

SUPPLEMENTARY FILE 2 – DATA QUALITY ANALYSIS

Note: all numbers are rounded to nearest 10, and all percentages to one decimal place, in compliance with statistical disclosure rules. This may result in rounding error.

The child in need (CiN) dataset, after cleaning, contained 8,066,880 rows after cleaning, with 3,170,040 unique children (based on the combined ID, which uses PMR where available, otherwise the LA child ID). There were 5,861,860 unique referrals.

The purpose of this analysis was to compare numbers of referrals in our extract of the CiN census with the Department for Education’s own statistical releases available at <https://www.gov.uk/government/collections/statistics-children-in-need>, where further methodological details, including data quality indicators, are available.

ID types

Children’s ID used was either the anonymised LA child ID or anonymised PMR. PMR was used where available. Otherwise, LA child ID was used. In terms of number of referrals, Table S2.1 shows the breakdown of ID type by age and calendar year of referral.

Table S2.1. ID type used on referrals (unit = referrals; row percentages)

	Anon LA child ID Only available	Anon LA child ID And Anonymous PMR available	Total Unique referrals
Total	1,744,480 (29.8%)	4,117,380 (70.2%)	5,861,860 (100.0%)
Age			
Pre-birth	200,270 (88.7%)	25,500 (11.3%)	225,760 (100.0%)
0-4 yr	1,036,840 (60.0%)	690,560 (40.0%)	1,727,400 (100.0%)
5-9 yr	214,780 (13.1%)	1,424,140 (86.9%)	1,638,920 (100.0%)
10-15 yr	197,380 (11.0%)	1,599,640 (89.0%)	1,797,020 (100.0%)
16-17 yr	95,220 (20.1%)	377,530 (79.9%)	472,750 (100.0%)
Referral calendar year			
2008	4,300 (11.0%)	34,950 (89.0%)	39,250 (100.0%)
2009	11,880 (3.7%)	308,060 (96.3%)	319,940 (100.0%)
2010	24,740 (6.6%)	351,920 (93.4%)	376,660 (100.0%)
2011	148,710 (27.1%)	400,740 (72.9%)	549,450 (100.0%)
2012	175,090 (29.8%)	412,990 (70.2%)	588,080 (100.0%)
2013	188,650 (29.4%)	452,650 (70.6%)	641,300 (100.0%)
2014	198,760 (31.1%)	440,240 (68.9%)	639,000 (100.0%)
2015	203,790 (33.1%)	412,700 (66.9%)	616,490 (100.0%)
2016	225,250 (35.6%)	408,160 (64.4%)	633,400 (100.0%)
2017	243,290 (37.3%)	408,740 (62.7%)	652,040 (100.0%)

SUPPLEMENTARY FILE 2 – DATA QUALITY ANALYSIS

2018	251,990 (39.1%)	391,930 (60.9%)	643,920 (100.0%)
2019	68,040 (41.9%)	94,290 (58.1%)	162,330 (100.0%)

Table S2.1 shows the LA child ID was most commonly the ID of choice among pre-birth referrals and children aged 0-4 years, as expected. However, not insignificant proportions of children in the older age groups used LA child ID (i.e., did not have PMR). Table S2.1 also shows a change in the distribution of ID types between 2010 and 2011, before which PMR was dominant.

Table S2.1 also reveals a definite undercount of referrals in 2008. This is because the dataset began in 2008 and all referrals before April 2008 were dropped in the cleaning phase. Likewise, 2019 shows an undercount of referrals as the last day of data collection was 31 March 2019. The completeness of the dataset is otherwise considered in the next section.

Dataset completeness

Table S2.2 shows the number of referrals received and the number of need episodes starting each financial year (1 April – 31 March). The data are given alongside figures published by DfE in their statistical first releases (SFRs) for comparison.

Table S2.2. Number of referrals in our dataset and in Department for Education published data

Financial year ending 31 Mar	Number of referrals (our data)	Number of referrals (DfE data)	Difference in referrals (our data – DfE data)
2009	69,760	-	-
2010	391,860	603,700	-211,840
2011	388,580	615,000	-226,420
2012	586,280	605,000	-18,720
2013	590,400	593,500	-3,100
2014	654,060	657,800	-3,740
2015	632,200	635,620	-3,420
2016	617,770	621,470	-3,700
2017	642,420	646,120	-3,700
2018	652,830	655,630	-2,800
2019	635,710	650,930	-15,220

DfE Department for Education, source:

<https://www.gov.uk/government/collections/statistics-children-in-need>

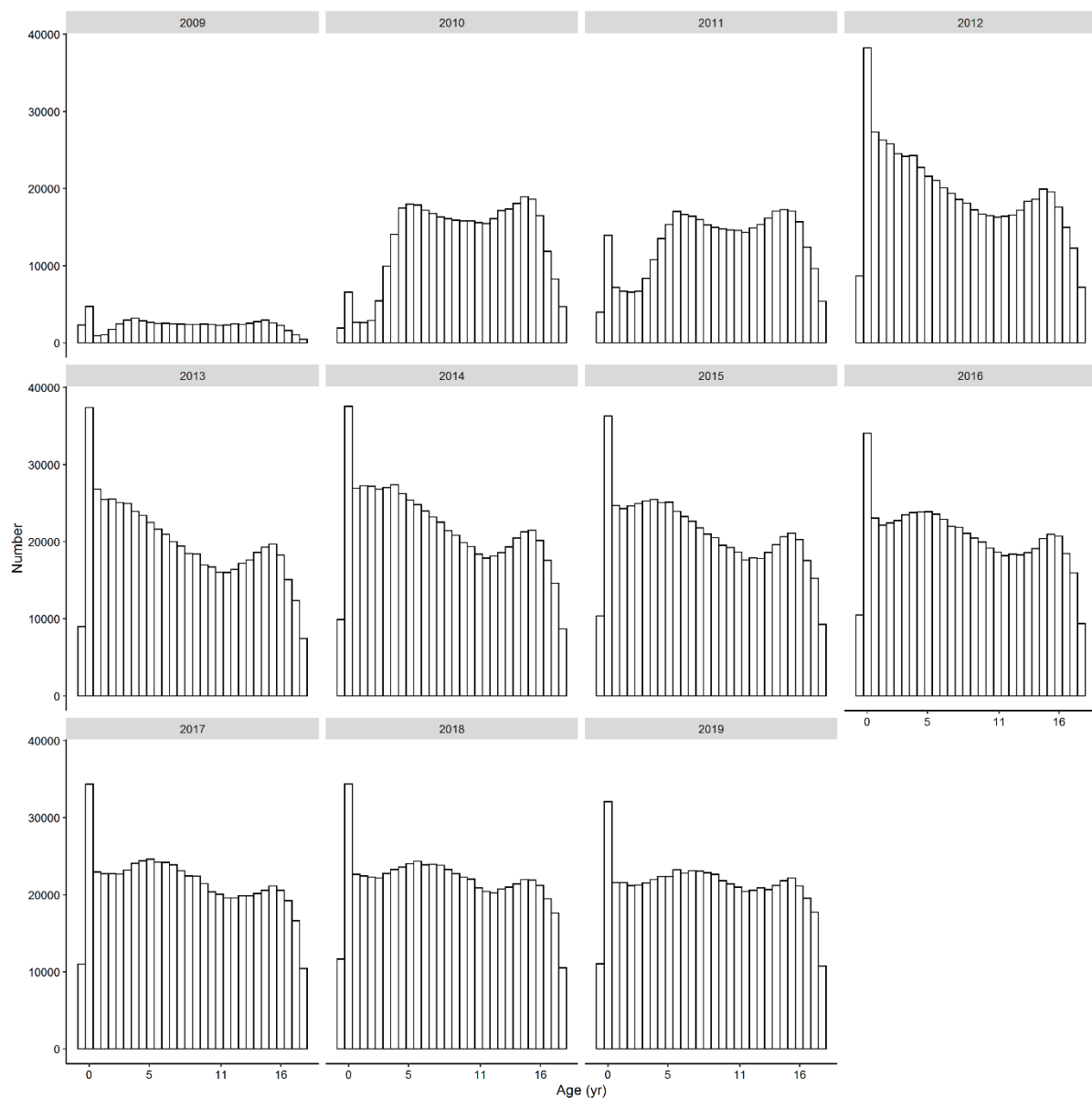
SUPPLEMENTARY FILE 2 – DATA QUALITY ANALYSIS

There was a significant undercount of referrals for the 2009/10 and 2010/11 financial years, and a slight undercount for the 2011/12 and 2018/19 financial years. Especially in the earlier years, given the distribution of IDs (Table S2.1), this is likely an undercount of younger children. The small differences between our counts and Department for Education published counts in other years may be due to different cleaning steps that we have applied.

Ages at referral by financial year

Finally, we examined the age distribution of all referrals, broken down by financial year of referral, as shown in Figure S2.1. Figure S2.1 reveals a likely significant undercount of referrals of younger children (aged <5 years) in years 2008/9 to 2010/11.

Figure 1. Age at referral by financial year (year ending 31 March)



SUPPLEMENTARY FILE 2 – DATA QUALITY ANALYSIS