

UK Biobank

Index of Multiple Deprivation Scores

Version 1.0

<http://www.ukbiobank.ac.uk/>

4th May 2018

CONTENTS

- 1. IMD (Index of Multiple Deprivation) scores2
- 2. Licensing2
- 3. Downloading, cleaning and matching the data2
 - Data download2
 - Cleaning the data3
 - Matching the data3

IMD SCORES IN SHOWCASE

1. IMD (INDEX OF MULTIPLE DEPRIVATION) SCORES

As of November 2016, the UKB Showcase allows researchers to download data on the “Townsend deprivation index” for all 502,000 participants at baseline (field 189). This field is very useful for researchers to get a general idea of the level of material deprivation of our participants, and has been requested in over 90% of our projects so far.

However, IMD (Index of Multiple Deprivation) scores are based on more factors than Townsend. It would be interesting to provide them to researchers, as they describe deprivation in a much more detailed way. An entire “Indices of multiple deprivation” folder could be created in Showcase, giving researchers access to the Townsend index (field 189), the general IMD score, as well as the detailed scores for each sub-domain of the IMD.

2. LICENSING

All the IMDs are published using the Open Government License (OGL) version 3.0, with the following conditions:

You are free to:

- *copy, publish, distribute and transmit the Information;*
- *adapt the Information;*
- *exploit the Information commercially and non-commercially for example, by combining it with other Information, or by including it in your own product or application.*

You must (where you do any of the above):

- *acknowledge the source of the Information in your product or application by including or linking to any attribution statement specified by the Information Provider(s) and, where possible, provide a link to this licence.*

This means we can redistribute the IMD data on Showcase, as long as we link back to the government’s website in the field notes.

3. DOWNLOADING, CLEANING AND MATCHING THE DATA

DATA DOWNLOAD

IMD scores are calculated per country in the UK. The data is available as Excel spreadsheets:

- English IMD: <https://www.gov.uk/government/collections/english-indices-of-deprivation>
- Scottish SIMD: <http://www.gov.scot/Topics/Statistics/SIMD>
- Welsh WIMD: <http://gov.wales/statistics-and-research/welsh-index-multiple-deprivation/>

8 files have been downloaded, to cover the full range of recruitment dates in all 3 countries:

- England: 2004, 2007, 2010
- Scotland: 2004, 2006, 2009
- Wales: 2005, 2008

CLEANING THE DATA

The Excel spreadsheets were imported, converted to standard delimited format and harmonised *for each country*. Indeed, it is not possible to match the IMD scores between countries, as they are not calculated based on the same input (variables and domains) and output (scales vary across countries and domains: 0 to 1, 0 to 100, or even distributions standardized around 0).¹

The variable names were made consistent across the different years for each country, resulting in the following subdomains:

| Country | England | Scotland | Wales |
|-------------------|-----------------------------------|---------------------------------|--|
| Subdomains | IMD | SIMD | WIMD |
| | Barriers to housing and services | Crime (<i>only from 2006</i>) | Access to services |
| | Crime and disorder | Current income | Community safety (<i>only from 2008</i>) |
| | Education, skills and training | Education, skills and training | Education |
| | Employment | Employment | Employment |
| | Health deprivation and disability | Geographic access | Health |
| | Income | Health | Housing |
| | Living environment | Housing | Income |
| | | | Physical environment |

MATCHING THE DATA

For each UKB participant, the scores were matched using:

- The corresponding LSOA for the postcode given at baseline.
- The IMD release closest to (and preceding) their year of recruitment for the country, e.g.:
 - o A participant recruited in **England in 2006** was matched to **English IMDs from 2004**
 - o A participant recruited in **England in 2008** was matched to **English IMDs from 2007**
 - o A participant recruited in **England in 2010** was matched to **English IMDs from 2010**
 - o A participant recruited in **Scotland in 2006** was matched to **Scottish IMDs from 2006**
 - o A participant recruited in **Scotland in 2008** was matched to **Scottish IMDs from 2006**
 - o A participant recruited in **Scotland in 2010** was matched to **Scottish IMDs from 2009**
 - o A participant recruited in **Wales in 2006** was matched to **Welsh IMDs from 2005**
 - o A participant recruited in **Wales in 2008** was matched to **Welsh IMDs from 2008**
 - o A participant recruited in **Wales in 2010** was matched to **Welsh IMDs from 2008**

The LSOA at baseline could not be found in the government’s data for:

- 12,261 participants in England
- 70 participants in Wales

The remaining 490,948 participants were all attributed IMD scores based on the matching above.

¹ See <http://www.neighbourhood.statistics.gov.uk/HTMLDocs/analysisandguidance/analysisarticles/indices-of-deprivation.htm> which advises to keep the country-specific IMD scores separate.