Residency episodes

- This is the starting point for all ALPHA analyses and is therefore essential for all study sites.
- It is used to compute person-year denominators for age-specific rates.
- We expect to see one record per episode of household residence for each individual in the data set, i.e. on average more than one record per person
- Only those individuals who have been resident continuously in the same household between first and last date of observation will have only one residence episode record.
- Individuals who have moved household within the DSS area, or have left and returned to the DSS area since the time they were first seen, will have two or more records, depending on the number of periods of absence from the study area and the number of times they have moved household. In this case, entry and exit dates and types in the classification below refer to the start and end of an episode of residence rather than to the first and last encounter with the individual in the study throughout the whole of his/her life.
- For records relating to consecutive residence episodes, where an individual moves within the study area, the (household) exit date of the earlier episode should be equal to the (household) entry date of the later episode.
- For records relating to individuals who moved out of the study area and then moved back in, the entry date of the later episode must be strictly greater than the exit date of the earlier episode.

ALPHA Network data specification: residency episodes

Variable name	Description	Type	Coding	Notes
study_name	The ALPHA name for your study site	string	site specific	Character – please be consistent across data sets
idno	Person ID number	string	site specific	Numeric IDs long integer format, unique for an individual
hhold_id	Household ID number	string	Site specific	Geographical location, Should be unique for each household
hhold_id_extra	Household ID number	string	Site specific	If your site has another definition for Households (e.g. social units) include it here
sex	Male or female	numeric	1 Male 2 Female	Must not vary between residence episodes
dob	Date of birth- best estimate	numeric (double)	in Stata format (days since 1st Jan 1960).	If actual month and day are not known it is OK to impute, e.g. assign to middle of the month or mid-year Must not vary between residence episodes
residence	Type of area within DSS	numeric	site-specific grouping, we expect most sites to have 2 to 4 categories	Aim to distinguish urban / rural, or among rural areas distinguish remote / roadside, or by dominant industry
entry_date	Date of start of residence episode	numeric (double)	in Stata format	This date should be known quite accurately – it could be the date of a household interview or in between two consecutive household interviews
entry_type	Type of entry	numeric	1 baseline recruitment 2 birth 3 in-migration 4 became eligible for study 5 start membership episode	
exit_date	Date of end of residence episode	numeric	in Stata format	This date should be known quite accurately. For individuals remaining resident in the study area this will be date of last household interview
exit_type	Type of exit	numeric	1 present in study site 2 death 3 out-migration 4 lost to follow-up 5 became ineligible for study 6 end membership episode	Code 1 refers to those alive and still resident in study area – they will be treated as censored at last interview date when known to be in study population Code 5 is for sites which have ceased follow up on certain people/households/areas after a set point.

9th April 2019