Scottish Longitudinal Study Disclosure Control Protocol

This document must be read by all persons wishing to use SLS data

The SLS is a linked database containing individual confidential data, managed by the Longitudinal Studies Centre – Scotland (LSCS). Any person using that data, whether in the role of a SLS SO providing data to a user, or as an end user receiving results of data for research, must comply with the confidentiality requirements stated in the undertaking.

Further to this, certain disclosure controls must be applied to the data to ensure that no individual can be identified within them. These controls will be applied to any data released to users by support personnel. If data are to be released in tabular form then the SO must ensure that any variables that alone, or in conjunction with others, may identify individuals are aggregated to the point where no identification is possible.

- No data on the birth dates of SLS members may be released, with the exception of year of birth. Where full date of birth is required for use in derivations (i.e. in such procedures as person years at risk analysis) only those SLS staff based at NRS with full database permissions will be allowed access to the data.

- Exposure times (e.g. person years at risk) may be included in aggregated datasets provided ***either*** there is more than 10 events in each cell ***or*** else the data has been subject to adjustment to prevent disclosure.

When releasing tabular data SLS support staff must ensure that cell counts are 10 or over for Intermediate and Final Outputs. If associated data allows the cell to be split then the support person must aggregate the data to the highest level consistent with the need to explain the results.

- Sample uniques are never allowed.

- Reporting residual values that identify individual cases will not be allowed when releasing data from statistical models.

- Particular care should be exercised with plots especially where there are outliers or extreme values. These should not be released.

- Histograms and bar charts can be reported if, and only if, the same data in frequency table form could be reported.

Restrictions will also be placed on the release of any variable deemed to be sensitive. These include variables which relate to small numbers of people in Scotland (i.e. local-area geographic identifiers, detailed ethnicity, rare causes of death etc.). Other variables, such as religion, may also be treated as sensitive, depending on the context of the research. It should be noted that selection criteria used in extracting data such as sex and age may be disclosive when used in conjunction with other variables.

If a SO believes that data may be disclosive they must bring this to the attention of the SLS data custodian who will decide on the procedure to follow. In most cases this will require further aggregation of the data.