

THE ALSPAC STUDY

Address data

**Prepared by
The ALSPAC study team**

Documentation summarising the data available and the variable names.

March 2019

© University of Bristol

Release file version history

March 2019 - Documentation Update – Removed Tables 1, 2 and 3 and geographical variables available to researchers due to temporal inaccuracy. Contact alspac-data@bris.ac.uk to get up-to-date information on what data is available.

Address data overview

The initial ALSPAC address data was collected at the start of the study upon recruitment of mothers and managed purely as an administrative database permitting researchers to keep in contact with participants. Since this initial address database was constructed the cohort address data has largely been collected and recorded in an ad-hoc manner - with changes/new addresses recorded initially in paper format and more recently electronically – due to the high frequency of questionnaires and clinics for mothers, children and partners and substudy contacts.

ALSPAC address data is collected from participant self-reported residential moves, with the time between a participant contacting ALSPAC and changes to the database typically being 1-3 days. Addresses are keyed into ALSPAC's admin database by members of the admin team in the exact form that they appear from the participant. Tracking and Tracing is a process now used by ALSPAC to follow up and identify participants' residential location where 'addressee unknown' responses are received back from Royal Mail, ensuring that the database is kept up to date as much as possible.

Because of the historical ad-hoc approach data has not been recorded in waves as other cohorts might do (a cross section of all addresses at a particular questionnaire/clinic) but the database has been updated as and when ALSPAC has been made aware of address changes. While this method has permitted extremely high temporal accuracy (more so than if data were collected every few years) it is also possible that some cases may only be legacy addresses; despite addressee unknown responses resulting in immediately 'closed off' (i.e. move out date appended to the date of addressee unknown receipt) residential address fields, some 'current' addresses may no longer be valid as new residents may accept old postage for forwarding or to dispose of. While this should not provide a problem when geographical data is linked with other ALSAPC data (as participants will not likely have returned questionnaires if their contact details were incorrect) researchers should be aware of this when using geographical data in linkage or as standalone data.

International addresses are not available to researchers as they cannot be geocoded accurately and are considered too identifiable.

Partners

Researchers need to be aware that partners address data is not available for release. The reason for this is that address data was held for the pregnancy (ALSPAC being a pregnancy based study) and therefore partners questionnaires and clinic invites were sent out to the mothers address. Partners were not asked to clarify their residence and have only had their own residence recorded since 2008. While researchers may choose to use mothers address data as a proxy for partners the data cannot be validated to a sufficient level to be released by ALSPAC.

Data available to researchers

Variables are also available on a number of UK deprivation indices, as follows:

Table 1: Available deprivation variables

Variable	Label
imd2000	2000 IMD Score
imd2000_income	2000 IMD income score
imd2000_employment	2000 IMD employment score
imd2000_health	2000 IMD health score
imd2000_education	2000 IMD education score
imd2000_housing	2000 IMD housing score
imd2000_access	2000 IMD access score
imd2000_childpoverty	2000 IMD child poverty score
imd2004	2004 IMD score
imd2004_income	2004 IMD income score
imd2004_employment	2004 IMD employment score
imd2004_health	2004 IMD health deprivation score
imd2004_education	2004 IMD education, skills, training score
imd2004_housing	2004 IMD barriers to housing and services score
imd2004_crime	2004 IMD crime and disorder score
imd2004_environment	2004 IMD living environment score
idaci2004	2004 IDACI score
idaopi2004	2004 IDAOPI score
imd2007	2007 IMD score
imd2007_income	2007 IMD income score
imd2007_employment	2007 IMD employment score
imd2007_health	2007 IMD health deprivation and disability score
imd2007_education	2007 IMD education, skills and training score
imd2007_housing	2007 IMD barriers to housing services score
imd2007_crime	2007 IMD crime and disorder score
imd2007_environment	2007 IMD living environment score
idaci2007	2007 IDACI score
idaopi2007	2007 IDAOPI score
idaci2010	2010 IDACI score
idaopi2010	2010 IDAOPI score
imd2010	2010 IMD score
imd2010_income	2010 IMD income score
imd2010_employment	2010 IMD employment score
imd2010_health	2010 IMD health deprivation and disability score
imd2010_education	2010 IMD education, skills and training score
imd2010_barriers	2010 IMD barriers to housing and services score
imd2010_crime	2010 IMD crime and disorder score

imd2010_environment	2010 IMD living environment score
ur01ind	2001 Census urban/rural indicator
ru11ind	2011 Census rural-urban classification
Townsend01	Townsend deprivation score
Townsend01q	Quintiles of Townsend deprivation score

Nb. All deprivation variables are only available recoded as quantiles and are available in quartiles (4), quintiles (5), and deciles (10).

Please note:

- Some variables will not be fully available raw due to low counts of data on certain variables that would risk participant disclosure. In such cases data will either be reverted to missing or sent prior to the main data request with a unique id if researchers wish to derive their own variables.
- Deprivation indices data are recoded as quantiles and are available in quartiles (4), quintiles (5), and deciles (10). Other quantile coding may be possible but may be subject to additional costs.
- The mapping of data to further geographies may also be possible but as with all data requests will depend on executive and ethical approval, and appropriate costs will have to be covered (costs will be provided at the time of approval).
- We will not provide standard administrative data available from other sources, only that which is immediately derived from our address data and cannot be derived by the user.
- We will not provide adjusted data for years between censuses, instead it will be up to individual researchers to decide how best to manage their data in the interim years between the census years.

All project proposals that include a request for address data may be subject to a longer turnaround time than the standard two weeks due to the additional work involved. The length of this extension will depend on the geographic variables required.

Because of the detailed nature of address data each submitted research proposal requesting ALSPAC address data must include detailed information on the reasoning and methodology behind the requested geography, and in the case of high-resolution geographies being requested why smaller resolution data is not acceptable.

Each proposal will be judged on its own merit; previous provision and the availability of geographical data are not a guarantee of data provision; approval rests with the ALSPAC executive and the ALSPAC Law and Ethics Committee (ALEC) and requests for specific geographies may be denied in cases where it is believed participants' disclosure may be at risk.

Please indicate clearly on your proposal form the exact variables you require, based on the list on the previous page and at which time points. Additional costs will apply.