

UK Biobank

Address change history data [retired]

Version 2.0

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

June 2024



Contents

##### This data has been retired #####	2
1. Introduction.....	2
2. Source of data	3
3. Data cleaning and derivation	3
4. Reference system.....	4

This data has been retired

Please see [Resource 2060](#) for information about the new participant location history data (Data Fields 32220 – 32224). This resource relates to the formerly available location history, Data Fields 22700 – 22704.

A number of inaccuracies were identified within the address history data fields 22700 – 22704. In particular:

- Some of the addresses are thought to reflect locations where the NHS interacted with the participant rather than their residential address
- A majority of dates were incorrectly processed in the construction of data-field 22700 and are up to 30 months different from the participant's change of address notification
- All 1km values were derived by rounding the 100m values, rather than truncating which is the normal practice for grid coordinates
- Historical values were generally based on postcode alone, giving a lower accuracy level

A new location history, spanning the period from the time a participant was invited to join UK Biobank, up to the present day, has been developed and this is now available in [Category 150](#) (data fields 32220 – 32224).

The following information is retained for information only, for researchers who have already used the old fields in their research.

Introduction

1.1: Some research projects require data on changes to participants' residential address over time. Address details (i.e. postcode) are not disclosed to researchers owing to their potential identifying nature; only designated members of staff have access to this information.

1.2: Instead, UK Biobank has provided grid co-ordinate data-fields, available in two levels of resolution: 100-m and 1-km grid postings. Researchers who request these data are encouraged to use those at the lower resolution (i.e. 1 km), which will be of sufficient granularity for most research purposes. Grid co-ordinates to a 100-m resolution will only be released for specific research projects that fully require this level of granularity.

Source of data

2.1: Data on address change history via GP registration for participants resident in England and Wales are provided to UK Biobank by the Administrative Data Liaison Service (ADLS) [<http://www.adls.ac.uk/adls-trusted-thirdparty-service/>]. The data are managed by NHS Digital [<https://www.digital.nhs.uk/>].

2.2: Data on address change history via GP registration for participants resident in Scotland are provided to UK Biobank by the Information Services Division (ISD) [<http://www.isdscotland.org/>].

2.3: The datasets from NHS Digital and ISD contain the date the address change was entered on the GP system and the postcode attached to the historical residential location. Full address, including the household number and street name, is only provided for the most recent address.

2.4: In addition to the two datasets mentioned above, UK Biobank participants can inform us directly of changes to their address (and are encouraged to do so through the annual newsletter).

Data cleaning and derivation

3.1: Where the full address is present and is verified to be valid, software package DataPlus (provided in the QuickAddress Batch packet), is used to transform the address data into the grid coordinates.

3.2: Where only the postcode is present, the grid coordinates were generated with the aid of on-line mapping tools: doogal [<https://www.doogal.co.uk/>] and uk-postcodes [<https://www.ukpostcodes.com/>]. Further details of this process are available from relevant websites.

Reference system

4.1: The grid coordinate data are provided in the British National Grid (i.e. OSGB1936) projection. OSGB1936 is the Ordnance Survey National Grid geographic reference system (i.e. it is not latitude and longitude).

4.2: The grid co-ordinate measurements refer to easting and northing with a reference point near the Isles of Sicily. The data is usually projected in the units of meters, and there are no negative values (as they are in the case of latitude/longitude convention).