**SLS Intermediate Outputs Clearance Form[[1]](#footnote-1)**

Project Number: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Researcher Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date of request: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Folder location and name: **Reports\To\_be\_cleared\**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(eg \*Reports\To\_be\_cleared\20xx\_xxx\_tbc(n)\_yymmdd*) (where 20xx\_xxx=project number, n=output number, yymmdd=date)

Date of SLS approval (SLS use only) \_\_\_\_\_\_\_\_\_\_\_\_\_\_

Approved by (Staff initials - SLS use only) \_\_\_\_\_\_\_\_\_\_\_\_\_\_

**I have read the SDC guidance v3.5 document ☐I have read the Intermediate Output Clearance Guidance and completed the Statistical Disclosure Control Checklist (below) ☐I confirm that the requested outputs fall within the scope of the project’s aims and objectives ☐**

Please complete the table below for each file to be released (insert additional rows if required)

|  |  |  |
| --- | --- | --- |
| File name | Description of file contents | Is this an update to a previously released file? If so, please provide details of changes |
|  |  |  |
|  |  |  |
|  |  |  |

**Intermediate Output Clearance GUIDANCE**

As far as possible, limit intermediate output requests to one of the following scenarios:

1. Pre-publication i.e. necessary to produce your final outputs
2. Discussions with team members, without which the analysis would be delayed. Team members should attend the safe setting where possible and researchers should make the majority of analysis decisions themselves

Check the following prior to requesting a file to be released from the SLS (see also the more detailed disclosure control checklists on p. 3-4):

* Titles and descriptions are clear and self-explanatory. Tables, charts[[2]](#footnote-2), variables and variable codes should be labelled in a meaningful way and have been formatted to a standard suitable for a final presentation i.e. not cut and pasted from log/output windows
* Outputs do not contain embedded data which could be made available after release
* There are no small, potentially-disclosive numbers - no tables should have a cell count of less than 10
* The sample used in each table/chart/model is shown i.e. description including exclusions and sample size (N). Any statistics should be based on a sample size of at least 10
* Differencing from previously released outputs: Have you produced similar analysis before that, combined with this output, could be used to identify someone?
* Every effort has been made to convert data results to descriptive text to avoid release of actual results. For example, a table of age in 1991 by age in 2001 would not be released due to small numbers but text could be written ‘*although in most cases age in 2001 is a corresponding number of years older than age in 1991, there are instances where this is not the case so a decision needs to be made to either exclude these cases or accept the 1991 age is being ‘correct’. Excluding such cases would reduce the sample size by <5%*’. This approach of summarising results also makes team discussions quicker and easier
* Any graph/chart/percentage must also include the data behind the graph/ chart/ percentage, have no outliers and be in a suitable format. Scatter plots are not released.

**Save this completed Intermediate Outputs checklist form with the outputs and email your SO to request clearance together with the location of the output**

**If output does not comply with the above, your SLS SO will cease output checking and you will need to return to the SLS Safe Setting to amend the output**

**Statistical Disclosure Control Checklist for SLS Intermediate Outputs**

Complete the following SDC checks: (Please answer all questions)

**Frequency tables/charts**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Yes | No | N/A |
| Are there any cells in the table with a value >0 and <10? | ☐ | ☐ | ☐ |
| Are there any columns or rows dominated by zeros or 100% of observations? | ☐ | ☐ | ☐ |
| Are there any cells with suppressed values/hidden columns or rows? | ☐ | ☐ | ☐ |
| Has the table used a different population base from previous similar tables? | ☐ | ☐ | ☐ |
| Has the table used a different variable breakdown from a previous similar table? | ☐ | ☐ | ☐ |
| Has the table used a different definition or source for a variable previously tabulated? | ☐ | ☐ | ☐ |
| Are there any minima/maxima present?  (for max/min separate frequency counts/univariate distribution table must be provided – ie for age could be disclosive for top age or certain ‘events’) | ☐ | ☐ | ☐ |

**Models**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Yes | No | N/A |
| Does the model have fewer than 10 residual degrees of freedom? | ☐ | ☐ | ☐ |
| Does the model description quote or plot any individual values, such as minimum or maximum values or outliers? Again, plots are discouraged and tables to be produced to produce away from the Safe Settings | ☐ | ☐ | ☐ |
| Does the model description include a residual plot or residual values? | ☐ | ☐ | ☐ |
| Has the model used a different population base (ie cohort) from a previously described (or cleared) model (or sets of models)? | ☐ | ☐ | ☐ |
| Is the regression undertaken on a single unit? | ☐ | ☐ | ☐ |
| Does the regression solely consist of categorical variables? | ☐ | ☐ | ☐ |

**Syntax files**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Yes | No | N/A |
| Is the code clearly annotated with comments to assist the reviewer? | ☐ | ☐ | ☐ |
| Are there any references or figures in the comments or code that could lead to potential identification of individuals? | ☐ | ☐ | ☐ |
| Are there any SLS numbers included in the code or the comments? | ☐ | ☐ | ☐ |
| Are there any counts from the data present in the comments or code? | ☐ | ☐ | ☐ |

* **Any white** boxes ticked, your output will fail SDC. If you think your output should still pass, please discuss this with your SLS support officer before leaving the safe setting
* **Any light grey** boxes ticked, your output may fail SDC. If your output fails, your support officer can also provide advice about how to re-design your outputs so they will pass
* **All dark grey** boxes ticked, your output is likely to pass SDC, but it still needs to be checked

The SLS support officer will compare your outputs to all outputs previously released from the same dataset to make sure it is not possible to identify any individuals or small cells by comparing outputs.

1. Please note this SLS form was adapted from the eDRIS doc: “*National Safe Haven Statistical Disclosure V3*” [↑](#footnote-ref-1)
2. Where possible please produce charts from tables away from the Safe Setting as charts should only be produced if the counts can be released [↑](#footnote-ref-2)