

SLS APPLICATION FORM

#### The Longitudinal Studies Centre – Scotland (LSCS) is a Scottish centre of excellence in the creation and use of longitudinal data. It is responsible for the Scottish Longitudinal Study (SLS).

#### Please look at the [SLS-DSU website](https://uoe.sharepoint.com/sites/SLScoreteam/Shared%20Documents/SLS%20forms/%28http%3A/sls.lscs.ac.uk/guides-resources/step-bystep-guide-to-accessing-sls-data-1) for steps involved in accessing SLS data. The completion of this form is step four of 'Guides & Resources' Step-by-Step Guide'.

**Your Privacy**

The information you provide on this form will be treated in the strictest confidence by SLS and NRS staff. An electronic copy of this form will be securely stored with NRS and the SLS-DSU. A hard copy will be securely stored in the SLS-DSU. For further information on how your data is held and used, see:

[www.nrscotland.gov.uk/record-keeping/legislation/primary-information-legislation/data-protection](http://www.nrscotland.gov.uk/record-keeping/legislation/primary-information-legislation/data-protection)

<https://sls.lscs.ac.uk/more/privacy-statement/>

Some key information is given on pages 2-10 of this document including FAQs to provide you with information and guidance about the SLS data sets.

* Please read pages 2-10 before you start to complete the form.
* Speak to your SLS Support Officer for advice on selecting your research sample of interest.
* Read through and familiarise yourself with the [Census forms](https://calls.ac.uk/guides-resources/census-form) so that you are aware of how questions were asked, changes between Censuses and question routing ie some questions are not completed by everyone e.g. working age only, some questions are not present at all censuses, some questions are fundamentally the same but differ in wording or available responses.
* To select your research variables, use the [SLS data dictionary](https://sls.lscs.ac.uk/variables/#load) online including the ‘more’ tab and create an online account in order to use ‘save’ tabs. With minor editing (including sorting by table and variable name) the saved project can be cut and pasted into this form. Additionally, there is an Access version downloadable from the link above which may be useful to search through all category/variable labels. However, logging in and creating an account is best for completing this form.

## Overview of the main SLS study structure

## Additionally, the SLS Birth Cohort of 1936 ([SLSBC1936](https://calls.ac.uk/2017/03/30/new-1936-birth-cohort-study/)) is now available. The SLSBC1936 includes: the Scottish Mental Survey of 1947 (SMS1947 - a cognitive ability test that included almost all Scottish children born in 1936) has been linked to the 1939 National Register, the NHS Central Register and the SLS. The outcome of the project is a powerful life-course dataset containing information from childhood to old age. For further information on this cohort please read the [blog and working paper](https://calls.ac.uk/output-entry/the-scottish-longitudinal-study-1936-birth-cohort/)

## The GP residential history data are also now available. Postcodes of residence and dates of new GP registrations are available for 2000-2015 (Table E10 in the [SLS data dictionary](https://sls.lscs.ac.uk/variables/#load) online), allowing study of internal migration and the type of areas that members have lived in. Please note that you will probably be using variables derived from those listed in E10 so speak to your support officer.

## Overview of the SLS Restricted variables – 3 levels of restriction



Please note that the day field of date variables have restriction level 1 and are not available to external researchers however the month and year fields are available so in the variable selection box write either ‘month and year fields only’ or ‘day field available for SLS staff’ as appropriate. For deaths only, if you require the ‘day’ part for date of death please discuss with SLS staff to make the case.

## Overview of the SLS tables – from the online [Data Dictionary](https://sls.lscs.ac.uk/variables/#load)

## Data contained in the SLS Database for members and non-members

| **Source** | **Short Table name** | **Description** |
| --- | --- | --- |
| General | CORE | Core contains info. from the source from where the SLS member ***first*** entered the SLS (i.e. birth, immigration, Census). File containing unique identifier for each SLS member, sex, year of birth, presence at Census & tracing indicators. *We recommend that you select variables such as sex from the Census tables*. |
| 1991 Census | 91MEMBER | 1991 Census person data on SLS members |
| 1991 Census | 91NONMEMBER | 1991 Census person data on SLS non-members (people living with SLS members) |
| 1991 Census | 91HOUSEHOLD | 1991 Census household data for SLS members and non-members |
| 1991 Census | 91COMMUNAL | 1991 Census communal data for SLS members living in communal households  |
| 1991 Census | 91FAMILY | 1991 Census data on persons living in families |
| 1991 Census | 91LSRELAT | 1991 Census data on relationships to SLS members in households |
| 2001 Census | 01MEMBER | 2001 Census person data on SLS members |
| 2001 Census | 01NONMEMBER | 2001 Census person data on SLS non-members (people living with SLS members) |
| 2001 Census | 01HOUSEHOLD | 2001 Census household data for SLS members and non-members |
| 2001 Census | 01COMMUNAL | 2001 Census communal data for SLS members living in communal households  |
| 2001 Census | 01LSRELAT | 2001 Census data on relationships to SLS members in households |
| 2011 Census | 11MEMBER | 2011 Census person data on SLS members |
| 2011 Census | 11NONMEMBER | 2011 Census person data on SLS non-members (people living with SLS members) |
| 2011 Census | 11HOUSEHOLD | 2011 Census household data for SLS members and non-members |
| 2011 Census | 11COMMUNAL | 2011 Census communal data for SLS members living in communal households  |
| 2011 Census | 11FAMILY | 2011 Census data on persons living in families |
| 2011 Census | 11LSRELAT | 2011 Census data on relationships to SLS members in households |
| Vital Events | BIRTH | Births of new SLS members |
| Vital Events | DEATH | Deaths of SLS members  |
| Vital Events | DEATH | Deaths of SLS members’ spouses |
| Vital Events | BIRTH | Live Births to female SLS members |
| Vital Events | BIRTH | Live Births to male SLS members |
| Vital Events | STILLBIRTH | Stillbirths to female SLS members |
| Vital Events | STILLBIRTH | Stillbirths to male SLS members |
| Vital Events | INFANT MORTALITY | Infant deaths to child of SLS mother |
| Vital Events | INFANT MORTALITY | Infant deaths to child of SLS father |
| Vital Events | MARRIAGE | Marriage of female SLS member |
| Vital Events | MARRIAGE | Marriage of male SLS member |
| --- | IMMIGRANT | Immigrations of new SLS members from outside UK |
| --- | MIGRATION | Migrations of SLS members (into or out of Scotland)  |
|  | Residential location data (created from GP histories, 2000 onwards). | NHS Postcode History for SLS Members (from 2000 onwards) |
| --- | Education dataTables H01 to H13 | Including: School Census, School Attendance, School Exclusions, SQA Attainment and SQA Qualifications |
|  | 1936SLSBC Tables M01 to M04 | Tables relating to the 1936 SLS Birth Cohort (1936SLSBC) including: SMS1947 and 1939 National Register |
| --- | ECOLOGICAL | Area-based variables about deprivation, population density and urban-rural indicator – please provide the geographic level eg datazone |
| --- | POLLUTION  | Pollution Variables  |
| --- | WEATHER  | Weather Variables  |

## Notes on requesting area-level data

For Ecological data such as deprivation, population density, SIMD and urban-rural please ensure you request which geographical level at which you require the data in section B1.7. Details of these variables can be found in the Ecological Table (G01) in the [SLS data dictionary](https://sls.lscs.ac.uk/variables/#load) online. However, only SIMD 2004 is available so you may wish to source the relevant SIMD that corresponds to your research cohort as it is freely available online.

For weather or pollution data please ensure that you specify the frequency (e.g. monthly, annual, day of Census) and geographical level at which you require the data in section B1.7. Details of these variables can be found in the Weather Table (G03) and Pollution Table (G02) of the [SLS data dictionary](https://sls.lscs.ac.uk/variables/#load) online. However, they are not up to date (and will not be updated as it is freely available online) – as such, it may be best to contact the relevant agency and bring in your own data.

## NHS Health data

For NHS health data linkage to the SLS (including Cancer registrations SMR06, Hospital admissions SMR01 and Maternity Inpatient and Day Case SMR02), information on getting started with health data and what is available for research use, please look at the online [Health and Social Care Data Dictionary](https://www.ndc.scot.nhs.uk/Data-Dictionary/) along with the downloadable [‘NHS Scotland PDF Pack’](https://sls.lscs.ac.uk/wp-content/uploads/ISDPDFpack1-downloaded-Feb2018.pdf) (10.4Mb). Note that if any health data linkage is required you will have to apply to the Public Benefit and Privacy Panel (HSC-[PBPP](https://www.informationgovernance.scot.nhs.uk/pbpphsc/home/for-applicants/)) for Health and Social Care after you have SLS project approval. You will need an Electronic Data Research and Innovation Service ([eDRIS](http://www.isdscotland.org/Products-and-Services/EDRIS/)) Research Coordinator to assist you with the HSC-PBPP application. Your SLS Support Officer can also advise you and should be copied in on your discussions with your eDRIS Research Co-ordinator. For your convenience, we have prepared a Part filled HSC-PBPP application form which can be found at Step 7 on the [SLS Website](https://sls.lscs.ac.uk/guides-resources/step-bystep-guide-to-accessing-sls-data-1/) which is regularly updated to correspond to the most recent PBPP Application Form on the PBPP website. More information on completing the PBPP application form can be found on the [PBPP](http://www.informationgovernance.scot.nhs.uk/pbpphsc/home/for-applicants/) website.

***NHS Central Register (NHSCR) data***

The SLS holds various variables that are either directly or derived from NHSCR data. Please note that your project must have an impact on health (directly or indirectly) as a condition of using variables sourced from NHSCR. This relates to variables from tables E06 (Immigrant), E08 (Migration) and E10 (Postcode). Please speak to SLS staff for more information once you have selected your variables in order to undertake the research.

## SLS working papers

In order to understand how the SLS was created, how large it is, along with tables with breakdowns by age, sex and marital status it may also be useful to read the [technical working papers](https://sls.lscs.ac.uk/outputs/working-papers/) on the SLS:

* [*The SLS, An Introduction*](http://calls.ac.uk/wp-content/uploads/2013/05/LSCS-WP-1.0.pdf)
* [*The SLS, Tracing rates and sample quality for the 1991 Census SLS sample*](http://calls.ac.uk/wp-content/uploads/LSCS-WP-2.0.pdf)
* [*The SLS, A technical guide to the creation, quality and linkage of the 2001 Census SLS sample*](http://calls.ac.uk/wp-content/uploads/LSCS-WP-3.0.pdf)
* [*The SLS, The 1991 - 2001 Scottish Longitudinal Study Census Link*](http://calls.ac.uk/wp-content/uploads/LSCS-WP-4.0.pdf)

**Overview of commonly selected research variables** – from the online [Data Dictionary](https://sls.lscs.ac.uk/variables/#load)

|  |  |  |
| --- | --- | --- |
| **Description** | **Variable name** | **Table name** |
| Unique identifier allocated to SLS member at entry to SLS | SLSNO | A01 / CORE[[1]](#footnote-2) |
| Year of birth | DOBYRDOBYRT9DOBYR0DOBYR1 | A01 / COREC10 / 91MemberC20 / 01MemberC30 / 11Member |
| Sex | SEXSEXTEN9 SEX0SEX1 | A01 / COREC10 / 91MemberC20 / 01MemberC30 / 11Member |
| Age at 1991 censusAge at 2001 censusAge at 2011 census | AGETEN9 AGEP0AGEP1 | C10 / 91MemberC20 / 01MemberC30 / 11Member |
| Presence at 1991 censusPresence at 2001 censusPresence at 2011 census | HISCEN9HISCEN0HISCEN1 | A01 / COREA01 / COREA01 / CORE |
| Tracing indicators (way of entry into and out of the SLS) | TRACEINTRACEOUT | A01 / CORE |
| Indicator whether a person lives in a private household or in a communal establishment | 1991: RECTYPE92001: PERTYP02011: RESIDENCE\_TYPE1 | C10 / 91Member and C11 / 91NonMemberC20 / 01Member and C21 / 01NonMemberC30 / 11Member and C31 / 11NonMember |
| Relationship(1991: of all household members to the head of household)(2001: of all household members to one another)(2011: of all household members to one another) | 1991: RELAT92001: RELAT02011: RELTSHIP1 | C10 / 91Member and C11 / 91NonMemberC20 / 01Member and C21 / 01NonMemberC30 / 11Member and C31 / 11NonMember |
| Relationship to the SLS member | LSRELAT9LSRELAT0LSRELAT1 | C15 / 91LSRELAT C25 / 01LSRELATC35 / 11LSRELAT |

*Useful info for selecting variables and selecting a sample:*

**1991 Census ‘SLS recoded variables’**

The 1991 Census took place before advances in machine readable forms and was transcribed. The ‘difficult to code’ Census information (e.g. occupation) was only captured electronically for 10% of the population and the SLS-DSU needed to transcribe these data for the remainder of our sample. In fact, we transcribed the information for our entire SLS sample, to allow us to check our results against the original transcribing for 10% of the cases and this procedure revealed some coding errors in the original Census transcription/coding. Both the original Census (10%) and the SLS recoding (100%) are available in the Data Dictionary. As a general rule we recommend that users select the variables with the ‘***SLS recoded variables***’ in the description (if available) these can be identified by ‘t9’ or ‘ten9’ at the end of the variable name. For example, AGETEN9 rather than AGE9 and ECONPOT9 rather than ECONPO9. However, not all 1991 variables were part of the original 10% sample transcribed and hence there are not two versions of all the 1991 Census variables.

**Dates of Birth from different sources and selecting a sample**

As a result of the SLS recoding (see section above), unless otherwise specified the SLS-DSU will create requested research samples based on age using the Census tables for a conditional age range. For example, age 18+ in 1991 would use the year part of Date of Birth DOBYRT9 [Year of birth (***SLS recoded***) from 1991 Census] as this is deemed more accurate than DOBYR9 [Year of birth from 1991 Census] or the Core table variable DOBYR.

Since information held in the Core table is the information as received when an SLS member first enters the SLS (i.e. birth, migrant, Census), for a small proportion of members, human error – either accidental misreporting of information or transcription errors from Census – may mean that the Date of Birth for sample members found in Core table is different from the date of birth in the census table. For the 1991 Census, the Date of Birth in the Core Table would match the 1991 Census date of birth but may not match the SLS coded Date of Birth for the 1991 Census. This means that a small proportion of people are entered into the SLS sample as having one of the 20 birth dates but it is subsequently found that there was an error in the birth date and it is not one of the 20 birth dates. Such people are retained as SLS members however there is no linkage to vital event data for these people (linkage based on the 20 dates). Please see ‘extra deaths’ below.

The implications of these differences depend on how your sample is specified, i.e. whether selected using variables from the Core table where members first entered the SLS, or from Census information. For example, if you request two (or more) versions of the Date of Birth variables (eg DOBYR1, DOBYR) and run cross-tabulations you will see minor differences. This is not an issue for other Censuses, as there is only one Date of Birth variable.

**2011 Imputation flags**

There are no missing data for some 2011 variables as the 2011 variables were released to SLS after the Census Team had carried out edit and imputation ie missing data had been replaced with non-missing data (imputation) or had been edited if data were found to be not consistent. Imputation flags indicating if values have been edited or imputed are available for some primary variables (ie variables that are directly obtained and not derived from the census questionnaires) with extension \_IMP. Researchers may wish to select these variables to either exclude edited/imputed values or carry out sensitivity analyses. Please note that census data before edit and imputation are not available.

**Extra Deaths as recorded by NHSCR (table E09)**

Deaths from the Vital Event data are added to the SLS via the NHS Central Register (NHSCR)[[2]](#footnote-3). All SLS members are flagged on the NHSCR on entry to the SLS be it via a census record, a birth record or an immigration (immigration is captured in the SLS by someone registering with a Scottish GP – ie recorded on the NHSCR). Vital Event death records are linked to the SLS for all individuals with one of the 20 SLS birth dates, however some cases will not be matched by this process.

SLS members can fail to be linked to their death record for a number of reasons;

1. **Identifying data used for linkage does not match** - if the identifying data (name, date of birth and sex) are different on the death record compared to the NHSCR then the records may not be matched. For example, if names are spelled differently, there has been a name change eg due to marriage or divorce which NHSCR has not been notified of or discrepancies in date of birth or sex.
2. **SLS members emigrate** - Only deaths which are recorded in Scotland have a vital event record that can be passed to NHSCR for matching, therefore if an SLS member has left Scotland and is living elsewhere when they die, their death cannot be traced using the normal method.
3. **SLS members subsequently found to not have one of the 20 SLS date of births** - In some cases, an individual is selected as an SLS member based on the date of birth which is subsequently found to be incorrect (ie not one of the 20 SLS birth dates). In these cases, the individual remains in the SLS sample but cannot be linked to the Vital Event data as only vital events death records with an SLS date of birth are received.

To try and trace these missing deaths, NHSCR data can be used. NHSCR in England can provide death details for SLS members who have emigrated to England (category 2). NHSCR in Scotland and England can provide recorded death details for individuals who fall into categories 1 and 3. Deaths that are identified from NHSCR data rather than vital events are termed ‘extra deaths’.

We encourage all SLS users interested in deaths to request this information to correctly account for people who died outside Scotland within the NHS system or who entered the SLS without one of the 20 SLS birth dates (for example, from incorrect details given at Census compared with NHSCR information). In practice these extra deaths as recorded by NHSCR accounts for just over 200 deaths per year for the period 1991-2010.

The 3 most useful variables are:

**DOD\_ND** - Date of Death of SLS member, as held by NHSCR (month & year only)

**LOCATION\_ND**  - Location of Death of SLS member, as held by NHSCR. Recorded as: Abroad, England/Wales, Northern Ireland or Scotland

**DEATH\_IND\_ND**  - Indicates whether the death is recorded by the SLS as well as NHSCR or NHSCR only.

*Limitations*

At present, we do not receive ICD-9 or ICD-10 codes for these extra deaths as recorded by NHSCR so information on cause of death is not available.

**Other often-used Census variables in the SLS**  – from the online [Data Dictionary](https://sls.lscs.ac.uk/variables/#load)

***Note****: variables in 1991, 2001 and 2011 are* ***not*** *necessarily comparable; they may have a different definition or different categories. Please check before you use these variables. \* denotes a Restricted variables*

|  |  |  |
| --- | --- | --- |
| **Topic** | **Description** | **Variable Name** |
| **1991** | **2001** | **2011** |
| Personal | Marital status | MSTATT9 | MSTP0 | MSTP1 |
|  | Student indicator | ECONSTU9 | STUP0 | STUP1 |
|  | Term time address indicator | TERMIND9 | TTIND0 | TTIND1 |
|  | Country of birth | COB9 | COBP0 | COBP1 |
|  | Dependent child indicator | DEPNCHN9 | DCHP0 | DCHP1 |
|  |
| Accommodation | Household tenure | TENURE9 | TENH0 | TENH1 |
|  | Owns/rents | --- | OWNRENT0 | OWNRENT1 |
|  | Landlord | --- | LANH0 | LANH1 |
|  | Number of cars | CARS9 | CAVH0 | CAVH1 |
|  | Central heating | CENHEAT9 | CNHH0 | CENHEATH1 |
|  | Availability of bath/shower | BATHSHW9 | BTSH0 | --- |
|  | Building type | BLDTYPE9 | ACCH0 | ACCH1 |
|  | Number of rooms | ROOMS9 | NORH0 | NORH1 |
|  |
| Communal establishment | Establishment type | ESTABTY9 | CETC0 | ESTNATURES1 |
|  | Position in communal establishment | RESCLAS9 | POSP0 | POSP1 |
|  |
| Industry | 1991 coding frame | IND919 | --- | --- |
|  | 2001 coding frame | IND019 | INDP0 | --- |
|  | 2011 coding frame | --- | --- | INDP1 |
|  |
| Occupation | SOC 90 code | SOC909 | SOC900 | --- |
|  | SOC 2000 code | SOC20009 | SOC20000 | --- |
|  | SOC 2010 code | --- | --- | OCCP1 |
|  | Employment status | EMPST9 | EMPST90, EMPST0 | EMPSTP1 |
|  | Supervisor status | --- | SUMP0 | SUMP1 |
|  | Company size | INDSIZE9 (10%) | SIZP0 | --- |
|  | Economic activity | ALWEEKT9 | ECOP0 | ECOP1 |
|  | Economic position | Econpot9 | --- | --- |
|  |
| Socio-economic | Social class | SCLAS9 | SCLAS90, SCLAS0 |  |
|  | Social Grade (approximated) |  |  | SCGP1 |
|  | Socio-economic group | SEG909 | SEG9090, SEG900 |  |
|  | Ns-sec code | NSSECR9 | NSSECR0SOCNSSP0 | NSSECR\_20101 |
|  |
| Qualifications | Qualifications | --- | QUPS10/20/30/40/50/60/80 | QUPS11/21/31/41/51/61/81/91/101/111 |
|  | Professional qualifications | --- | QUPS70 | QUPS71 |
|  | Highest qualification | QMLVHQT9 | HLQP0 | HLQP1 |
|  |
| Health | Long-term illness | LTILL9 | ILLP0 | ILLP1 |
|  | Carer | --- | HELP0 | HELP1 |
|  | General health | --- | HEAP0 | HEAP1 |
| Ethnicity | Ethnicity  | ETHNIC9 | ETHGRP0, ETHP0 | ETHP1 |
|  |
| Religion | Practising religion | --- | RELIGP0 | RELPS1 |
|  | Raised religion | --- | RELIGR0 | --- |
|  |
| Family and household | Household size | DJHH9 | SIZH0 | SIZH1 |
|  | Relationship  | RELAT9 (to the head of household) | RELAT0, RELP0 (within household) | RELTSHIP1(within household) |
|  | Relationship to the SLS member | LSRELAT9 | LSRELAT0 | LSRELAT1 |
|  | Head of household / household reference person | HDHEADT9 | HRPP0 | HRPP1 |
|  | Head of family / family reference person | FAMHEDT9 | FRPP0 | FRPP1 |
|  | Family status | FAMPOHT9 | FMSP0 | FMSP1 |
|  | Family type | FAMTYPT9 | FMTFAM0 | FMTFAM1 |
|  | Generation in family | GENINFM9 | GENINF0 | GENINF1 |
|  | Household composition | HHFAMTYT9, FHHTYPT9 | HHCH0 | HHCH1 |
|  | Number of dependent children in household | HDPCHAT9 | DPCH0 | DPCH1 |
|  | Minimal household type | MHUTYPE9 | MHUTYPE0 | MHUTYPE1 |
|  | Position of person within MHU | MHUPOS9 | MHUPOS0 | MHUPOS1 |
|  | Combination of MHUs within household | MHUPOS9 | MHUPOS0 | MHUPOS1 |
|  | Number of families in household | FAMINHT9 | FMCNT20 | FAMINHH1 |
|  | Number of adults in household | ADULTS9 | ADTH0 | ADTH1 |
|  | Number of adults in household in employment | EMPERSH9 | ADEMH0 | ADEMH1 |
|  | Age of youngest dependent child in household | AGYOCH9 | AGYOCH0 | AGYOCH0 |
|  |
| Migration | Migrant indicator (current residence is different than residence one year before) | MGPOP9, URONIND9 | MIGP0 | MIGP1 |
|  | Wholly moving household indicator | WMOHHIN9 | MIGH0 | MIGH1 |
|  | Enumeration postcode | ENUMPC9\* | ENUMPC0\* | ENUMPC1\* |
|  | Usual address one year ago (postcode) | URONDET9\* | MIGPCP0\* | MIGPCP1\* |
|  | Migration distance (current residence to residence one year before) | MPDOM9 | MIGDOMA0 | MIGDOMA1 |
|  | Migrant indicator (1991 to 2001) |   | MBCFLAG0 | --- |
|  | Migration distance (1991 to 2001) |   | MBCDOM0 | --- |
|  | Migrant indicator (2001 to 2011) |   |  | MBCFLAG1 |
|  | Migration distance (2001 to 2011) |   |  | MBCDOM1 |

Please refer to the SLS-DSU website for the SLS governance arrangements that can be downloaded [here](https://sls.lscs.ac.uk/wp-content/uploads/SLS-Governancev3.2_noheader.doc) which are used to assess a SLS research proposal.

**PART A: USER AND STUDY DETAILS**

Section A1 Application details

*Study title*

*(this title should be a brief informative summary of the research proposal)*

## Head of proposed study

|  |  |
| --- | --- |
| Title:  | Name:  |
| Position:  |
| Organisation:  |
| Address:  |
| Telephone:  | Date and provider/organisation of last Safe Researcher Training (SRT):  |
| Email:  |

### *Names and contact details of all members of the research team*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Name | Position | University/ organisation | Email & Telephone | Date of last IG training | Main Contact /chief researcher (please tick) | Will access SLS data(tick) |
| 1 |       |       |       |       |  |  |  |
| 2 |       |       |       |       |  |  |  |
| 3 |       |       |       |       |  |  |  |
| 4 |       |       |       |       |  |  |  |
| 5 |       |       |       |       |  |  |  |
| 6 |       |       |       |       |  |  |  |

### *Expected duration of project (dd/mm/yyyy)*

#### From: Click or tap to enter a date. To: Click or tap to enter a date.

Please ensure that you propose a realistic end date. In order to comply with current guidance on data retention in line with GDPR, the SLS-DSU will archive your project extract data for a period of 5 years beyond the stated end date, after this it will be deleted. This is to enable if need be, to further revise models for publications etc. If you require access to your extract beyond the stated end data a Project Extension Form needs to be completed (and PBPP amendment if health data is linked to the extract).

#### Section A2

#### Funding details

It is not necessary to have funding to conduct research with the SLS, as the data are freely available to academic researchers. However, many researchers request funding to undertake projects using the SLS.

*Source of funding (if applicable)*

*Is funding confirmed?* (if yes, tick box) [ ]

#### Section A3 Background and aims of study

#### The information given here in section A3.1 will be published on the SLS-DSU website at: <http://sls.lscs.ac.uk/projects/> if the project is approved

A3.1 *Explain the precise aims of your study (up to approx. 250 words)*

This should describe the specific aims of your project and must include the hypothesis that you hope to test and/or your detailed research questions.

Type or paste your text in the box below.

|  |
| --- |
|  |

A3.2 *Ethical considerations (up to approx. 100 words)*

Please detail any ethical issues which may arise from your project.

We now require that all academic users require ethical approval from your institution. If you are a non-academic user, we can work with you on this part. Please note this subsection will not be included on the SLS-DSU website, only A3.1.

|  |
| --- |
|  |

A3.3 *Give an account of the background to your planned study (up to approx. 400 words)*

This should include a brief summary of what is known to date about your chosen topic and what and how your study contributes to what is known to date.

|  |
| --- |
|  |

A3.4 *References (as required for section A3.3 above)*

|  |
| --- |
|  |

A3.5 *Explain why the SLS data are useful for this study (up to approx. 200 words)*

(You may wish to explain why other datasets are not as useful as the SLS for the study you are planning)

|  |
| --- |
|  |

A3.6 *For projects using NHSCR variables please indicate how the research proposal relates to Health research (up to approx. 100 words)*

|  |
| --- |
|  |

#### Section A4 Dissemination and Impact

**A4.1** *Please indicate how you will publicise and disseminate the findings of your proposed study and identify any deadlines.* (e.g., plans for journal papers or conference presentations)*. (up to approx. 200 words)*

|  |
| --- |
|  |

**Would you be happy to give a presentation if requested to Scottish Government?**

Yes [ ]  No [ ]

**A4.2** *Please indicate how your research will have impact both for academia and for a wider audience (e.g., Government and third-sector bodies). Give some idea of exactly who might benefit from the work, and the steps you plan to take to achieve this impact. For example, if your work could influence policy-making, you might plan to write a research briefing for a Government department, give a presentation to your target audience, or offer advice on a consultancy basis.*

Your Support Officer can advise you on completing this section. *(up to approx. 200 words)*

|  |
| --- |
|  |

**Would you be happy to share your project code for reuse?**

Yes [ ]  No [ ]

**A4.3** *The SLS-DSU is funded by the ESRC, due to the restrictions on datasets which form part of the SLS we cannot deposit research data with the UK Data Service (data archive). As such, we are asking researchers to share code used for data management or modelling for reuse by others. This could be all project syntax or key parts (ie specific analysis). Code from all statistical packages accepted. The SLS-DSU will then deposit code for reuse in GitHub (or similar).*

|  |
| --- |
|  |

*If you are not willing to share code please give details below:*

#### PART B: SPECIFYING EXTRACT AND PRACTICAL ARRANGEMENTS

#### Please complete this section of the form with the help of your Support Officer.

#### Section B1 Study population

**B1.1** *Provide a description of the population selected for the study*

(For example, all male SLS members enumerated at the 1991 and 2001 Censuses and aged 16 – 64 years at the 1991 Census).

If your project requires more than one sample (for example, non-members resident with SLS members at Census time) please describe the samples separately in the box.

|  |
| --- |
|  |

**B1.2** *Estimated sample size (ask your Support Officer for help)*

Please include details of the estimated size of the sample and any subgroups that may be the focus of any analysis. Alternatively provide power calculations. Where possible, please include references for your estimates and show the calculations used to generate your sample size (ie NRS/SG or ISD/Public Health Scotland published tables/reports then estimating 5%, how many expected over time, etc).

If your project requires more than one sample (except for non-member samples) please give details in the box detailing it separately.

|  |
| --- |
|  |

**B1.3** *What tables will the data be drawn from?*

(See **pages2-4** for an overview of the SLS study structure and brief description of the data in each of the SLS files) Most data unless specified is from 1991 onwards. As the SLS is a dynamic dataset, data is updated on a cycle. Please speak to SLS staff to check which years are available and indicate in the Future Holdings box if you would like to use any updates to the SLS database once new data becomes available.

***SLS Core datasets***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Census data** | **Vital events data** |  | **Years**  | **Future holdings** |
| 1991 [ ]  | Births of SLS Member |[ ]  from       | to       | [ ]  |
| 2001 [ ]  | Births to SLS Member |[ ]  from       | to       |  [ ]  |
| 2011 [ ]  | Stillbirths |[ ]  from       | to       |  [ ]  |
|  | **Vital events data** |  | **Years** |  | **Future holdings** |
|  | Infant Mortality |[ ]  from       | to       | [ ]  |
|  | Marriages |[ ]  from       | to       | [ ]  |
|  | Widow(er)hoods | [ ]  | from       | to  | [ ]  |
|  | Death | [ ]  | from       | to       | [ ]  |
| **Other SLS data** |  |  |  |  |
| Immigration |  |[ ]  from  | to       |[ ]
| Emigration |  |[ ]  from       | to  |[ ]
| Annual Residential location data (created from GP histories, 2000 onwards) |[ ]  from       | to  |[ ]
| **Education data (2007 onwards)** |  |[ ]
| School Census |[ ] [ ]
| School Attendance | [ ]  School Exclusions [ ]  | [ ] (either/both) |
| SQA Attainment | [ ]  SQA Qualifications [ ]  | [ ]  (either/both) |
|  |  |  |
| **SLS 1936 Birth Cohort (SMS1947)** |  [ ]  |  |
|  |  |  |
| Ecological data (table ECOLOGICAL) |  [ ]  |  |
|  |  |
| ***Health data (variables on request only)*** |  |
|  *PBPP application attached* |[ ]   |
|  *PBPP application being prepared* |[ ]   |
|  |  |  |
| Hospital admissions SMR01 |[ ]   |
| Cancer registrations SMR06 |[ ]   |
| Mental Health Inpatient & Day Case SMR04 |[ ]   |
| Maternity Inpatient and Day Case SMR02 |[ ]   |
| Outpatient Attendance dataset SMR00 |[ ]   |
| Prescribing Information System (PIS) data |[ ]   |
| Child Health data |[ ]   |

## Other health data - please specify: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## B1.4 What research variables will you need?

Please list in more detail the variables you need

This section should be completed using the SLS Data Dictionary, which can be found at: <http://sls.lscs.ac.uk/variables/>. The easiest way is to create an account then click to select the variables and work in Excel on the variable selection (along with your Support Officer) until finalised. Please order by table and variable name before cutting and pasting the Excel file to create the table required below.

The tables in **page 2-5** at the start of this form (a summary of the most widely used variables and general information on selecting variables) may also be helpful. If you are not sure of the exact variables you want, give a broad indication of the type of variables you are likely to use and discuss this in more detail with your SLS Support Officer.

Full details will be required before data extraction can begin. Please note that a project with over 60 variables in total is considered a large project. It may also be useful to consider the restriction levels of variables (in [SLS data dictionary](https://sls.lscs.ac.uk/variables/#load) online) in order to operationalize them for your research. You can check this with your Support Officer or see **page 2** for an overview of the restricted variable levels.

If your project requires more than one sample (for example, non-members resident with SLS members at Census time) please list these variables in separate tables.

Justification for inclusion may include linkage, commonly accepted risk or cofactor, suspected risk or cofactor (provide some evidence), required for SLS produced derived variable in Box 1.3

***SLS members***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table name** | **Variable name** | **Short description** | **Restriction****level** | **Justification for Inclusion** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

***SLS non-members (i.e. people living with SLS members at Census time)***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table name** | **Variable name** | **Short description** | **Restriction****level** | **Justification for Inclusion** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**B1.5** *Derived variable(s)*

You may wish to specify your own derived variables (i.e. variables that do not exist already on the SLS database, which cannot readily be created from existing variables due to restriction levels). List the input variables to be used in the derivation and, where possible, show how the variable will be derived. You do ***not*** need to detail any variables you will create from the SLS research extract.

*As an example, if an SLS member is living with a partner (cohabiting/marriage) in both 1991 and 2001, we could use the full date of birth for the partner at each point in order to check whether it is the same person. This would create a 0/1 indicator: 0=different partner, 1=same - based on partner's DOB.*

|  |
| --- |
|  |

**B1.6** *Independently sourced data*

If you wish to link independently sourced data to your SLS sample, please describe the data (including source, such as SIMD) and the geographical level and time point (e.g. 1991 postcode, 2001 output area, 2011 datazone) at which you would like to link it below. Please also indicate the format that you will provide the data in.

|  |
| --- |
|  |

**B1.7** *Briefly explain the variable/data selection for this study (up to approx. 300 words)*

* Users of ISD health data: please explain in general why these variables are required for the study. If you have not completed a PBPP form please indicate the sources that may be requested.
* Users of weather and/or pollution data: please specify the frequency of (e.g. monthly, annual, day of Census) and geographical level at which you require the data.
* All users: please explain broadly why each of the SLS datasets selected is required.

(For example: if you have requested the Stillbirths, Infant Mortality data and have not discussed the need for this elsewhere in the application form, please state briefly why this information is necessary to answer your research question. Or if Marriages and Widow(er)hoods or Immigration/Emigration data requested explain briefly why it is essential)

|  |
| --- |
|       |

#### Section B2 Service and access arrangements

Currently, there are three possible methods for accessing SLS data, but all do require initial data prep work to be carried out within the NRS Safe Setting room (Ladywell House, Edinburgh)

1. Visiting the NRS Safe Setting room and working alongside a SLS Support Officer. We do not supply individual-level data for users to analyse themselves outside the safe setting. All researchers will start off with initial work on data management, deriving variables and exploratory data analysis at the safe setting.
2. Synthetic version of research data to work on remotely. Once the single analytic data set has been produced with one record per SLS member ie. variables have been derived, data extracts linked and the research cohort has been finalised etc, then SLS staff may create a synthetic version in order that work can be done away from the safe setting. For more info please see: <https://sls.lscs.ac.uk/guides-resources/synthetic-data/.> As code has been tested on the synthetic data, this will reduce the number of Safe Settings visits required to run this code on the real data.
3. Remote access will involve users submitting syntax files (command files). These will be run by SLS staff on the data. This service is offered only to more established researchers who require models (which do not have complicated SDC requirements etc).

Please discuss with your Support Officer what will work best for you during the lifecycle of your research project. Additionally, we are in the process of finalising permissions to allow SLS research extracts to be held in accredited Scottish National Safe Havens (ie Aberdeen, Glasgow or Dundee).

## What format do you want your extracts to be?

 SPSS [ ]  Stata [ ]

Please note we also have R and mapping software in the SLS-DSU Safe Setting, however **your initial SLS labelled extracts will only be provided in either** **SPSS or Stata**.

If required we can use package Stat Transfer to convert these extracts to other formats though loss of labelling may result. We will provide you with the syntax that we used to label the extract(s) (in SPSS or Stata) so that you can adapt this to re-label in your chosen package.

SLS-DSU staff offer support in SPSS or Stata (ie for running remote access code), but welcome R users to work on the SLS by using the Safe Settings. If you prefer another package (ie SAS, Matlab etc) you may be able to work on a ‘bring your own licence basis’, but please consult your Support Officer.

***Covid update:*** Due to changes in restrictions/lock downs we are now asking that all new SLS researchers prepare their SLS research data extracts to be synthesised, then SLS-DSU staff will create a synthetic version to enable working remotely from home. This will allow project research to continue if things shut down. Further, despite having 8 SLS-DSU Safe Setting PCs, owing to operating a low density safe working environment, presently we can only allow 2 of these to be used. Please work with SLS-DSU staff for support on producing a suitable dataset to be synthesised. The synthetic versions of your research extract will be provided in R, we can help convert back to Stata/SPSS using Stat Transfer – however, if variables were derived from part of your project data management phase then the labels will be missing, as such you will need to apply your own code to label them again.

#### Section B3 Application Checks

For office use only:

|  |  |  |
| --- | --- | --- |
|  | **Please comment or check box** | **Date** |
| **Checked by SLS SO**  |[ ]   |
| **Comments from SLS SO to SLS PM / SLS RB** |[ ]
| **Checked by SLS PM** |[ ]   |
| **Comments from SLS PM to SLS RB** |[ ]   |
| **Sent to SLS RB** |[ ]   |
| **Comments from SLS RB to research team** |[ ]
| **Application approved by SLS RB** |[ ]   |

1. The Core table contains info on the source from where the SLS member first entered the SLS (i.e. birth, migrant, Census). [↑](#footnote-ref-2)
2. THE NHSCR contains basic demographic details of everyone who was born or has died in Scotland plus anyone else who is (or has been) on the list of a general medical practitioner in Scotland. [↑](#footnote-ref-3)