Requesting Outputs Form

Table of contents

1   Processing Time

Two members of ODAP data access team check all outputs before release. Please allow sufficient time to review and assess your outputs.

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

- Make sure your titles and descriptions are clear and self-explanatory

- Make sure your outputs do not contain embedded data which could be made available after release

- Check for small, potentially disclosive numbers

- Differencing from previously released outputs: Have you produced similar analysis before that combined with this output could be used to identify someone?

- Do you need it released? Do you want a draft output to be released so that you can discuss with your project team? You could consider reviewing it within the ODAP alongside any members of your project team that are named in your governance documentation. If you do need output released for internal use, please be aware this will be for sharing within the research team only.

On completion, email the ODAP data access team to request release of your files. Your output files that have been cleared for disclosure by the ODAP data management team will be sent to you via email.

2   Acknowledgements in publications:

The ODAP data access team ask that you acknowledge the use of ODAP data in any publications/presentations where appropriate. A copy of the ODAP publication policy will be attached in your output request.

3   ODAP Output Request Form

Study Number:

Applicant Name:

Date of Request:

Date of ODAP sign out *(ODAP use only)*:

Name and date of ODAP data manager sign out – if applicable *(ODAP use only)*:

3.1 Please confirm the following:

☐ I have completed the Disclosure Checklist (below)
☐ I confirm that the requested outputs fall within the scope of the project's aims and objectives

3.2 Please provide a brief description of study cohort including dates and geographical coverage of data:

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

File Name	File type (e.g. PDF, Word, Excel, R code etc.)	Which datasets is this output derived from? (e.g. SMR01, NRS Census)	Description of file contents	Is this an update to a previously released file? If so, please provide details of changes	Reason for release (e.g. internal/draft/publication)

4 Disclosure checklist for ODAP Outputs Requests

Complete the following checks (please answer all questions):

4.1 Frequency tables/charts

4.1.1 Are there any cells in the table with a value >0 and less than or equal to 5?

☐ Yes
☐ No

4.1.2 If there is any sensitive information or low-level geography, are there any cells in the table with a value >0 and less than or equal to 10?

☐ Yes
☐ No

4.1.3 Are there any columns or rows dominated by zeros or 100% of observations?

☐ Yes
☐ No

4.1.4   Are there any cells with suppressed values/hidden columns or rows?

☐ Yes
☐ No

4.1.5   Has the table used a different population base from previous similar tables?

☐ Yes
☐ No

4.1.6   Has the table used a different variable breakdown from a previous similar table?

☐ Yes
☐ No

4.1.7   Has the table used a different definition or source for a variable previously tabulated?

☐ Yes
☐ No

4.1.8   Are there any minima/maxima present?

☐ Yes
☐ No

4.2   Models

4.2.1   Does the model have fewer than 10 residual degrees of freedom?

☐ Yes
☐ No

4.2.2   Does the model description quote or plot any individual values, such as minimum or maximum values or outliers?

☐ Yes
☐ No

4.2.3   Does the model description include a residual plot or residual values?

☐ Yes
☐ No

4.2.4   Has the model used a different population base from a previously described model?

☐ Yes
☐ No

4.2.5   Is the regression undertaken on a single unit?

☐ Yes
☐ No

4.2.6   Does the regression solely consist of categorical variables?

☐ Yes
☐ No

4.3 Coding files (syntax)

4.3.1 Is the code clearly annotated with comments to assist the reviewer?

☐ Yes
☐ No

4.3.2 Are there any references or figures in the comments or code that could lead to potential identification of individuals?

☐ Yes
☐ No

4.3.3 Are there any pseudo anonymised ID numbers included in the code or the comments?

☐ Yes
☐ No

4.3.4 Has the volume of code needed been minimised?

☐ Yes
☐ No

4.3.5 Are there any counts from the data present in the comments or code?

☐ Yes

☐ No

- Any white boxes ticked; your output will fail Statistical Disclosure Check (SDC). If you think your output should still pass, please discuss this with an ODAP data manager.
- Any light grey boxes ticked; your output may fail SDC. If your output fails, an ODAP data manager 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 data manager 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.

Because of this, it is a good idea to think carefully about the population base and variable breakdowns you will need to use later on, before you release your first outputs.

5 Example scenarios:

5.1 Changing variable breakdowns

You release a table showing a variable against age and sex, in wide age bands (i.e. 16-39, 40-64, 65+). Later you decide to release some tables with different age bands (16-39, 40-59, 60+), to coincide with routine screening for a particular disease starting at age 60. The new tables can be compared against the old tables to produce a table of information about people in your sample aged 60-64. Your new tables will only pass SDC if they would pass for the 60-64 age group in their own right.

5.2 Excluding data points

You release some tables of results based on your whole sample of linked records. You later discover that a particular medical condition interacts with the effect you are investigating, so you decide to exclude all individuals with this medical condition from your analysis. Your new tables could now be compared against the original

tables you released, to reveal information about the people you excluded from your analysis. Your new tables will now only pass SDC if the equivalent table for the excluded group also passes in its own right. If the group you want to exclude is small, this will limit your chances of any tables passing SDC.

5.3   Adding data points

You release some data on all individuals registered as blind or partially sighted, and later decide to widen the scope to all individuals with any visual impairment in order to have a larger sample size for your model. Your new outputs could be used to produce results about all people with visual impairments who are not registered as blind or partially sighted. These outputs will also need to pass an SDC check.

5.4   Changing definitions

- You release some results on all individuals in your sample who said they were retired at a certain date. You later change your definitions to include only people who both said they were retired and were over the state pension age at that date. Any new outputs will reveal some information about people who retired under the state pension age, and this will also need to pass an SDC check.

- You release some data on all individuals in your sample who said they were retired at a certain date. You later change your definitions to use HMRC records showing whether or not an individual received a state pension. This would not necessarily create a problem for SDC, as it is possible to be either retired with no state pension or receiving a state pension while not retired.