



Small area level data based on practice postcode

Documentation and Data Dictionary

Version 3.4

Date: 19 April 2022

Documentation Control Sheet

Over time, it may be necessary to issue amendments or clarifications to parts of this document. This form must be updated whenever changes are made.

Version	Affected Areas Summary of Change	Prepared By	Reviewed By
0.9	Initial Draft	Dan Dedman	
0.9.2	Modified	Dan Dedman	
0.9.3	Modified	Dan Dedman	Helen Strongman
1.0.0	Final version for release	Dan Dedman	
1.0.1	Modified	Dan Dedman	
1.0.2	Formatted	Grant Lee	Sophia Amjad
1.0.3	Modified	Shivani Padmanabhan	Daniel Dedman
2.0	Modified	Helen Strongman	Arlene Gallagher
2.1	Modified	Helen Strongman	Shivani Padmanabhan
2.2	Modified	Helen Strongman	
2.3	Modified	Rebecca Ghosh	Helen Strongman
2.4	Modified and formatted	Rebecca Ghosh	Arlene Gallagher
2.5	Reviewed	Arlene Gallagher	
2.6	Modified	Melissa Cabecinha	Rebecca Ghosh
2.7	Reviewed	Susan Hodgson	Elizabeth Crellin
2.8	Reviewed	Susan Hodgson	Rebecca Ghosh
2.8	Reviewed	Susan Hodgson	Rebecca Ghosh
2.9	Reviewed	Susan Hodgson	Rebecca Ghosh
3.0	Reviewed	Susan Hodgson	Jessie Oyinlola
3.1	Reviewed	Susan Hodgson	Eleanor Axson
3.2	Modified	Susan Hodgson	
3.3	Updated	Eleanor Axson	Preveina Mahadevan Susan Hodgson
3.4	Modified	Eleanor Axson	Susan Hodgson

Summary of Changes

Version 0.9.2

- Revised draft

Version 0.9.3

- Revisions, January 2013

Version 1.0.0

- Release version

Version 1.0.1

- Added note on Patient Choice of GP Practices

Version 1.0.2

- Formatted header and footer with new agency branding

Version 1.0.3

- Removed reference to set version, and amended related links section and hyperlinks

Version 2.0

- Updated versions of IMD scores

Version 2.1

- Updated postcode lookup file

Version 2.2

- Added builds from which changes to versions of IMD scores and the postcode lookup were implemented
- Changed reference to HSCIC to NHS Digital to reflect rebranding

Version 2.3

- Updated LSOA numbers and clarified some text

Version 2.4

- Updated versions of IMD scores

Version 2.5

- Updated to include individual IMD domains
- Updated to include Carstairs 2011 Index of Deprivation
- Updated to include Urban-Rural Classification
- Reviewed for set 16

Version 2.7

- Reviewed for set 17

Version 2.8

- Reviewed for set 18

Version 2.9

- Reviewed for set 19

Version 3.0

- Reviewed for set 20

Version 3.1

- Updated to include CCG pseudonym and DOIs

Version 3.2

- Updated to revise reference to ISAC

Version 3.3

- Updated IMD and Townsend measures
- Updated file and variables names
- Updated references
- Updated for set 22

Version 3.3

- Updated Townsend Deprivation Score comparability information

Background

Classifications based on the population characteristics of small areas or neighbourhoods (and the individuals who live there) have been in use for several decades. In health research they have many applications, including: as a proxy for individual level measures of socioeconomic status; for planning and targeting of health and social care services; for ecological studies of environmental effects on health; and for individual level studies where characteristics of place of residence are of particular interest. [1] There are a wide range on small area data available from many sources. CPRD has linked the GP practice postcode for both CPRD GOLD and CPRD Aurum to some of the most commonly requested area level data. This includes several measures of area level deprivation and a rural-urban classification, and a pseudonym for Clinical Commissioning Group, for studies which might require this for the purposes of matching and/or cluster level analyses.

Census geography [2]:

The small area data provided is based on census geography which is the main geography directly associated with the UK Census in England, Wales, and Northern Ireland. The base unit of this geography is the Output Areas (OA) which are built from clusters of adjacent postcode units [3]. Output areas usually contain around 110-140 households and are designed to be similar in population size and social characteristics based on tenure of household and dwelling type. Output Areas can be aggregated into Super Output Areas (SOA) which can be sub-divided into lower layer super output areas (LSOA) and middle layer super output areas (MSOA) in England and Wales [2].

The small area data provided by CPRD is at LSOA level which are typically built up from 4-6 OAs and have a notional minimum size of 1,000 residents and 300 households, and an average of 1,600 residents [4]. In Northern Ireland, data Super Output Area (SOA) are used which have a population of 2,100 people [5]. In Scotland, a similar geography known as data zones (DZ) are used which are slightly smaller than LSOAs, with the majority having a population of 500-1,000 residents.

A note on Patient Choice of GP Practice

Since January 5th 2015, GP practices in England are free to register new patients who live outside their practice boundary area [6]. This might mean that some patients will register at a practice near their place of work or children's schools, rather than a one near their home. Researchers should consider the implications of this change, for example if they are considering using the practice level deprivation score as a proxy for patient level deprivation or socioeconomic status. Linkage based on the patient postcode is available separately for patients registered with a subset of general practices in England, as detailed below, and described in the documentation on small area data for patients.

A note on GP practices with multiple locations

GP practices may have multiple locations. In these cases, the GP practice will have only one practice postcode based on the location of the main practice. Researchers should consider how this could affect the deprivation linkage and rural-urban classification, as the main practice might have branch practices in more rural or deprived areas.

Area level measures of deprivation:

There are several well-known area-based measures of deprivation available at the LSOA level for linkage to CPRD primary care data through the practice postcode. These measures are:

The Index of Multiple Deprivation (IMD):

One of the most used measures of deprivation is the Index of Multiple Deprivation (IMD). This is a composite measure derived from a number of indicators covering different aspects ('domains') of material deprivation: income, employment, education and skills, health, housing, crime, access to services, and living environment. Each domain index can itself be a composite score derived from two or more sub-domain indicators. The overall composite index, the Index of Multiple Deprivation (IMD), is calculated as a weighted sum of the domain indices.

The first official 'Indices of Deprivation' for England [7] were produced by the UK Department for Communities and Local Government in 2000, replacing the 1998 Index of Local Deprivation. Updates were produced in 2004, 2007, 2010, 2015, and 2019. Similar indices are also available for Wales [8], Scotland [9] and Northern Ireland [5] [10]. It is important to note that differences in methodology and source data mean that the indices are not directly comparable between different countries. Note also that all the indices measure relative rather than absolute deprivation. As such it is the ranking of areas provided by the IMD score, rather than the actual score itself, which is of primary interest.

Composite IMD ranks are available linked to CPRD practice postcode at the LSOA level in England and Wales for 2019, SOA level in Northern Ireland for 2017, and at the DZ level in Scotland for 2020. CPRD can provide individual domain IMD ranks on request on a study-by-study basis with an approved protocol. Please contact CPRD Enquiries enquiries@cprd.com, before any application, to discuss.

Townsend Score: [11] [12]

This indicator was devised by Townsend *et al* in 1988 as an index of material deprivation and disadvantage. Originally calculated at ward level, it was based on 4 measures from the 1981 census:

- Unemployment: proportion of the economically active population aged 16-59/64 who are unemployed.
- Car ownership: proportion of households with no car.
- Home ownership: proportion of households not owning their own home.
- Household overcrowding: proportion of private households with ≥ 1 resident per room.

The index is created as the sum of scores for each standardised measure. Townsend scores have been recalculated using data from the 1991, 2001, and 2011 censuses, and for different geographies – for example census output areas (OA) and LSOA. As with IMD scores, it is the ranking of areas provided by the Townsend score, rather than the actual score itself, which is of primary interest. The Townsend Score is comparable across UK countries, but not over time (i.e. it is inappropriate to compare an area measure from 2001 to a measure from 2011) [13].

The Townsend Score is available linked to CPRD practice postcodes in England, Wales, Scotland, and Northern Ireland at the LSOA/DZ/SOA level for the 2011 census.

Carstairs Index: [14]

This indicator was devised in the 1980s by Carstairs and Morris as an index of material deprivation at the small area level. Originally developed for Scotland, and now covering England and Wales as well, it is based on four measures from the UK census:

- Male unemployment: proportion of the economically active population of males age 16-74 who are unemployed.
- Car ownership: proportion of households with no car
- Overcrowding: proportion of all persons living in private households with a density of more than one person per room.
- Social Class: proportion of persons in private households with an economically active head of household in a low social class. Social class score was originally created using the Registrar General's Social Class [15], but as of 2001, this measure is approximated from the operational categories of the National Statistics Socio-Economic Classification

The index is created from the sum of scores for each standardised measure. Carstairs scores have been recalculated using data from the 1991, 2001, and 2011 census. In England and Wales, Carstairs scores are calculated at LSOA level, while scores for Scotland are calculated using data zones. As with IMD scores, it is the ranking of areas provided by the Carstairs Index, rather than the actual score itself, that is of primary interest. Carstairs 2011 scores are comparable between the different countries because the scores are calculated based on the total distribution across all LSOA/DZs with data available. The Carstairs index covering Scotland, England, and Wales is available linked to CPRD practice postcode at the LSOA level for the 2011 census [16].

The Carstairs Index is available linked to CPRD practice postcodes in England, Wales, and Scotland at the LSOA level for the 2011 census [16].

Rural Urban Classification:

It may be important to distinguish between rural and urban areas when investigating differences in social and economic characteristics of small areas. Populations can vary in their composition between urban and rural areas, as can access to services, employment and educational opportunities, and quality of life.

The Rural Urban classifications (RUC) for England and Wales are produced by the Office for National Statistics and are based on census population data [17]. The "Urban" or "Rural" designation is given based on resident population only, and do not reflect the land use, policy, or financial characteristics of an area. Similar classifications exist for Scotland [18] and Northern Ireland [19]. It is important to note that the methodologies and source data for England and Wales differ from those for Northern Ireland and Scotland. This means that the classifications for England and Wales are comparable, however the classifications for England and Wales, Scotland, and Northern Ireland are not comparable.

In England and Wales, the Urban/Rural classification is available at the LSOA level using the 2011 census [17]. In Scotland, the classification is available by data zone for 2016 [18] and, in Northern Ireland, the classification is available at the SOA level for 2015 [16]. The classifications are available as binary (rural/urban) variables, the full classification has up to eight categories, four urban and four rural.

The binary Rural Urban Classification from the 2011 census is available linked to CPRD practice postcodes in England, Wales, Scotland, and Northern Ireland at the LSOA/DZ/SOA level. CPRD may be able to provide different groupings of the categories on request on a study-by-study basis with an approved protocol. Please contact CPRD Enquiries enquiries@cprd.com, before any application, to discuss.

Clinical Commissioning Groups

Clinical Commissioning Groups (CCGs) were created following the Health and Social Care Act in 2012 replacing Primary Care Trusts on 1 April 2013. CCGs are clinically-led, statutory NHS bodies responsible for the planning and commissioning of health care services for their local area. As of 1 April 2020, following a series of mergers, there were 135 CCGs in England [20].

In order to prevent the possibility of deductive disclosure of the location of a practice, researchers will be provided with a CCG pseudonym, rather than the actual geographical identifier, which allows practices within the same CCG to be identified for the purposes of matching and/or cluster level analyses. The CCG pseudonyms are based the November 2020 CCG update.

What data are available through this linkage?

All General Practices which contribute data to the CPRD GOLD and CPRD Aurum databases are eligible for inclusion in this linkage. The PRACTICE postcode is mapped to LSOA, SOA (Northern Ireland only), or datazone DZ (Scotland only) using a postcode lookup file. This then allows linkage to the following small-area level deprivation measures:

- English Index of Multiple Deprivation [7]
 - 2019 composite and individual domains (including the 'Indoor' and 'Outdoor' Living Environment subdomains*)
- Northern Ireland Multiple Deprivation Measure (MDM) [5]
 - 2017 composite and individual domains (including the 'Housing Quality' and 'Outdoor Physical Environment' Living Environment subdomains*)
- Scottish Index of Multiple Deprivation [9]
 - 2020 composite and individual domains
- Welsh Index of Multiple Deprivation [8]
 - 2019 composite and individual domains
- Carstairs Index
 - 2011
- Townsend Deprivation Score
 - 2011
- England and Wales Rural Urban Classification
 - 2011
- Northern Ireland Rural Urban Classification
 - 2015
- Scottish Rural Urban Classification
 - 2016

* There is no 'Environment' IMD domain for Scotland, and no 'Housing' IMD domain for England and Northern Ireland. For studies which include practices in Scotland, England, Wales and/or Northern Ireland, and for which a measure of living environment/housing deprivation is required, the 'Housing' domain (Scotland and Wales) might be used with the 'Indoor' (England) and 'Housing quality' (Northern Ireland) sub-domains.

Linkage based on the patient postcode is available separately for patients registered with a subset of general practices in England. This uses the patient postcode of residence, which is linked via LSOA, to

the English IMD 2019, Townsend 2011 scores, Carstairs 2011 Index, and Rural Urban Classification. These data are described in the documentation on small area data for patients.

Which area-based classification should I use?

Only one practice level classification will be provided for a single study (see section on disclosure control). In deciding which classification is most appropriate, there are several things to consider:

- theoretical considerations: the IMD classifications summarise a larger range of 'domains' of deprivation than Townsend or Carstairs which are intended to focus on material deprivation. It may be helpful to review the literature on the derivation of the different classifications.
- external validity: you may wish to select a classification that allows your results to be most comparable with other published work.

Aside from these issues, it is worth noting that all the available measures are very highly correlated. Spearman's rank correlation coefficients are between 0.94-0.97 for deciles of the different IMD classifications, and between 0.83-0.89 for deciles of Townsend and 0.88-0.90 for decile of Carstairs versus the IMD classifications. This means that for many applications, the choice of area-based deprivation measure is unlikely to have a significant impact on the results or interpretation.

As standard, the most recent deprivation scores will be provided for approved studies. It is worth noting that the available measures are very highly correlated to historic measures. For example, the Spearman's rank correlation coefficient for deciles of the IMD classifications is 0.97 between English IMD 2010 and 2015, 0.98 between Scottish IMD 2009 and 2012, 0.93 between NI 2010 and 2017 and 0.97 between Welsh IMD 2011 and 2014. This means that for many applications, the year of an area-based deprivation measure is unlikely to have a significant impact on the results or interpretation.

Linkages to older versions of the IMD, Townsend, Carstairs, and RUC may be requested on a study-by-study basis with an approved protocol. Please contact CPRD Enquiries enquiries@cprd.com, before any application, to discuss.

Disclosure control

The Rural Urban Classification and deprivation scores and rankings are in the public domain and can be used to identify individual LSOAs. Therefore, for IMD and Carstairs, each LSOA, SOA or DZ is classified into quintiles and deciles of the deprivation score or (equivalently) rank. The score or rank measures are not provided as these may identify the actual LSOA, SOA or DZ of an individual general practice. This ranking is not restricted to the CPRD population, no further processing is done on the data and the quantiles are not weighted.

Technical note: different approaches are available for assigning quantile membership (quintile, deciles) when the number of units to be grouped is not an exact multiple of the number of groups. The quintile and decile groupings were created with the `-xtile-` command in Stata version 13 [21], with the `-nquantile()` option to specify the number of equal sized groups to be created (quintiles=5 groups; deciles=10 groups).

By cross-tabulating two or more classifications it is possible to identify very small groups of lower super output areas (LSOA). To minimise the possibility of deductive disclosure of a patients' area of residence, CPRD will only supply one of the area-based measures for any one study. If you feel you have a compelling justification for using two or more classifications in the same study, you should contact CPRD to discuss your requirements.

Eligibility for inclusion in practice-level postcode linkage

Practices are eligible for inclusion if all the following criteria are satisfied:

- Practice full postcode is available and has a valid format.
- Practice postcode can be linked to a LSOA in England and Wales, SOA in Northern Ireland or data zone (DZ) in Scotland, using the current postcode lookup file.

How often is this linkage updated?

The linkage requires updating when one of the following events occurs:

- A new practice is recruited and begins contributing to patient data to the CPRD GOLD or CPRD Aurum databases.
- An existing practice moves to new premises with a different postcode.
- An LSOA boundary change occurs such that an existing postcode is assigned to a different LSOA.
- An existing postcode designation is changed - for example it may be split into two postcodes as a result of a new residential development being built.
- A new area-based classification system is introduced (e.g. IMD 2019)

The practice level postcode linkage will be updated as part of the monthly database build.

Related files

NHS Postcode Directory (NHSPD) [22]: The Office for National Statistics (ONS) supplies postcode-related data to the Organisation Data Service (ODS). The Organisation Data Service (ODS) is provided by NHS Digital and is responsible for the publication of all organisation and practitioner codes and NHS data standards. NHSPD is updated on a quarterly basis and is available as a downloadable zip file (gridall.csv) from: <https://digital.nhs.uk/organisation-data-service/data-downloads/national-statistics> . The current version used is the 2022 November postcode directory for the deprivation measures, and the 2021 August postcode directory for the CCG linkage.

Linkage summary results

Coverage period for linkage: undefined. The linkage is based on the practice postcode at the time of the last data collection for that practice, and which may change with each monthly database build.

Researchers should bear in mind the time period to which the source data used in the IMD classification relate and consider the extent to which this is valid for their particular study period. Historic postcodes for practices are not maintained.

DOI

Please cite the appropriate DOI in any publications using these data, available: <https://cprd.com/digital-object-identifiers-dois-datasets>

Dataset specification

Index of Multiple Deprivation¹

File name: practice_imd.txt

Column name	Description	Type	Format
pracid	Encrypted unique key given to a practice in CPRD GOLD or CPRD Aurum [primary key]	INTEGER	5
country	1=England, 2=Northern Ireland, 3=Scotland, 4=Wales	INTEGER	4
e2019_imd_5	England: IMD2019: quintile (1=LEAST deprived)	INTEGER	1
e2019_imd_10	England: IMD2019: decile (1=LEAST deprived)	INTEGER	2
ni2017_imd_5	Northern Ireland: MDM2017: quintile (1=LEAST deprived)	INTEGER	1
ni2017_imd_10	Northern Ireland: MDM2017: decile (1=LEAST deprived)	INTEGER	2
s2020_imd_5	Scotland: IMD2020: quintile (1=LEAST deprived)	INTEGER	1
s2020_imd_10	Scotland: IMD2020: decile (1=LEAST deprived)	INTEGER	2
w2019_imd_5	Wales: IMD2019: quintile (1=LEAST deprived)	INTEGER	1
w2019_imd_10	Wales: IMD2019: decile (1=LEAST deprived)	INTEGER	2

File name: imd_domains.txt

Column name	Description	Type	Format
pracid	Encrypted unique key given to a practice in CPRD GOLD or CPRD Aurum [primary key]	INTEGER	5
country	1=England, 2=Northern Ireland, 3=Scotland, 4=Wales	INTEGER	4
e2019_income_5	England: IMD2019 income domain: quintile (1=LEAST deprived)	INTEGER	1
e2019_income_10	England: IMD2019 income domain: decile (1=LEAST deprived)	INTEGER	2
e2019_employment_5	England: IMD2019 employment domain: quintile (1=LEAST deprived)	INTEGER	1
e2019_employment_10	England: IMD2019 employment domain: decile (1=LEAST deprived)	INTEGER	2
e2019_education_5	England: IMD2019 education domain: quintile (1=LEAST deprived)	INTEGER	1
e2019_education_10	England: IMD2019 education domain: decile (1=LEAST deprived)	INTEGER	2
e2019_health_5	England: IMD2019 health domain: quintile (1=LEAST deprived)	INTEGER	1
e2019_health_10	England: IMD2019 health domain: decile (1=LEAST deprived)	INTEGER	2
e2019_crime_5	England: IMD2019 crime domain: quintile (1=LEAST deprived)	INTEGER	1
e2019_crime_10	England: IMD2019 crime domain: decile (1=LEAST deprived)	INTEGER	2
e2019_access_5	England: IMD2019 access domain: quintile (1=LEAST deprived)	INTEGER	1
e2019_access_10	England: IMD2019 access domain: decile (1=LEAST deprived)	INTEGER	2
e2019_environment_5	England: IMD2019 environment domain: quintile (1=LEAST deprived)	INTEGER	1
e2019_environment_10	England: IMD2019 environment domain: decile (1=LEAST deprived)	INTEGER	2
e2019_housing_5	England: IMD2019 housing sub-domain: quintile	INTEGER	1

¹ Please note researchers can only obtain Deprivation data in quintiles OR deciles i.e. e2019_XXX_5 **or** e2019_XXX_10

	(1=LEAST deprived)		
e2019_housing_10	England: IMD2019 housing sub-domain: decile (1=LEAST deprived)	INTEGER	2
e2019_outdoor_environment_5	England: IMD2019 outdoor environment sub-domain: quintile (1=LEAST deprived)	INTEGER	1
e2019_outdoor_environment_10	England: IMD2019 outdoor environment sub-domain: decile (1=LEAST deprived)	INTEGER	2
ni2017_income_5	Northern Ireland: MDM2017 income domain: quintile (1=LEAST deprived)	INTEGER	1
ni2017_income_10	Northern Ireland: MDM2017 income domain: decile (1=LEAST deprived)	INTEGER	2
ni2017_employment_5	Northern Ireland: MDM2017 employment domain: quintile (1=LEAST deprived)	INTEGER	1
ni2017_employment_10	Northern Ireland: MDM2017 employment domain: decile (1=LEAST deprived)	INTEGER	2
ni2017_education_5	Northern Ireland: MDM2017 education domain: quintile (1=LEAST deprived)	INTEGER	1
ni2017_education_10	Northern Ireland: MDM2017 education domain: decile (1=LEAST deprived)	INTEGER	2
ni2017_health_5	Northern Ireland: MDM2017 health domain: quintile (1=LEAST deprived)	INTEGER	1
ni2017_health_10	Northern Ireland: MDM2017 health domain: decile (1=LEAST deprived)	INTEGER	2
ni2017_crime_5	Northern Ireland: MDM2017 crime domain: quintile (1=LEAST deprived)	INTEGER	1
ni2017_crime_10	Northern Ireland: MDM2017 crime domain: decile (1=LEAST deprived)	INTEGER	2
ni2017_access_5	Northern Ireland: MDM2017 access domain: quintile (1=LEAST deprived)	INTEGER	1
ni2017_access_10	Northern Ireland: MDM2017 access domain: decile (1=LEAST deprived)	INTEGER	2
ni2017_environment_5	Northern Ireland: IMD2017 environment domain: quintile (1=LEAST deprived)	INTEGER	1
ni2017_environment_10	Northern Ireland: IMD2017 environment domain: decile (1=LEAST deprived)	INTEGER	2
ni2017_housing_5	Northern Ireland: MDM2017 housing sub-domain: quintile (1=LEAST deprived)	INTEGER	1
ni2017_housing_10	Northern Ireland: MDM2017 housing sub-domain: decile (1=LEAST deprived)	INTEGER	2
ni2017_outdoor_environment_5	Northern Ireland: MDM2017 outdoor environment sub-domain: quintile (1=LEAST deprived)	INTEGER	1
ni2017_outdoor_environment_10	Northern Ireland: MDM2017 outdoor environment sub-domain: decile (1=LEAST deprived)	INTEGER	2
s2020_income_5	Scotland: IMD2020 income domain: quintile (1=LEAST deprived)	INTEGER	1
s2020_income_10	Scotland: IMD2020 income domain: decile (1=LEAST deprived)	INTEGER	2
s2020_employment_5	Scotland: IMD2020 employment domain: quintile (1=LEAST deprived)	INTEGER	1
s2020_employment_10	Scotland: IMD2020 employment domain: decile (1=LEAST deprived)	INTEGER	2
s2020_education_5	Scotland: IMD2020 education domain: quintile (1=LEAST deprived)	INTEGER	1
s2020_education_10	Scotland: IMD2020 education domain: decile (1=LEAST deprived)	INTEGER	2
s2020_health_5	Scotland: IMD2020 health domain: quintile (1=LEAST deprived)	INTEGER	1
s2020_health_10	Scotland: IMD2020 health domain: decile (1=LEAST deprived)	INTEGER	2

s2020_crime_5	Scotland: IMD2020 crime domain: quintile (1=LEAST deprived)	INTEGER	1
s2020_crime_10	Scotland: IMD2020 crime domain: decile (1=LEAST deprived)	INTEGER	2
s2020_access_5	Scotland: IMD2020 access domain: quintile (1=LEAST deprived)	INTEGER	1
s2020_access_10	Scotland: IMD2020 access domain: decile (1=LEAST deprived)	INTEGER	2
s2020_housing_5	Scotland: IMD2020 housing domain: quintile (1=LEAST deprived)	INTEGER	1
s2020_housing_10	Scotland: IMD2020 housing domain: decile (1=LEAST deprived)	INTEGER	2
w2019_income_5	Wales: IMD2019 income domain: quintile (1=LEAST deprived)	INTEGER	1
w2019_income_10	Wales: IMD2019 income domain: decile (1=LEAST deprived)	INTEGER	2
w2019_employment_5	Wales: IMD2019 employment domain: quintile (1=LEAST deprived)	INTEGER	1
w2019_employment_10	Wales: IMD2019 employment domain: decile (1=LEAST deprived)	INTEGER	2
w2019_education_5	Wales: IMD2019 education domain: quintile (1=LEAST deprived)	INTEGER	1
w2019_education_10	Wales: IMD2019 education domain: decile (1=LEAST deprived)	INTEGER	2
w2019_health_5	Wales: IMD2019 health domain: quintile (1=LEAST deprived)	INTEGER	1
w2019_health_10	Wales: IMD2019 health domain: decile (1=LEAST deprived)	INTEGER	2
w2019_crime_5	Wales: IMD2019 crime domain: quintile (1=LEAST deprived)	INTEGER	1
w2019_crime_10	Wales: IMD2019 crime domain: decile (1=LEAST deprived)	INTEGER	2
w2019_access_5	Wales: IMD2019 access domain: quintile (1=LEAST deprived)	INTEGER	1
w2019_access_10	Wales: IMD2019 access domain: decile (1=LEAST deprived)	INTEGER	2
w2019_housing_5	Wales: IMD2019 housing domain: quintile (1=LEAST deprived)	INTEGER	1
w2019_housing_10	Wales: IMD2019 housing domain: decile (1=LEAST deprived)	INTEGER	2
w2019_outdoor_environment_5	Wales: IMD2019 outdoor environment domain: quintile (1=LEAST deprived)	INTEGER	1
w2019_outdoor_environment_10	Wales: IMD2019 outdoor environment domain: decile (1=LEAST deprived)	INTEGER	2

Townsend¹:

File name: practice_townsend2011.txt

<i>Column name</i>	<i>Description</i>	<i>Type</i>	<i>Format</i>
pracid	The encrypted unique key given to a practice in CPRD GOLD or CPRD Aurum	INTEGER	5
uk2011_townsend_5	United Kingdom: Townsend 2011 quintile (1=LEAST deprived)	INTEGER	1
uk2011_townsend_10	United Kingdom: Townsend 2011 decile (1=LEAST deprived)	INTEGER	2

¹ Please note researchers can only obtain Deprivation data in quintiles OR deciles

Carstairs¹:

File name: practice_carstairs.txt

<i>Column name</i>	<i>Description</i>	<i>Type</i>	<i>Format</i>
pracid	The encrypted unique key given to a practice in CPRD GOLD or CPRD Aurum	INTEGER	5
gb2011_carstairs_5	Great Britain: Carstairs 2011: quintile (1=LEAST deprived)	INTEGER	1
gb2011_carstairs_10	Great Britain: Carstairs 2011: decile (1=LEAST deprived)	INTEGER	2

Rural urban classification

File name: practice_urbanrural.txt

<i>Column name</i>	<i>Description</i>	<i>Type</i>	<i>Format</i>
pracid	The encrypted unique key given to a practice in CPRD GOLD or CPRD Aurum	INTEGER	5
e2011_urban_rural	England: 2011 Urban-Rural classification 1=Urban 2=Rural	INTEGER	1
w2011_urban_rural	Wales: 2011 Urban-Rural classification 1=Urban 2=Rural	INTEGER	1
s2016_urban_rural	Scotland: 2016 Urban-Rural classification 1=Urban 2=Rural	INTEGER	1
ni2015_urban_rural	Northern Ireland: 2015 Urban-Rural classification 1=Urban 2=Rural	INTEGER	1

Clinical Commissioning Group pseudonym

File name: practice_ccg.txt

<i>Column name</i>	<i>Description</i>	<i>Type</i>	<i>Format</i>
pracid	Encrypted unique key given to a practice in CPRD GOLD or CPRD Aurum [primary key]	INTEGER	5
ccg	CCG pseudonym	INTEGER	4

¹ Please note researchers can only obtain Deprivation data in quintiles OR deciles

Reference List

- [1] G. D. Smith, E. Whitley, D. Dorling and D. Gunnell, "Area based measures of social and economic circumstances: cause specific mortality patterns depend on the choice of index," *J. Epidemiol. Community Health*, vol. 55, no. 2, p. 149–50, 2001.
- [2] Office for National Statistics, "Census Geography - Office for National Statistics," 2017. [Online]. Available: <https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeography>.
- [3] Office for National Statistics, "Output Areas (OAs)," 2013. [Online]. Available: <http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/output-area--oas/index.html>.
- [4] Office for National Statistics, "2011 Census: Population and Household Estimates for Small Areas in England and Wales, March 2011," 2012. [Online]. Available: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/2011censuspopulationandhouseholdestimatesforsmallareasinenglandandwales/2012-11-23>.
- [5] Northern Ireland Statistics and Research Agency, "Northern Ireland Multiple Deprivation Measures 2017," 2017. [Online]. Available: [https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/NIMDM17-with ns.pdf](https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/NIMDM17-with%20ns.pdf).
- [6] NHS England, "NHS Choices: Patient Choice of GP Practice," [Online]. Available: <http://www.nhs.uk/nhsengland/aboutnhservices/doctors/pages/patient-choice-gp-practices.aspx>. [Accessed April 2019].
- [7] Department for Communities and Local Government, "The English Indices of Deprivation 2015," *Neighb. Stat. Release*, p. 1–20, 2015.
- [8] Statistics for Wales, "Welsh Index of Multiple Deprivation," 2015. [Online]. Available: <https://statswales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation>.
- [9] Scottish Government, "Scottish Index of Multiple Deprivation," 2017. [Online]. Available: <http://www.gov.scot/Topics/Statistics/SIMD>.
- [10] Northern Ireland Statistics and Research Agency, "Northern Ireland Multiple Deprivation Measure 2010," 2010. [Online]. Available: <https://www.nisra.gov.uk/statistics/deprivation/northern-ireland-multiple-deprivation-measure-2010-nimdm2010>.
- [11] P. Townsend, P. Phillimore and A. Beattie, *Health and deprivation: Inequality and the North*, London: Routledge, 1988.
- [12] A. Gartner and N. Lester, "LSOA Townsend deprivation scores calculated from unadjusted Census data," 2008. [Online]. Available: <http://www.publichealthwalesobservatory.wales.nhs.uk/lsoa-townsend-deprivation-scores>. [Accessed April 2019].
- [13] S. Yousaf and A. Bonsall, "UK Townsend Deprivation Scores from 2011 census data," UK Data Service, 2017.
- [14] V. Carstairs and R. Morris, "Deprivation: explaining differences in mortality between Scotland and England and Wales," *BMJ*, vol. 299, no. 6704, p. 886–9, Oct 1989.
- [15] Office for National Statistics, "The National Statistics Socio-economic classification (NS-SEC)," 2017. [Online]. Available: <https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifications/thenationalstatistics socioeconomicclassificationnssecrebasedonsoc2010>.
- [16] B. Wheeler, "Carstairs Index 2011 for LSOAs," 2014. [Online]. Available: <http://reshare.ukdataservice.ac.uk/851497/>.
- [17] P. Bibby and P. Brindley, "Urban and Rural Area Definitions for Policy Purposes in England and Wales: Methodology (v1.0)," 2013. [Online]. Available:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/239477/RUC11methodologypaperaug_28_Aug.pdf. [Accessed April 2019].

- [18] Scottish Government, "Scottish Government Urban Rural Classification 2016," 2018. [Online]. Available: <http://www.gov.scot/Publications/2018/03/6040>.
- [19] Northern Ireland Statistics and Research Agency, "Northern Ireland Urban - Rural Classification 2015," 2016. [Online]. Available: <https://www.nisra.gov.uk/support/geography/urban-rural-classification>.
- [20] NHS Digital, "Change summary - 2020 STP, CCG and Commissioning Hub reconfiguration," 4 March 2020. [Online]. Available: <https://digital.nhs.uk/services/organisation-data-service/change-summary---stp-reconfiguration>.
- [21] Statacorp, *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP, 2013.
- [22] NHS Digital, "Organisation Data Service Postcode Files," [Online]. Available: <https://digital.nhs.uk/organisation-data-service/data-downloads/postcode-files>.