the data between and within countries, and to link the data across the three main types of administrative data sources.

## Forms of Intervention

The forms of intervention we focused on, children who were the subject of a substantiated child protection concern (on a child protection plan or register) or who were looked after, ‘in care’, are framed by the four countries’ legislation on child protection and children’s services more broadly. Both of these intervention categories illustrate some of the difficulties that are faced in ensuring that inter-country analyses are comparing like with like. In Northern Ireland, Scotland and Wales child protection registers (CPR) record all children who have been identified as being at risk of significant harm. All such children will be the subject of a child protection plan (CPP). In England, however, child protection registers are not kept; but the list of all children subject to a CPP effectively constitutes the register. Subsequently, for simplicity, we refer only to CPPs whichever country we are discussing. For further discussion see Bunting et al., 2017.

There are significant differences as well as substantial similarities in the legal frameworks operating in the four countries with respect to looked after children (LAC) (see McGhee et al., 2017). One concrete example is that in Scotland the label of LAC includes a group of children who are on a Supervision Order to the Local Authority and who are placed at home, while none of the other countries now has an equivalent court order. In England and Wales, children on Special Guardianship Orders – usually placed with former foster carers or relatives – are not included in the LAC numbers and this Order does not apply in Scotland and Northern Ireland. These differences, while important to clarify for quantitative comparison, should not obscure more profound differences in the underlying cultures,

structures and legal systems. These differences are likely to increase rather than decrease as devolution develops. For example, in April 2016 the Social Services and Well-being (Wales) Act came into force and, while still having much in common with English law on children's services and child protection, signalled a step change in the ownership of policy in Wales and at least the beginnings of more divergent policy and practice from England.

## Data Collection

### National Children's Services Data Collections

The national children's services data sets in each country provided the framework and definitions for the data collected in this study. The main focus of current administrative data collection on CPPs or LAC in all four countries is on the children and young people concerned and how cases are processed by local authorities. No data are reported about, for example, the age, educational background, health, income, housing or employment of the children's parents, about family structures or about the quality of the neighbourhoods in which the children live. Rather, the child-level data include information about children's identities, about the various stages through which concerns about children are processed (referrals, assessments, decisions to place children on plans or registers) and the time taken over such stages, about their legal status and the type of placement they are in if they are LAC, the destinations to which children leave LAC status, and some measures of childhood LAC outcomes. Annual reporting of these data at a national and sometimes a local authority level allow for the construction of trends over time, but some changes in the content and analysis of data have taken place in recent years making trends more difficult to ascertain in some cases.

In the absence of socio-economic data about parents, we employed as a proxy the deprivation score of the home neighbourhood of the child. If they were LAC, this was based on their original home address not their current placement address. However, in England and Wales data are only reported by local authorities (LAs) to the DfE on the home addresses of the children involved for looked after children in order to facilitate the calculation of the distance from home that the child is placed. No address data are returned to central government for CPP in England, Scotland and Wales. Identifying the home neighbourhoods of children was crucial for linking child welfare data with deprivation indices. As a result, in England and Scotland individual LAs were approached for all the data required rather than going through the national government departments. In Wales the national government provided geographical data for LAC but local authorities had to be approached for CPP data on home addresses. In Northern Ireland the central repository was able to provide all data, with both positive and negative consequences, as outlined below.

Local authorities collect data on whether children are disabled or not, with detailed sub-categories for the type of impairment(s) involved. However, previous attempts to collect and analyse disability data (Bywaters et al., 2016) showed completely inconsistent results were produced between LAs as to the proportions of children who were disabled. Indeed, LAs with lower overall levels of deprivation were systematically more likely to report higher levels of childhood disability, a finding which contradicts other sources of evidence (Bywaters et al., 2014). This suggests that the data are more a product of the LA's capacity to focus on the issue than anything meaningful in terms of the prevalence of disability. So the disability data – although recorded – were of insufficient quality to be of value and were excluded from our analysis. This – like the absence of data on parental circumstances is a major weakness in the administrative data sets in the UK.

### Census date

Our focus was only on elements of the entire data set, in order to answer our central concern which was to compare and examine CPP and LAC intervention rates at a point in time. For this purpose we selected 31<sup>st</sup> of March 2015, the census date for annual returns in England, Northern Ireland and

Wales. In Scotland, July 31<sup>st</sup> is used for annual reporting so that was the chosen date for the Scottish data. The data we requested concerned all children who were on a CPP or who were LAC on the given dates. In England and Northern Ireland, data were also returned on all Children In Need at March 31<sup>st</sup>. This was beyond the scope of our enquiry in Wales, and the category does not exist in Scotland.

### Content

We requested information about all the children's identities (age, gender, and ethnicity). For LAC, we sought data on their legal status and placement type, in order to support the reconciliation of known differences in approach between the four countries. For children who were on CPPs, we requested data about the type of abuse they were recorded as having experienced.

### Neighbourhoods

In addition to the personal characteristics and case features listed above, we requested an indicator of the neighbourhood in which the child lived, or from which they had entered care, if they were LAC living away from home. The small geographies used for statistical purposes in the four countries varies, as will be described below. We adopted the following areas as providing the best available equivalence:

- In England and Wales: Lower Super Output Areas, average population around 1600
- In Northern Ireland: Super Output Areas, average population around 2000
- In Scotland: Data Zones, average population around 800.

### Deprivation

Each of these small neighbourhoods could be linked to an Index of Multiple Deprivation (IMD) score and therefore to a rank, within each country. Indices are based on multiple domains such as employment, housing, health, income and the environment and often more than one measure is utilised within each domain. The various domains are weighted so that some aspects contribute a larger element of the overall IMD score than others.

The indices provide *relative* measures of deprivation. An IMD score is not an absolute measure of deprivation, but identifies whether any given area is more or less deprived than another. A number of alternative lists of IMD scores and ranks are published, including measures of the concentration of deprivation and indices focused on children. The Income Deprivation Affecting Children Index (IDACI) in England measures the proportion of all children aged 0 to 15 living in income deprived families. It is a subset of the Income Deprivation Domain which measures the proportion of households in an area experiencing low income. We chose not to use IDACI as our preferred Index because it only relates to children up to 15 rather than 17 and because it only uses income as an indicator of deprivation rather than a wider range of factors. We conducted some pilot tests using alternative deprivation measures and the overall results were not significantly different. However, it is undoubtedly the case that LAs' relative ranking is affected to some extent by the deprivation measure adopted, for example, some areas are particularly income poor while in others housing may be a critical issue.

The national IMDs across the four countries use different measures and allocate different weights to the domains used in producing an overall score. For that reason, we used two different indices for the analysis. For the analysis of data *within* each single country, we used the relevant overall national IMD scores at neighbourhood and LA levels (calculating population weighted LA scores where necessary). For the purposes of the *cross-country* analysis we constructed a UK-wide index at the level of each small neighbourhood and LA, following a methodology established by Payne and Abel (2012). After our analysis was under way, Abel et al. (2016) revised their analysis to include new IMD data sets that had been produced in England and Wales. This updating was published too late to be adopted for our

purposes. However, when we checked, the amended methodology produced negligible differences in IMD rankings.

The four national Deprivation Indices are updated at different times. We used the latest available data in each case:

- England: 2015
- Northern Ireland: 2010
- Scotland: 2012 (the 2016 update was published too late for our analysis)
- Wales: 2014

The indicators on which the IMDs are based employ data which are usually recorded two or three years earlier than the published date of the Index.

In Scotland, at the time of data collection, the most recent small areas boundaries were the 2011 Data Zones which are based on the 2011 Scotland Census and released in November 2014. However, LAs who did not send us postcodes (8 of 10), sent data using the 2001 Data Zones (based on the 2001 Scotland Census) rather than the more recent 2011 Data Zones. As LAs knew we were interested in applying Scottish Index of Multiple Deprivation (SIMD) scores, they may have provided us with 2001 Data Zones rather than the more recent 2014 Data Zones as this was the geography used in the most recent SIMD at the time of data collection - the SIMD12. In addition, support for conversion of post codes to Data Zones at this time was limited for the 2011 areas, making post code conversions to 2011 Data Zones much more challenging. There was a high degree of boundary and population change from 2001 to 2011 Data Zones - only 399 of the 6976 new Data Zones remain the same as the earlier 6505 Data Zones, whilst only 50.2% of new Data Zones have one to one population matches with the older zones (Scottish Government, 2015). Support for converting postcode to 2011 Data Zones was increased upon the release of the SIMD16 in late August 2016 (after our data collection was completed), due to the new indices using 2011 Data Zones instead of the older 2001 areas. In future studies it should be possible to use the more recent SIMD scores.

### Population

Population data are available at the level of small neighbourhoods, categorised by age, gender and ethnicity. Although, as we outline below, some further difficulties exist in respect to detailed information about ethnic sub-categories, information was available about the population of children (aged 0-17) in each neighbourhood, which enabled us to calculate intervention rates: the number of children subject to the intervention per 10,000 children in the population. Mid-year population estimates for 2014 have been used throughout except where detailed ethnic categories are being analysed, when data from the 2011 Census is the most up to date available at the neighbourhood level.

### Local Authorities

In England, Scotland and Wales, local authorities are responsible for managing children's services. In England it is upper tier or unitary authorities who hold this responsibility, rather than district councils. In Northern Ireland, health and social care services are integrated and children's services are the responsibility of 5 Health and Social Care Trusts (HSCTs) rather than LAs.

### Sampling

The sampling strategy was designed to meet a number of overlapping aims. It was necessary to secure representative samples of children subject to children's services interventions in each of the four countries, while also ensuring that we had appropriately sized and constructed samples to support

comparison between the countries. In addition we needed sufficient numbers of LAs (or HSCTs) to investigate patterns of intervention.

Taking into account the different population sizes, separate but linked sampling approaches were adopted in each country.

Table 1: The Four Country Samples: LAs, Neighbourhoods and Child Populations

	England	Northern Ireland	Scotland	Wales	All
Total child population	11,591,701	433,161	1,033,183	629,609	13,685,256
Child population in sample	1,432,180	433,161	548,020	629,609	3,042,673
Number of LAs in sample	18	5	10	22	55
Total number of small neighbourhoods	32,844	990	6,505	1,909	42,248
Neighbourhoods in the sample	4,115	990	3433	1909	10,447
Sample child population as percentage of total	12.4%	100.0%	54.1%	100.0%	

In England, the objective was to include a minimum of 10% of all children and 10% of all local authorities. A regional structure divides the country into 10 groups and securing data from 2 LAs in each region was the objective set. This would have provided data from 20 LAs, 13% of all LAs. Additional criteria were that the sample included a spread of LAs from amongst those with low, medium and high average deprivation, that a minimum of 10% of children on CPPs or who were LAC in England would be included, and that a minimum of 10% of children from each of the five broad ethnic categories would be included. All these criteria were met in the final sample except that data were secured from only 18 LAs. Consequently, two regions (North West and South East) had only one LA included.

In Wales and Northern Ireland a strategy based on a similar sample of LAs would have produced very low numbers of children in the data set, which would have made comparisons between countries almost meaningless. In each case the total child population was lower than that for a 10% sample in England. Therefore in Wales and Northern Ireland all LAs (HSCTs) were included covering 100% of children.

In Scotland's 32 LAs, population sizes are very variable. 5 LAs contain over 38% of the population, with more than 7% in each LA, while the remaining 27 LAs contain only 62% of the population with an average size of under 2.5%. The sample was designed to include a minimum of 50% of all children in at least 10 of 32 LAs, including all of the 5 largest LAs. This meant that the largest urban areas were all included, in addition to some of the predominantly rural LAs.

The overall sample sizes resulting from this strategy are shown in Table 2.

Table 2: The Four Country Samples: Children on Child Protection Plans/Registers and Looked After Children as a proportion of the national published data.

At 31.3.15	Population 0-17 (Mid-year estimate, 2014)	Children on Child Protection Plans/Registers	Looked after Children
<b>England – published data</b>	11,591,701	49,700	69,540
Sample - published data	1,432,180	6,716	8,865
Sample as % of England	12	14	13
<b>Northern Ireland - published data</b>	433,161	1969	2875
SOSCARE cleaned data <sup>3</sup>	433,161	1845	2882
Sample as % of NI	100	94	100
<b>Scotland – published data</b>	1032698	2799 <sup>2</sup>	15,404
Sample - published data	548020	1593 <sup>2</sup>	9285
Sample as % of Scotland	53	57	60
<b>Wales – published data</b>	629609	2936	5615
Sample – published data	629609 (CP) 590036 (LAC) <sup>1</sup>	2936	5350
Sample as a % of Wales	100 (CP) 94 (LAC)	100	95

1. Over 45% of postcodes in two LAs were missing from centrally held LAC data and these LAs were excluded from the analysis.

2. One LA gave us details for children on the child protection register as of 31.07.14 rather than 31.07.15. The published figures for children on child protection plans in this table are adjusted accordingly so will differ from those published on the 31.07.15.

3. In Northern Ireland some discrepancies emerged between the centrally held SOSCARE records we accessed and the data published by the Department of Health (DoH). For more details see below.

Table 3: The Four Country Samples: Ethnicity: England

	White	Mixed	Asian	Black	Other	All
England Population 0-17	8,903,313	590,017	1,136,293	562,333	145,004	11,336,960
Sample Population 0-17	1,100,936	70,426	133,404	68,112	26,268	1,399,146
Population 0-17 by ethnic category, %	78.5	5.2	10.0	5.0	1.3	100
Population 0-17 by ethnic category, Sample, %	78.7	5.0	9.5	4.9	1.9	100
Sample as % of England	12.4	11.9	11.7	12.1	18.1	12.3

Table 4: The Four Country Samples: Ethnicity, Scotland

	White	Mixed	Asian	Black	Other	All
--	-------	-------	-------	-------	-------	-----

Scotland Population 0-17	920674	7735	28054	8156	3054	967673
Sample Population 0-17	476424	4584	19927	5753	2214	508902
Population 0-17 by ethnic category, %	95.1	0.8	2.9	0.8	0.3	100
Sample population 0-17 by ethnic category, %	93.6	0.9	3.9	1.1	0.4	100
Sample as % of Scotland	51.7	59.3	71.0	70.5	72.5	52.6

Table 5: The Four Country Samples: Ethnicity, Wales

	White	Mixed	Asian	Black	Other	All
Wales Population 0-17	591862	13379	18990	4750	4426	629609
Sample Population 0-17 for CP	591862	13379	18990	4750	4426	629609
Sample Population 0-17 for LAC	553807	12381	18201	4613	4255	590036
Population 0-17 by ethnic category, %	93.9	2.1	3.1	0.8	0.7	100
Population 0-17 by ethnic category, CP sample, %	85.6	3.0	1.9	0.7	8.9	100
Population 0-17 by ethnic category, LAC sample, %	92.0	3.0	1.2	0.9	2.8	100
Sample as % of Wales for CP	97.9	95.5	98.2	95.2	99.2	98.0*
Sample as % of Wales for LAC	93.4	81.1	90.0	80.0	82.9	92.6**

\*There is some inconsistency in published ethnic data which make comparisons somewhat unreliable.

\*\* This percentage differs from Table 7 because of rounding in the published ethnic group statistics

Ethnic background was available in the Northern Ireland data set, however it could not be reported because of risk of disclosure due to the very small numbers of children in minority ethnic groups.

Table 6: The Four Country Samples: Ethnicity, NI

2011 Census	White	Mixed	Asian	Black	Other	All
Northern Ireland Population 0-17 years	420,436	3,531	5,259	986	551	430,763
Population 0-17 by ethnic category, Northern Ireland, %	97.6	0.8	1.2	0.2	0.1	100

## Ethics and Access

### Ethical approval

The project received ethical approval from Coventry University but required separate approval in each of the four UK countries.

In England the study was approved by the ADCS Research Group, negating the need for further research governance approvals from individual LAs. The ADCS were assured that no child would be identifiable by the researchers or from any published data produced. Discussion took place as to whether individual local authorities were in support of being identified in reports of the study. Had we attempted to anonymise Local Authorities by using alternative names or numbers it is likely that this would not have been effective as it would be possible to use, say, the published LAC rates to identify which LAs had participated. Complete anonymity would have been difficult to assure and so published data do not include details of individual LAs. Following ADCS approval, Directors of Children's Services were approached to participate in the research and, if agreeable to link us to the relevant local data manager. LA's were reassured that only three named members of the research team had access to the data. Consent to take part was therefore sought at three levels.

In Northern Ireland data were accessed through the Honest Broker Service (HBS). The HBS provides access to anonymised ethically approved health and social care data routinely collected by the Department of Health and associated Health and Social Care organisations. A mandatory application to the Office of Research Ethics Northern Ireland (ORECNI) for proportionate review was part of the application process. From application to data access took around six months, with ethical approval gained in six weeks.

In Scotland ethical approval was granted by the University of Stirling REC. All, but one LA accepted the University REC approval as sufficient in line with the existing national agreement with the Association of Directors of Social Work to accept University REC decisions. We therefore sought and gained separate approval from one LA ethics committee. All consent and information sheets were adapted for the Scottish context.

### Access

Methods used for accessing data were different in each nation.

In *England*, LAs were selected according to the sampling frame outlined above and approached on an individual basis through Directors. The majority of LAs approached agreed to take part. Refusals were mainly because of pressures resulting from impending Ofsted inspections, the consequences of recent inspections or staffing pressures in the data management area. Where alternative LAs were sought, replacements were identified that met the criteria for maintaining the balance between regions, levels of deprivation and population size. All participating LAs were provided with a template in Excel

indicating the data sought (based on mandatory LA returns to the DfE). Most LAs used this format but in some cases we had to extract the data from a wider data set..

LAs who agreed to take part were sent written details about the research as a whole and detailed written instructions about the data sought. Often participating LA's raised both ethical issues and questions for clarification regarding the purpose of collecting the data, the anonymity of individual children's data and the security of transferring and storing data. Provision of LSOAs was problematic for some who used postcodes for internal purposes, in which case we accepted postcodes and converted them to LSOAs ourselves.

Data were transferred by LAs in a variety of ways, using secure sites, or encrypted files and, on occasion, by email. Data or Information Sharing Agreements were required by a small number of LA's. Finalising such agreements was sometimes a lengthy process, involving Coventry University's Research Support Partner and legal team and the legal team of the LA. Negotiations between CU and one LA were particularly protracted. It was important to LA's to be kept informed of the progress of the research and the outcomes, all were invited to the end of project launch of findings in London.

*Northern Ireland:* The HBS linked data from the Social Services Client Administration Retrieval Environment (SOSCARE) database to the family of origin address at the time of referral to children's social services using the neighbourhood Super Output Area (SOA) identifier. The HBS charged a standard daily rate for preparing the data (3 days @ £475 per day). All members of the research team also had to undergo safe researcher training once data access was approved. All data analysis was conducted within the HBS 'safe haven' with supervised access on PCs with both internet and USB connections disabled. Outputs had to be requested and cleared for approval to avoid disclosure, and we were not permitted to analyse data offsite. The HBS operate two workstations to facilitate this which could be booked in advance.

In *Scotland* selected local authorities were approached individually and all bar one were happy to share data. Protocols for data access varied between LAs with some requesting additional information and permissions including a copy of the research proposal, ethical approval, privacy impact assessments and data sharing agreements. Senior managers approved access and delegated to information managers responsible for the Children Looked After Survey (CLAS) and Child Protection (CP) returns. An information data collection information sheet containing details of the study aims and information required was then shared with the appropriate personnel. LAs were asked to provide CLAS and CP data as well as the Data Zone of the child's address of origin. One LA raised concerns about potential data disclosure. With agreement they provided age data in age bands and excluded unique identifiers such as the Scottish Candidate Number, using randomly assigned pseudo identifiers instead. This LA also carried out some minor recoding of demographic characteristics with occurrences of 5 or less. Data were requested in Excel spreadsheet format, a sample template was provided but most LAs provided the data in their own format.

Most LAs were happy to use the University of Stirling's secure drop facility to submit their data online, however one LA requested a copy of the IT security policy to send to their IT department. It agreed to release the data using this method as long as the data were downloaded and then deleted at a designated time. All files were removed from the secure drop immediately after collection. Other LAs used alternative methods of secure data transfer included remote access to an encrypted server, emailing encrypted files, or password encrypted file transfer. All data received was then encrypted. Data collection took approximately seven months, and was completed in June 2016.

*Wales:* The data used for the study were drawn from two sources, the Children in Need census data on children on the child protection register and the SSDA903 dataset, which is the annual return to

Government in relation to children and young people looked after by local authorities. The SSDA903 data were provided by the Data Unit within Welsh Government and the child protection data were supplied by each of the 22 Welsh local authorities individually. As part of the agreement Welsh Government converted post code data to LSOA code; child ID numbers were replaced with auto generated anonymised ID numbers; and DOB was changed to age at 31st March 2015. The child protection data were supplied by each of the 22 Welsh local authorities individually, following an approach made with a supporting letter from the Welsh Government encouraging LAs to take part in the study. A template was provided for provision of data but local authorities in fact provided the data in a range of different formats. Data were transferred using secure file transfer systems – AFON for the Welsh Government and FastFile (via Cardiff University) for local authorities.

### Data Cleaning

Although cleaning guidelines were established for project as a whole, each nation encountered individual difficulties preparing the data for analysis.

*England:* As outlined above, data were provided in different formats across the participating LAs. As data was provided, back-up copies were made where data sharing agreements permitted and each child was established as a single case rather than separated number of episodes. Checks were also made for cases that had ceased before the 31 March 2015 census date and were removed. Age and date of birth data had to be standardised and codes for LSOA, MSOA, District/LA, ethnicity, initial abuse category and legal status had to be identified and standardised. Further checks were undertaken to remove unborn children on CPPs, those aged 18 or over, cases missing age details or LSOAs and cases where the LSOA was outside the LA in question. For those cases missing an LSOA but with a postcode, GeoConvert software was used to convert this to LSOA 2011 (Google, [statistics.data.gov.uk](https://statistics.data.gov.uk) and [opendatacommunities.org](https://opendatacommunities.org) were checked where GeoConvert had not been able to source an LSOA). Some authorities provided data based on LSOA 2001 boundaries which had to be converted to 2011 codes. Using the 'LSOA 2001 to 2011 lookup table' provided us with the new LSOA 2011 codes and names in most cases. If there was only a change in name or a merger of LSOAs, the case was changed to the current LSOA. Where the LSOA had split into two or more new LSOAs and there were multiple cases affected, these were evenly distributed across the new LSOAs. A 'random allocation generator' (<https://www.randomizer.org>) or rolling a dice was used when there was a single child in the 2001 LSOA but cut across several different boundaries in the updated 2011 LSOAs.

Where there were queries, a list was compiled for the LA and submitted as a single request for clarification so as not to burden the LA with multiple requests.

*Northern Ireland:* All data were provided in Excel spreadsheets, with multiple rows for each child episode. These files were converted to SPSS Version 22 and separated into three different data sets: CP, LAC and Children in Need. Multiple duplicate cases with different SOAs were recorded. A request for additional data were approved which identified the SOA at the date of referral; for children in need or children on the Child Protection Register. The SOA on or nearest to the census date was then used to select the primary case. For looked after children, the earliest date of referral was considered to be indicative of the birth family's proxy measure of deprivation, and this date was used. Age was provided in 'years of age' format, however this had been calculated on the date the data were extracted and not the census date (13 March 2015), this was re-requested and a number of cases aged 18+ were excluded at this stage. A small number of cases were removed because an age discrepancy could not be resolved. Unborn children and children without an SOA were also removed. LAC placement type, legal status, type of abuse and ethnicity was recoded to correspond with other nation coding frames to facilitate comparison. NIMDM 2010 at SOA level was used to calculate deprivation ranks, deciles

and quintiles. Using the mid-year population estimates for 2014, rates for interventions were calculated. Ethnicity and religious background population data were based on 2011 census returns.

*Scotland:* Although a sample template was submitted to LAs if requested, most decided to submit data in an alternative format which meant that every data set differed in style and content. Some LAs presented data as one Excel sheet with one child per row, but most submitted data across a number of Excel sheets. These Excel sheets were imported into the software package SPSS as separate data sets which were then merged to create one data set for each LA. Some LAs submitted the data in numeric format using Scottish Governments ScotXed 2015 Data Specification coding. When this was the case a simple pre-made syntax file could be used to add value labels to the variables. However, many LAs submitted data in string format, requiring data to then be converted into numeric format. A template syntax file was created for this task, however, all files contained idiosyncrasies which meant this syntax needed adapting slightly for each data set. A child level data set was then created each for the CP and LAC data by combining the data from the LAs. This data set was then used to aggregate child details to Data Zone level and then merged with Data Zone population and SIMD12 files.

Where address information was missing or multiple addresses were recorded, the most recent parental home address was used. Missing address information was minimal. Two LAs had higher levels of missing data and this exercise raised awareness of data quality issues that the LAs were keen to address.

Most LAs converted postcodes to Data Zones themselves, for those having difficulty locating Data Zones, additional permissions were granted to allow us to convert postcodes to Data Zones. This was done by creating a V-LOOKUP Excel spreadsheet from the Scottish Postcode Directory 2015\_2 file, downloaded from the ISD Scotland website ([www.isdscotland.org/](http://www.isdscotland.org/)). The LAs that had provided Data Zones used 2001 Data Zone boundaries. This may have been for a number of reasons: the Scottish Index of Multiple Deprivation (SIMD12) scores used the 2001 Data Zones; in addition, support for the more recent 2011 Data Zones was limited and more difficult to convert postcodes.

There was difficulty identifying certain categories of children by legal status/reason by placement type in 3-5% of children, in the under 5s. Some ethnicity data were missing, along with non-specified gender in a small proportion of the data.

In Scotland, a child may appear in both the CP and LAC data sets. Individuals were matched in the CP and LAC data sets and a number of inconsistencies in this reporting were identified. Additional data were requested from a number of LAs to resolve errors. Once the final data set was completed, it became apparent that for one LA, the sample size for individuals on the CPR at 31 July 2015 was 50% less than the Scottish government published figure. This LA was approached and it was discovered that the data surrounding the children who were on the CP register was as of the 31<sup>st</sup> July 2014 and not the 31<sup>st</sup> July 2015. The data supplied were details of what had happened to these children only, during the year 1<sup>st</sup> August 2014-31<sup>st</sup> July 2015. As this was the case it was decided that these individuals could still be included in the main analysis of CP data, with the acknowledgment that the data were for all individuals on the register for the 31<sup>st</sup> July 2014 and not 2015. However, as data given were a follow up of these individuals over the year after they were initially registered, abuse concerns identified at the individuals' initial case conferences were missing from the data. This resulted in exclusion of this LA from analysis in the area.

*Wales:* The data on children 'looked-after' were provided by Welsh Government and had undergone a process of cleaning by submitting local authorities and the Welsh Government data unit as part of the SSDA903 return process. As a result there was no missing values or duplication of cases. The data were a 100% sample of all children looked-after in Wales on the census date. However, not all cases could be included as some did not have a usable LSOA code as although they had a code it related to the postcode having been suppressed, in the case of children placed for adoption (Deleted); not provided by the local authority (Unknown); or it related to a child whose postcode at time of becoming 'looked-after' was outside of Wales (Outside W). These cases were removed from the analysis. The cases where the LSOA was not provided were also concentrated in two local authorities and as a result these LAs were removed from the analysis in relation to 'looked-after' children.

An information sheet was produced and sent to each of the 22 local authorities outlining the data being requested for each child on the child protection register on 31st March 2015. Whilst this summarised the information requested there was variation in the format in which these data were supplied by local authorities. Age and/or date of birth data had to be standardised and codes for LSOA, local authority, ethnicity, initial abuse category and legal status had to be identified and standardised as these were provided in a number of formats. Further checks were undertaken to remove unborn babies from the dataset and any young people who were 18 years of age or older on the census date. In the information sheet the initial abuse category was requested and coded as one of the following Neglect, Physical Abuse, Sexual Abuse, Emotional Abuse or Multiple categories. Some LAs coded the last category as 'Multiple', whilst others provided a list of all the relevant categories for each case, for example Neglect and Emotional Abuse. As there was a discrepancy between the published figures on registration category and the data collected, this issue was discussed with Welsh Government. Through these discussions it became apparent that when publicly reporting the figures for child protection, in the case example above they would not record this case as multiple categories, but as Neglect. The assumption being that all abuse and neglect has some element of Emotional Abuse. A case registered as Neglect and Physical Abuse would on this basis be reported as 'Multiple'. Where it was possible, based on this advice, cases where the multiple categories included Emotional Abuse were recoded to the other predominant category in order to bring numbers in line with published figures.

## Final Sample

### England

In England the final sample (Table 7) comprised 18 LAs, two from each of 8 regions and 1 each from two regions (North West and South East), treating Inner- and Outer-London as separate regions. The cleaned data were checked against the officially reported data for CIN, CPP and LAC and comprised 94% of CPP and 91% of LAC. In all, clean data were secured on 6310 CPP and 8090 LAC.

Table 7: Cleaned data as a percentage of published data in four country samples.

	Population 0-17	Children on Child Protection Plans/Registers	Looked after Children
England Sample: Published data	1432180	6716	8865

England Sample: Cleaned data		6301	8090
England Sample: Cleaned data as a % of published data		93.8	91.3
NI Sample: Published data	433,161	1969	2875
NI Sample: Cleaned data		1845	2882
NI Sample: Cleaned data as a % of published data		93.7	100.2
Scotland Sample: Published data	548020	1593	9285
Scotland Sample: Cleaned data		1531	8418
Scotland Sample: Cleaned data as a % of published data		96.1	90.7
Wales Sample: Published data	629609	2936	5350
Wales Sample: Cleaned data		2847	4965
Wales Sample: Cleaned data as a % of published data		97.0	92.8

In terms of the child population covered by the data, the sample LAs included 12.4% of all children in England, based on the mid-year estimate for 2014. The proportion of children who were CPP or LAC after cleaning, were 12.7% and 11.6% respectively. In order to allow for the loss of cases in the cleaning process, in the final analysis some results were adjusted for reduced sample size by uprating each LA's totals to create the equivalent of an 100% sample from each LA. This involves the assumption that all missing data are distributed by the same pattern as the cleaned data, i.e. if 5% of data are missing from one LA then 5% are missing in every deprivation decile. The final adjusted figures then give rates which better reflect the published data for each LA and also for the relative deprivation pattern in each LA. For example, the two LAs in which we had the lowest proportion of cleaned LAC data to published data are both within the highest third of English LAs by overall deprivation. Not adjusting for the lower returns would produce underestimates of the impact of high overall deprivation on the total picture.

#### Northern Ireland

In Northern Ireland, a total population sample was used. Population data were based on the Northern Ireland Statistics Research Agency (NISRA) mid-year population estimates for 2014 (available online [www.nisra.gov.uk/publications/2014-mid-year-population-estimates-northern-ireland](http://www.nisra.gov.uk/publications/2014-mid-year-population-estimates-northern-ireland)). Although the total cleaned sample was within a 7% tolerance of the published statistics (see Table 8), when broken down by Health and Social Care Trust, there was greater variation between the data set and published data.

Table 8: Comparison of SOS CARE Data with Published Data by Intervention and HSCT (2014-15)

	Children on CPR	Looked after children
--	-----------------	-----------------------

HSCT	SOSCARE cleaned data	NI official publication	Difference (%)	SOSCARE cleaned data	NI official publication	Difference (%)
Belfast	287	382	+ 25	559	742	+ 25
Northern	483	505	+ 4	662	679	+ 3
South Eastern	385	377	- 2	567	464	- 22
Southern	281	297	+ 5	520	520	0
Western	409	408	0	574	470	- 22
Total	1845	1969	+ 6	2882	2875	0

Reasons to explain the variation were considered by statisticians and senior managers in both the Department of Health and the Health and Social Care Board. Figures for the Belfast HSCT were substantially different for CPR, and the LAC population in Belfast, South Eastern and Western Trusts also varied considerably from the published official statistics. Anomalies in data collection procedures were highlighted with some children's addresses based on foster carers' addresses rather than the birth family addresses, which was the reverse of our approach. In addition, the official statistics are based on aggregate returns which may be more vulnerable to coding errors; the SOSCARE sample is based on individual child data which may contribute to some variation.

#### Scotland

In Scotland, based on the 2014 mid-year estimates the sample LA's contained 53.1% of the 0-17 child population. The study sample once cleaned contained 54.6% of all LAC, and 54.7% of all children on CPPs in Scotland as of the 31st July 2015. For the sample LAs, this represented 90.7% of the published figures for LAC and 96.1% of the published figures for CPPs. As described for England, the final results were adjusted to account for the discrepancy observed between the published figures and those in the cleaned study sample.

#### Wales

The data on child protection registrations in Wales cover all 22 LAs, whilst for LAC the analysis only used data from 20 local authorities as there was too much missing postcode data from two local authorities (Ceredigion - 80% missing; Vale of Glamorgan - 47% missing) for them to be included. As a result, the child protection analyses include local authorities accounting for all 1909 Lower Super Output Areas (LSOAs) in Wales, whilst the 'looked-after' children analyses represent a sample of 1784 LSOAs.

The sample of children and young people included in the Welsh child protection analysis represented 97% of all children on the child protection register in Wales on 31st March 2015. The sample of 'looked-after' children included in the analysis represent 88% of all children in care in Wales on the census day and 93% of all the cases within the 20 local authorities included in the analysis. As there was close to a complete sample at the country level in Wales no adjustment factor was used.

#### Spread of deprivation within each nation

In England, 6 sample LAs were in the least deprived third of all LAs (excluding the City of London and Isles of Scilly on size); 5 were in the middle third and 7 in the least deprived third. However, because the high deprivation LAs were smaller in size on average, the population was not evenly spread between the three groups.

Table 9: Spread of deprivation in England sample LAs (England IMD)

	Number of LAs	Child population: sample LAs	% of all sample population
Low Deprivation	6	528404	37
Mid Deprivation	5	490900	35
High Deprivation	7	398612	28

In Scotland, 4 sample LAs were in the least deprived third of all LAs; 3 were in the middle third and 3 in the least deprived third. The largest proportion of the sample population resided in the high deprivation third. This was in part driven by one particularly large LA with high levels of deprivation and by the extra LA in the high deprivation band.

Table 10: Spread of deprivation in Scotland sample LAs (Scotland IMD)

	Number of LAs	Child population: sample LAs	% of all sample population
Low Deprivation	3	167482	31
Mid Deprivation	3	145368	27
High Deprivation	4	235170	43

In Wales and NI, the 100% samples meant that the whole range of LA/HSCT level deprivation was represented.

#### Spread of Deprivation across the UK

In total, data from 55 local authorities and Health and Social Care Trusts out of a total of 213 were collected. In the comparative elements of the quantitative study this was analysed using the UK Index that we had constructed. As would be expected, deprivation affecting children is not evenly spread across the neighbourhoods, local authorities and countries of the UK.

We found very large differences in deprivation at LA level between the countries. Four out of the five NI Trusts fall within the most deprived third of all authorities in the UK and none in the least deprived third. In Wales, the least deprived LA (Monmouthshire) ranks 45<sup>th</sup> in the UK, outside the top 20%, and only three out of 22 Welsh LAs are in the least deprived third in the UK, when at least seven might be expected. In Scotland, despite some very intense areas of deprivation, notoriously in Glasgow, at a LA level, LAs were over-represented in the least deprived third of UK LAs and under-represented in the most deprived third.

Table 11: Spread of Deprivation in All Sample LAs (UK IMD)

Deprivation	England	NI	Scotland	Wales	All
Low	51	0	16	3	70

Mid	48	1	9	12	70
High	53	4	7	7	71
All	152	5	32	22	211

At the neighbourhood level, the greater relative deprivation of Northern Ireland and Wales is again apparent, but the distribution of Scottish neighbourhoods is more evenly spread than is the case for LAs, because the low deprivation LAs in Scotland tend to have smaller average populations than the high deprivation LAs.

Table 12: Distribution of Neighbourhood Deprivation in Four UK Countries (UK IMD)

		Deprivation quintiles, UK IMD					
		1	2	3	4	5	All
England	Number	6920	6691	6427	6428	6378	32844
(LSOAs)	Percentage	21	20	20	20	19	100
NI	Number	7	130	175	254	327	893
(SOAs)	Percentage	1	15	20	28	37	100
Scotland	Number	1360	1265	1329	1258	1293	6505
(DZs)	Percentage	21	19	20	19	20	100
Wales	Number	143	343	499	488	436	1909
(LSOAs)	Percentage	7	18	26	26	23	100

Similarly, the spread of the child population varies by deprivation and country. As Chart 1 shows, in Northern Ireland and, to a lesser extent Wales, few children were living in the least deprived 20% of neighbourhoods in the UK. And correspondingly, the proportion living in the most deprived 40% of neighbourhoods was very high, especially in Northern Ireland: 68%. In all countries children were more likely to be living in the most deprived 20% of neighbourhoods than any other quintile but, beyond this, in England and Scotland, children's home addresses were fairly evenly spread.

The effect this has on LAs in the thirds of LAs with the lowest, mid or highest deprivation can also be seen in the subsequent chart (data for England only). Fifty three percent of children in the high deprivation LAs lived in the most deprived 20% of neighbourhoods, but less than 5% did so in the least deprived LAs. These differences in the population and in the lives of children and their parents have very substantial implications for the level of demand on children's services.

Chart 1

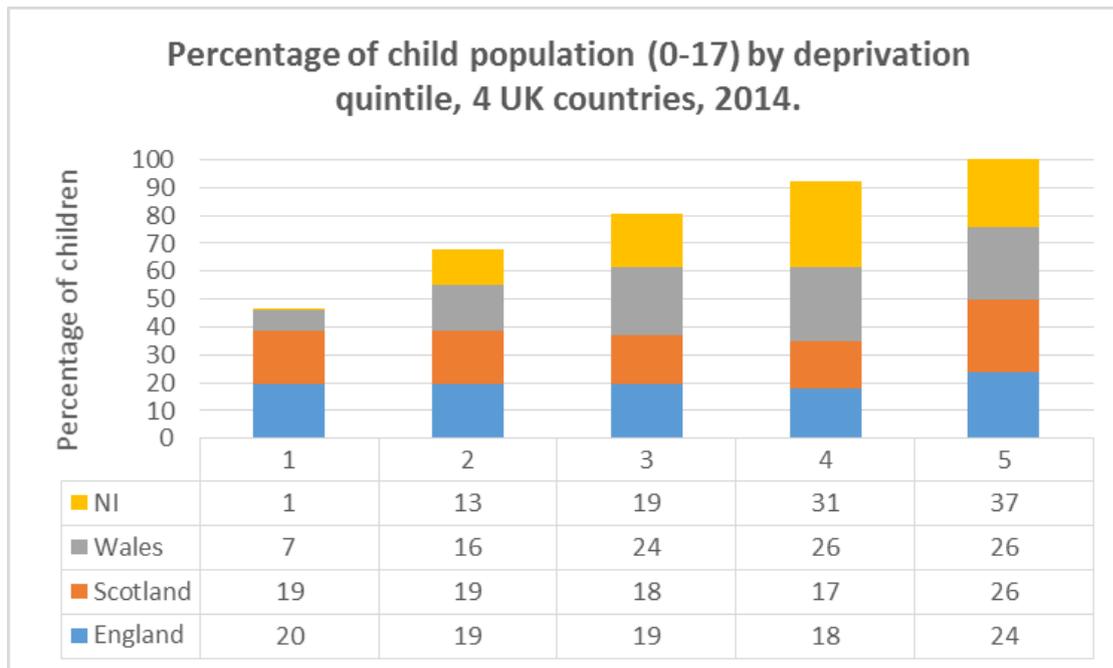
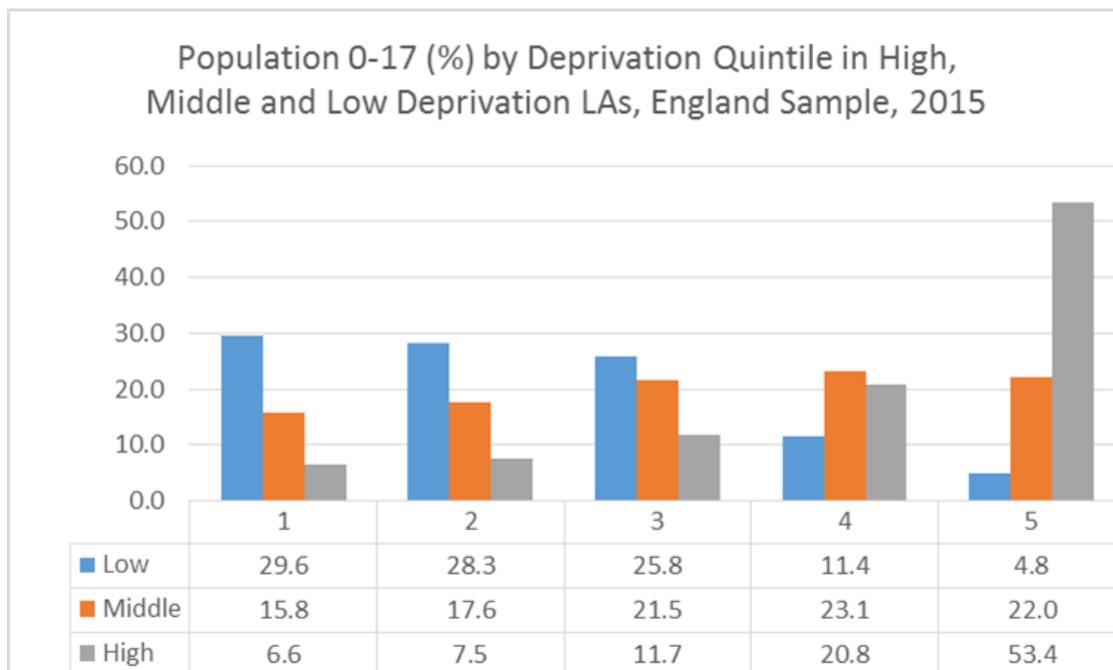


Chart 2



### Ethnicity

A much larger proportion of children in England are identified in population data as being from minority ethnic groups than is the case for the other three countries. From the 2011 census over 20% of English children were identified as non-White compared to less than 10% of children in the rest of the UK. This means that numbers of ethnic minority children in the children's services samples in Scotland and Wales are very small, and not made available for Northern Ireland for this reason. The data for ethnic categories are therefore essentially data for England.

A second problem arose from the categories adopted in nationally reported data on children's services: White, Mixed, Asian, Black, Other. There are a number of problems with these broad categories: for example, how they are ascribed and by whom; whether distinctions between, for

example, Black Mixed heritage children and Black children are – or could be – consistently applied; whether categories than elide children with different national origins, religions, birth places and skin colours can be meaningful.

Thirdly, the mid-year population estimates which we used for the main analysis are not available by ethnic group at the level of small neighbourhoods and so we had to rely on the 2011 census for this part of the work. Clearly, population data affecting young children in minority groups may have changed significantly in the four years between the census and this study. The impact of this problem is reduced by the scale of the study and the amalgamation of neighbourhoods into deciles or quintiles, but it remains the case that rates have to be treated with a degree of caution. It is likely that rates for ethnic minority children are a little overestimated, because the population denominator used was smaller than the reality.

We were given data for 18 ethnic sub-categories in England, and some analysis is carried out at this level, while, as seemed most appropriate to the issue under discussion, the broad categories were also sometimes reported.

Table 13: Sample by ethnic category as a proportion of total population, England.

England	White	Mixed	Asian	Black	Other	All
Overall sample child population, 0-17 (2011 Census)	1,156,682	75,542	143,843	71,466	26,960	1,474,493
Percentage by ethnic category - sample	78	5	10	5	2	100
Percentage by ethnic category - England	78	5	10	5	1	100
Sample population as % of England	13	13	13	13	19	13

### Aligning country data

Separate papers have been written about the issues involved in aligning the data between the four UK countries (Bunting et al., 2017; McGhee et al., 2017) and so we report the main issues here only in outline. There are no perfect solutions to the task of creating like-for-like comparisons of child protection and looked after children data across the four UK countries. For analytical purposes a series of decisions had to be made about how to handle discrepancies in what data were collected and the underlying principles of law and policy in each nation.

### Child abuse and neglect categories

As Appendix 1 sets out, the four UK countries employ different ways of categorising cases where concerns about child abuse or neglect have led to children being placed on child protection plans or child protection registers. All four countries include the four categories of physical abuse, sexual abuse, emotional abuse and neglect but have different strategies to manage the possible occurrence of multiple forms of abuse and Scotland's approach is now radically different to that in the other countries.

In England, only one category is permitted to be recorded for each child but the additional category of 'multiple' is available 'when more than one category of abuse is relevant to the child's current protection plan' (Department for Education (DfE), 2015a, p. 38). However, there is considerable

inconsistency in the use of this category between local authorities with many never using it, while it is used by a few others in more than half of all cases (DfE, 2014b). In our sample LAs, 8 did not use this category at all; 4 used it in fewer than 3% of cases but 2 LAs used it in at least 1 case in 6. The overall national rate of ‘multiple’ abuse was 8% in 2015 but given this variation in practice, this has little meaning as a national proportion. In Wales a similar approach is adopted in the sense that only one category per child is recorded, but some possible combinations of multiple forms of abuse and neglect are named, leading to eight categories overall (see also the discussion of multiple categories above). The same problem in the inconsistent use by LAs of the multiple categories applies as it does in England. This dilemma between recording single or multiple forms of CAN is addressed differently in Northern Ireland where a main and sub-category strategy is employed (Department of Health, Social Services and Public Safety (DHSSPS), 2014, p. 55), generating 17 categories overall. In all these three countries, it is possible to collapse the data into five categories, including a multiple category, which is valuable for analytical purposes. However, the validity and consistency of the data, given the different approaches between the four countries and between local authorities within England and Wales, is a cause for concern.

In Scotland, a further overlapping set of distinctions have been made since 2009–10, with several ‘concerns’ rather than a single reason being recorded for each case, if appropriate, resulting in an average of 2.6 concerns per case conference (Scottish Government, 2015). As can be seen in Appendix 1, this list not only includes the four core abuse categories but what would usually be seen as risk factors such as ‘non engaging family’ or ‘parental substance misuse’ (but no measures of socio-economic circumstances). In other words the categories are signifiers of a diverse set of issues, some describing what has been done to a child, some implying that the child’s behaviour is the problem, some describing parental behavioural issues and some detailing mental but not physical health as an issue. It is not possible to reconcile the Scottish approach with that in the other countries and so no such comparisons are attempted in our analysis.

In addition to these differences in the categories for recording child abuse and neglect, there are also differences in definitions which we outlined in the JRF review (Bywaters et al. 2016). It is unclear whether these and other differences in definition between the four countries result in or reflect differences in identifying or recording instances of abuse in practice. There is little evidence of attempts being made to enable statistics on the incidence of child protection cases to be comparable, despite the government in England commissioning a report on comparability (Munro et al., 2011b). Munro et al.’s call for greater consistency has not been acted on.

### Placement types

Within the LAC data, one of the main previously identified aspects of difference between countries, is the proportion of children classified as LAC who are placed at home with their parents. Looked after children in Scotland commonly live at home on Compulsory Supervision Orders, a legal measure which does not have an equivalent in the other UK countries, where children in similar circumstances would not be considered ‘looked after’. Although some children who are LAC in the other countries are placed with parents, this is a decision made locally that is not reflected in the legal status of the child.

A second known variable, although given less widespread attention in previous inter-country comparisons, is the proportion of children placed by the state with relatives or family friends. These children are looked after – the state has elements of parental responsibility – but are not living either with an approved foster carer previously unknown to the child’s family or in any kind of

residential care. This group does not include children placed with relatives and friends on an 'informal' or private basis who are not legally 'looked after'.

Against the background of different legal systems and placement patterns, in order to compare LAC rates we decided to use two different dimensions to analyse the data: placement types and legal status. We created five groups by placement type which we called LAC 1 to LAC 5.

LAC 1: All children classified as LAC

LAC 2: All children classified as LAC but excluding those placed at home with at least one parent

LAC 3: All children classified as LAC but excluding those placed either at home with at least one parent or placed away from home with relatives and friends

LAC 4: Only those children classified as LAC who were placed at home with at least one parent

LAC 5: Only those children classified as LAC who were placed away from home with relatives and friends.

This provided alternative ways of examining practice in the four countries. For example, it revealed profoundly different proportions of LAC were being placed with parents, relatives or friends and that the proportions so placed in NI were almost as large as those in Scotland, despite the different legal framework.

#### Legal categories

Within each of those 5 classifications of LAC we also analysed the proportions of children who were subject to different kinds of legal intervention. Again (Appendix 2) we grouped the legal categories of the four countries to make comparisons more appropriate. This resulted in four broad groupings (where X could be LAC 1, 2, 3,4 or 5):

LAC X.1: Child Protection Measures

LAC X.2: Adoption Measures

LAC X.3: Voluntary Accommodation

LAC X.4: Youth Justice Measures

The combination of LAC 1 to 5 and LAC X.1 to X.4 allow us to examine, for example, the proportions of children of different kinds in each country who were accommodated on a 'voluntary' basis rather than through a court order, and the proportions of those who were placed at home or with relatives and friends.

In Scotland, two additional legal categories were included to allow for those not comparable to the other countries to be included. These were for children on Compulsory Supervision Orders (LAC X.5) and a category for other legal reasons (LAC X.6), which included Permanence Orders without the authority to adopt.

#### Analysis

##### Data files

The core datasets, i.e. the individual case-by-case files provided by each LA, the population data, and the deprivation data, were each cleaned using Microsoft Excel, before eventually being converted into SPSS files.

Throughout the study, SPSS Version 24 was used. Each separate SPSS file was edited (recoded) for the purposes of the study and then, where applicable, merged through the SPSS merge files function. Population data and IMD data files were merged together to create an overarching population file, which provided each population breakdown and deprivation values for each neighbourhood (LSOA, SOA or DZ), and in turn this was merged with separate files for CPP or LAC (and LAC split by different placement types). Each country would then have six individual main files (one CPP and five LAC files).

The merged data files were based on one-row-per-LSOA (in line with population and deprivation data), whereas previously the master file for child cases had been one-row-per-child. So each LSOA had a total number of children on a CPP, for example, an IMD score (UK and individual country), and several population breakdown values (among other variables). For each of the main files, the purpose was that they would be used to create tables of values for each desired breakdown (e.g. by gender, or age) and population totals (see below).

#### Syntax – cross-country consistency

The majority of the work completed in SPSS was written or pasted into syntax files, specific to each type of file (i.e. dependent on CPP or LAC, UK IMD or individual country IMD etc.). The core syntax files for creating and processing the main data files discussed above were created by the team working on the English data (and later on the cross-country comparisons) at Coventry University and forwarded to the teams in the other countries, so that data formatting and analysis could be consistent between countries, and in preparation for the cross-country comparisons. Some tweaks were made through inter-team discussions in order to respond to data collection differences, such as those around legal codes, placement types and abuse categories used in each country, but essentially the process was repeated step-by-step in each of the four countries.

The syntax files also provided each team with the process for creating custom tables that would therefore resemble each other between countries. So-called custom tables (in SPSS) were produced for CPP cases and (the varying) LAC cases in order to understand the relationships/patterns influenced by the following:

- Age
- Gender
- Ethnicity
- Age and gender (combined)
- Age and ethnicity (combined)
- Age/gender/ethnicity by abuse type (for CPPs)
- Age/gender/ethnicity by placement types (for LAC)
- Age/gender/ethnicity by legal status (for LAC)
- Overall rates.

The above list refers to the input variables at child-level, but all tables were broken down by individual LAs or LAs grouped by average deprivation levels (high, mid or low) and by UK IMD deprivation deciles/quintiles. This process allowed us to consider data in the following ways:

- LA by UK IMD deprivation deciles
- LA by UK IMD deprivation quintiles
- LAs grouped by high, mid, low deprivation by UK IMD deprivation deciles
- LAs grouped by high, mid, low deprivation by UK IMD deprivation quintiles
- LAs as totals
- Whole countries (for the sample obtained) by deciles/quintiles

### Outputs and rate calculations

The purpose of the custom tables explained above was to be able to calculate rates for each area of interest, these rates indicating the proportion of children per 10,000 who were on a CPP or were classified as a LAC.

For example, to calculate the overall CPP numbers and rates in each deprivation decile for the 18 LAs in England we initially produced two custom tables. First, the SPSS output would be a table with the CPP numbers in each LA by each deprivation decile.

	CPP NUMBERS IN UK IMD Deprivation deciles										
LA	1	2	3	4	5	6	7	8	9	10	All children
LA #1	...	...	...	...	...	...	...	...	...	...	All children in the LA
LA #2 (etc.)	...	...	...	...	...	...	...	...	...	...	All children in the LA
All children	...	...	...	...	...	...	...	...	...	...	The total of children in the England sample by each deprivation decile (and total)

Second, we would generate a population table. All children aged 0-17 in each LA by each deprivation split would be covered in the following way.

POPULATION AGED 0-17	POPULATION AGED 0-17 in UK IMD Deprivation deciles										
LA	1	2	3	4	5	6	7	8	9	10	All children
LA #1	...	...	...	...	...	...	...	...	...	...	All children in the LA
LA #2 (etc.)	...	...	...	...	...	...	...	...	...	...	...
All children	...	...	...	...	...	...	...	...	...	...	The total of children in the population in the England sample by each deprivation decile (and total)

These two tables were then be exported to a Microsoft Excel file, where rates were calculated using specific formulae. So each CPP value (i.e. the number of CPP cases in each authority/decile, say) was computed into a rate per 10,000. The calculation would be like so:

$$\left(\frac{a}{b}\right) \times 10,000$$

Where  $a$  is the CPP number in each cell (e.g. LA #2, decile 4, CPP value) and  $b$  is the child population (aged 0-17) in the equivalent population cell. This process gives a table with the same structure as the two above but containing rates per 10,000 children.

Having all of these tables broken down in these different ways provided the team with the capacity to interrogate the data by all of the available breakdown variables (e.g. gender and age). Examples of some possible statements from the tables might be:

- $x$  Asian children per 10,000 in LA 1 were subject to a CPP, compared to  $y$  White children in the same area; or
- CPP rates for children in deprivation decile 10 were  $x$  per 10,000 compared to CPP rates for children in decile 1, which were  $y$  per 10,000.

Any graphs/charts produced from these rates were created in Microsoft Excel rather than SPSS, and Microsoft Word was often incorporated for the presentation of these to share between team members.

### Statistical tests/analysis

The tables, statements and graphs described above present the data in a descriptive manner. However, once the rates had been produced, the team interrogated the data for significance and effect sizes using a variety of statistical techniques appropriate to the variables under interrogation. These techniques included correlation and regression. Decile level summary data on child intervention rates typically showed an exponential type growth against decile number, thus log transformation prior to analysis was typically used. The comparison of these relationships between countries or between groups of local authorities was made using an extension of regression to test for differences in intercepts (to test for consistent differences in country/group intervention rates) and in slopes (to test for differences in country/group responses to increasing deprivation).

### Adjusted rates

As discussed earlier in the paper, for most of the LAs in England, Wales and Scotland we had some missing data which were not evenly distributed between LAs. If a child was not given a corresponding LSOA they were removed from analysis, for example, or indeed if there were other recording problems on their case. For LAC in England, our highest match up for an LA was 100% and our lowest 71% (but 16 of the 18 had at least an 84% match); for CPP (England) our matches ranged from 86% to 98%. Thus, we also produced 'adjusted rates' tables and statements.

In Northern Ireland, although there were discrepancies between the published data and the SOS CARE data, rates were not adjusted as the SOS CARE data was considered to reflect more accurately the actual numbers on CPR and LAC rather than the aggregate figures returned for official publication. In other countries there were similar or, in cases, larger discrepancies also accounted for once the data were sent to the team at Coventry University.

Taking an example local authority, we had an overall LAC rate of 26.9 per 10,000 children in the area, compared to an official (published) rate of 30.0; meaning we had 89.8% of the official rate covered in our data. Thus, the following was calculated across the table (i.e. in all deciles and overall) for the row representing the LA in question:

$$\left(\frac{x}{y}\right) \times 100$$

Where  $x$  is the original rate we had produced and  $y$  is the percentage of the official rate that we had previously had. So this LA's 'all children' rate:

$$\left(\frac{26.9}{89.83}\right) \times 100$$

This looked like:

<b>Previous and adjusted rates</b>	1	2	3	4	5	6	7	8	9	10	Total
<b>LA unadjusted rate</b>	11.5	13.3	40.7	44.7	36.5	70.8	37.6	114.6	n/a	n/a	26.9
<b>LA Adjusted rate</b>	12.8	14.9	45.4	49.8	40.6	78.8	41.9	127.6	n/a	n/a	30.0

As acknowledged earlier, this is a crude way of estimating the adjusted rates, as the missing data might not have been evenly spread across deprivation deciles, males and females, or ethnic groups etc. but if we had not done this high deprivation LAs, which had more missing data, would have been under-represented in comparisons. Thus, this was seen as the most appropriate response. These tables did not replace other outputs, but were rather to go alongside them as appropriate to the issue under discussion.

This process is also relevant to some of the barriers faced in the data process, discussed in the next section.

#### Cross-country comparisons

As discussed, all four of the teams (countries) were working to produce parallel files to each other within their individual countries at the same time, and between July and September 2016 outputs from these four country files were shared between the wider project team so the process of comparing the countries could begin. Each country team produced two parallel sets of outputs, one based on their own country IMD and one based on UK wide IMD scores.

There were some important differences in data sharing agreements and publically available data across the four countries; for example, in Wales once the sub-levels went as far as ethnic groups by LA by deprivation decile, some of the numbers were too small to report (meaning that other teams could only be provided with Welsh rates, not Welsh numbers). In Northern Ireland, some sub-groups were not available by the smaller neighbourhoods for mid-2014 population estimates; this led to problems in splitting rates across age groups and ethnic groups. We followed, as far as possible, government guidance on the reporting of data designed to protect any possibility of individuals being identified and to prevent inappropriate conclusions being drawn based on small numbers.

#### Individual country analysis

In addition to the whole project team carrying out the UK-wide comparisons, each individual country interrogated their own data using their national deprivation measures.

In England, then, the processes discussed above were replicated on the English data but using England IMD scores rather than UK IMD scores. So main files and their subsequent tables were produced in the same way but by England IMD. This was managed separately in each country, as each team had control over how they interrogated their own data, beyond a common framework. Each team would then go on to produce individual country reports as part of the larger project.

In England, we also developed the work on ethnicity. Whereas the tables in the cross-country work had looked at ethnic differences, these were only looked at between five overarching groups (White,

Mixed, Asian, Black, Other) and excluded NI. But the scale of the England data allowed us to look at smaller subgroups of ethnicities (17, in fact). This is discussed in the next section.

#### Detailed ethnicity analysis

In England we decided to run some of the analysis not only by the 5 broad ethnic categories but by 17 groupings in line with official ONS categories. The below table presents the process of matching these categories to the data we had received from the local authorities:

Codes from master file	The names of the ethnic groups in our data	Matching the numbers to the ONS data			The ONS official categories
WBRI	White British	E1	1	E1	White: English/Welsh/Scottish/Northern Irish/British
WIRI	White Irish	E2	2	E2	White: Irish
WROM	White Roman Traveller	E3 (Combined)	3	E3	White: Gypsy or Irish Traveller
WIRT	White Irish Traveller				
WOTH	White Other	E4	4	E4	White: Other White
MWBC	Mixed White and Black Caribbean	E5	5	E5	Mixed/multiple ethnic group: White and Black Caribbean
MWBA	Mixed White and Black African	E6	6	E6	Mixed/multiple ethnic group: White and Black African
MWAS	Mixed White and Asian	E7	7	E7	Mixed/multiple ethnic group: White and Asian
MOTH	Mixed Other	E8	8	E8	Mixed/multiple ethnic group: Other Mixed
AIND	Asian Indian	E9	9	E9	Asian/Asian British: Indian
APKN	Asian Pakistani	E10	10	E10	Asian/Asian British: Pakistani
ABAN	Asian Bangladeshi	E11	11	E11	Asian/Asian British: Bangladeshi
CHNE	Chinese	E12	12	E12	Asian/Asian British: Chinese
AOTH	Asian Other	E13	13	E13	Asian/Asian British: Other Asian
BAFR	Black African	E14	14	E14	Black/African/Caribbean/Black British: African
BCRB	Black Caribbean	E15	15	E15	Black/African/Caribbean/Black British: Caribbean
BOTH	Black Other	E16	16	E16	Black/African/Caribbean/Black British: Other Black
OOTH	Other 'other' Including Other Arab	Now referring to these children all as E17			Other ethnic group: Any other ethnic group

This allowed us to draw different conclusions in some cases. For example, of the five ethnic groups, we know that Asian children had lower intervention rates. But this finding therefore disguised the differences in intervention rates between children of Pakistani origin and Indian origin. Similarly, the differences between Black African and Black Caribbean rates were disguised in the broad ethnic group findings.

An important change occurred in this data, though, as although ethnicity data are provided at LSOA level, the combination of ethnicity and gender is not. Thus, the team made the decision to use the data at MSOA level (a slightly larger population group size). This meant population-weighting and re-calculating deprivation scores for MSOAs (from the LSOA values) and tweaking the syntax files to suit this process.

### Religious Background

In Northern Ireland, we examined intervention rates by religious background. Population statistics were available at SOA level for religion and gender for children aged 0-15 years, so the same age range in the SOSKARE sample was selected to calculate intervention rates. Cell sizes were small for some categories, and these were collapsed to avoid disclosure. Analyses were presented by deprivation quintile to avoid disclosure as cell sizes were very small by deprivation decile for the 'Other Non-Christian' category.

Of the four nations, Northern Ireland was the only one to record and report religion in child welfare statistics, clearly a legacy of the conflict. Official population data were not available by ethnic group to allow for ethnicity and religion comparisons. Although there is little ethnic diversity in Northern Ireland, in the other nations, recording and understanding the interaction between ethnicity, religion and child welfare interventions may provide further valuable insights.

<b>NISRA Census Religion Category</b>	<b>Analysis Category</b>
Roman Catholic	Roman Catholic
Church of Ireland	Other Christian
Church of England	
Presbyterian	
Methodist	
Other Christian	
Muslim	Other Non-Christian
Other Non-Christian	
Jewish	
None	None
Not Completed	
Not Known	
Refused	

## References

- Abel, G. A., Barclay, M. E., & Payne, R. A. (2016). Adjusted indices of multiple deprivation to enable comparisons within and between constituent countries of the UK including an illustration using mortality rates. *BMJ Open*, 6(11), e012750. <http://doi.org/10.1136/bmjopen-2016-012750>
- Bunting, L., McCartan, C.J., McGhee, J., Bywaters, P., Daniel, B., Featherstone, B. and Slater, T. (forthcoming) Trends in Child Protection across the UK – A Comparative Analysis', *British Journal of Social Work*, forthcoming.
- Bywaters, P., Brady, G., Sparks, T., and Bos, E. (2014) Inequalities in child welfare intervention rates: the intersection of deprivation and identity, *Child and Family Social Work*, doi:10.1111/cfs.12161
- Bywaters, P., Brady, G., Sparks, T., and Bos, E. (2016a) Child welfare inequalities: new evidence, further questions, *Child and Family Social Work*, vol. 21, pp. 369-80, 2016; doi:10.1111/ cfs.12154
- McGhee, J., Bunting, L., McCartan, C., Elliot, M., Bywaters, P. and Featherstone, B. (Forthcoming) Looking after children in the UK – convergence or divergence? *British Journal of Social Work*.
- Payne, R. & Abel, G. (2012). UK indices of multiple deprivation - a way to make comparisons across constituent countries easier. *Health Statistics Quarterly / Office for National Statistics*, (53), 22–37. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/22410690>
- Scottish Government (2015) Data Zone 2001 and 2011 Matching – Quick Guide <http://www.gov.scot/Topics/Statistics/sns/SNSRef/DZMatchingQGuide>

## Appendix 1 Categories of maltreatment in official data in the UK

### England

Physical abuse

Sexual abuse

Emotional abuse

Neglect

Multiple

### Northern Ireland

Neglect, physical abuse and sexual abuse –

- Main category – neglect
- Main category – physical abuse
- Main category – sexual abuse

Neglect and physical abuse

- Main category – neglect
- Main category – physical abuse

Neglect and sexual abuse

- Main category – neglect
- Main category – sexual abuse

Physical and sexual abuse

- Main category – physical abuse
- Main category – sexual abuse

Neglect (only)

Physical abuse (only)

Sexual abuse (only)

Emotional abuse (only)

### Scotland

Neglect

Parental substance misuse

- drug misuse

- alcohol misuse

Parental mental health problems

Non-engaging family

Domestic abuse

Sexual abuse

Physical abuse

Emotional abuse

Child placing themselves at risk

Child exploitation

Other concerns

## Wales

Neglect and physical abuse

Neglect and sexual abuse

Neglect with physical and sexual abuse

Physical and sexual abuse without neglect

Physical abuse

Sexual abuse

Emotional abuse

Neglect

## Appendix 2: LAC Legal Status Coding

LAC legal status across the four countries is grouped into four main categories, with Supervision Orders in Scotland being dealt with separately. They cover respectively children on orders relating to child protection, on orders relating to adoption, who are in voluntary accommodation, or who are on remand, detained or sentenced under youth justice provisions.

		England	Wales	Scotland	Northern Ireland
LAC X.1	Child Protection Measures	C1=Interim care order C2=Full care order L1=Under police protection and in local authority accommodation L2=Emergency protection order L3=Under child assessment order and in local authority accommodation	C1=Interim care order C2=Full care order L1=Under police protection and in local authority accommodation L2=Emergency protection order L3=Under child assessment order and in local authority accommodation	Child Protection Measures	Emergency Protection Order (Article 63 & 64) Care Order (Article 50 or 59) Interim Care Orders (Article 57) Supervision Order Interim Supervision Orders (Article 57)
LAC X.2	Adoption Measures	D1=Freeing order granted E1=Placement order granted	D1=Freeing order granted E1=Placement order granted	Freed for Adoption Permanence order with authority to place for adoption	Freeing order Placement order
LAC X.3	Voluntary Accommodation	V2=Single period of accommodation under Section 20	V2=Single period of accommodation under Section 20	Accommodated Under Section 25	Accommodated (Article 21)
LAC X.4	Youth Justice measures	J1=On remand, or committed for trial or sentence, and accommodated by LA J2=Detained in LA accommodation under PACE J3=Sentenced to CYPA 1969 supervision order with residence requirement	J1=On remand, or committed for trial or sentence, and accommodated by LA J2=Detained in LA accommodation under PACE J3=Sentenced to CYPA 1969 supervision order with residence requirement	Criminal Court Provision	Not presented in national statistics
	Compulsory Supervision measures	NA	NA	Compulsory supervision order at home Compulsory supervision away from home (excluding Residential Establishment) Compulsory supervision away from home (in a	NA

				Residential Establishment but excluding Secure) Compulsory supervision away from home with a Secure Condition Interim compulsory supervision order	
	<b>Other</b>			Other Legal Reason Permanence order Parental Responsibilities Order (this is an historic code and was replaced by Permanence Orders but may still appear in the data for those looked after longer term)	Other