|  |  |  |
| --- | --- | --- |
| **Cancer and Adult Screening**  **Team** |  | The Public Health Scotland logo |

STANDARD OPERATING PROCEDURE

Scottish Cancer QPIs Dashboard Update

Version 1.0

|  |  |  |  |
| --- | --- | --- | --- |
| **Version history** | **Date** | **Comments** | **Author(s)** |
| V0.1 | May 2023 | This is an unfinished draft | Angus Morton |
| V0.2 | Feb 2024 | Minor changes | Pauline Ward |
| V0.3 | Mar 2024 | Clarified filenames | Pauline Ward |
| V1.0 | April 2024 | Process for new QPIs | Pauline Ward |
| V1.1 | May 2024 | Updates on check\_submissions and note about QPI changes. | Pauline Ward |
| V1.1 | Aug 2024 | New script to build summary table. | Pauline Ward |

# Purpose:

The purpose of this SOP is to describe how to update the Scottish Cancer QPIs Dashboard.

# Overview:

|  |  |  |
| --- | --- | --- |
| **Topic** | **Description** | **Comment** |
| Analysis purpose: | Processing QPIs Data into the form in which Tableau requires it |  |
| Output: | hb\_hosp\_qpi.xlsx   * Data relating to the performance against each QPI for each location   Background\_Data\_Age\_Gender.xlsx   * Case numbers broken down by age and gender   Background\_Data\_Case.xlsx   * Case ascertainment figures broken down by board |  |
| Reporting: | SCRIS users |  |
| Regularity/Schedule: | Variable. Every time data is available for a new tumour group. |  |
| Datasets required: | Previous output files |  |
| Look-up files required: | Created as part of the process. See below |  |

# Background:

# Brief overview of procedure:

The cancer QPIs dashboard is updated using data provided by the three cancer networks. The cancer QPIs team in PHS needs to create a ‘lookup’ file which describes the nature of each QPI in each year for a specific cancer type. This lookup is then used to create excel templates for each of the networks to fill out. The networks are given a few weeks to pull out their data and send it back to us. Once we have the data for all three networks the scripts can be run to update the data files for the dashboard. The dashboard file itself can then be updated and migrated so that it shows up on the ‘live’ site.

# Useful information:

## Relevant folders:

Confidential area

[\\Isdsf00d03\quality\_indicators\Benchmarking\Cancer QPIs\Data](file:///\\Isdsf00d03\quality_indicators\Benchmarking\Cancer%20QPIs\Data)

## This SOP document is stored in:

\\Isdsf00d03\quality\_indicators\Benchmarking\Cancer QPIs\Data\new\_process

... and a copy of the latest version should also always be added to the /docs folder in the GitHub repository where the code is stored under version control:

<https://github.com/Public-Health-Scotland/qpi-dashboard>

# Procedure:

## Step 1: Set Up Folders

All files used by the R scripts will be read from/written to the ‘new\_process’ folder. In new\_process there is a folder called ‘blank\_folder\_structure’ which should be copied for each update and populated with the relevant files. The R scripts for updating the dashboard assume that the structure of the folders will remain the same each time. It is therefore important that folder and file names are kept identical between updates. Blank files with the appropriate names are in the ‘blank\_folder\_structure’ example and the paths can also be found/edited in the scripts. The most important file to put in at the start is the most recent HB\_Hosp\_QPI.xlsx file which can be found in tabvol or in the output folder from the most recent update prior to the current one. This file should be put in “/excels\_for\_tableau/initial\_run/input/”.

## Step 2: Create Lookup

The ‘lookup’ file is an excel file which contains general information relating to each QPI. This information includes the QPI name, descriptions of the data recorded in the numerator and denominator, and the percentage target for each QPI. PHS creates the lookup before creating the data collection templates to be sent out to the networks.

Each TSG has a lookup associated with it. For a given update there will already be a lookup for that TSG from the previous update however because the details of the QPIs change between years, updated information will need to be added to the lookup for the more recent years. If the previous lookup for the TSG can’t be found, or there have been changes to the old data in hb\_hosp\_qpi.xlsx, then the create\_lookup.R script can be run to generate a new lookup based on the data in HB\_Hosp\_QPI.xlsx.

In the case of completely new TSGs where there is no pre-existing lookup information, we have taken the simple approach of manually creating the lookup file, by taking information (QPI names, targets etc) from the HIS QPI definition document (“Clinical Quality Performance Indicators”), and copying the layout from another TSG’s lookup file.

It is often the case that there have been changes to the definitions of some of the QPIs from year to year. Manual editing of the lookup file is required in these cases.

The most useful resources for making manual edits to the lookup are the [WoSCAN](https://www.woscan.scot.nhs.uk/) regional reports for each TSG. If a QPI in the update differs from previous in definitions like numerator1 , denominator1, exclusions1 or target then edit the old QPI name (ie ‘QPI’ columns) to be QPI plus relevant years e.g., “QPI 7: Nephron Sparing Surgery: Years1:6” (for the years 2012 to 2017). Other changes to a name are that the QPI is archived, archiving means that the name is changed by prefixing with a Z and adding archived at end e.g. “ZQPI 6: Neo-adjuvant Radiotherapy (archived)”. This is necessary so that Tableau will treat the data from the different years as not being comparable, and will list them separately in drop-downs. If this splitting of the QPIs is not done, it can result in Tableau incorrectly putting the data on the same chart. In the case of a change to the target, Tableau will then display an inappropriately-averaged reference line.

## Step 3: Run Code

The code for the dashboard updates is kept in [this](https://github.com/Public-Health-Scotland/qpi-dashboard) GitHub repository. Clone this repository into somewhere you can access from posit. It doesn’t matter exactly where because all file paths are defined within the scripts.

These scripts pull information from three places to perform the update:

* The housekeeping.R file
* The current (‘live’) dashboard data files. The three required files are HB\_Hosp\_QPI.xlsx, background\_data\_case.xlsx and background\_data\_age\_gender.xlsx. These files need to be copied into the folder /excels\_for\_tableau/initial\_run/input/.
* The lookup.xlsx file, which needs to be copied into the folder /lookup/

### housekeeping.R

The top of this file contains the variables which will need to be edited with each new run. This script contains comments detailing each variable. At the start of each dashboard update edit these variables and save the file.

### create\_lookup.R

As mentioned above, this script creates an initial version of the lookup based on HB\_Hosp\_QPI.xlsx. Manual edits can be made to the resulting lookup if required. This script is not strictly necessary, as you could reuse the lookup from the last update of this TSG but in cases where edits have been made to older data in HB\_Hosp\_QPI.xlsx this script can be used to create a lookup which reflects those edits. Once created, this script places the file in the /lookup/ folder.

### create\_templates.R

This script writes the blank excel templates which are sent to the networks to be populated. The majority of the information this script needs comes from the lookup however a few pieces of info, such as the relevant hospitals for each TSG, come from the existing HB\_Hosp\_QPI.xlsx file and the housekeeping.R file. Once the templates have been created in this script they are saved in the /templates/ folder.

Once the networks have sent these templates back with the data filled in the completed submissions should be copied into the /data\_submissions/ folder.

### check\_submissions.R

Once the data submissions have been sent back by the networks, run this script to check that the totals match.

As you run the lines from the start of the script, they will call a function to carry out basic checks by counting (‘tallying’) the rows for different categories. The script creates a folder called quality\_checking/ alongside the lookup/ folder, and saves this output, in CSV format, as files with the word ‘tally’ in the filename.

The sum of the boards within a network should equal the network figure given in the submission. For surgical QPIs the hospital figures must also add up to the network total. This is checked for each of five numeric columns – numerator, denominator, and the not-recorded and exclusions columns. Any differences will be flagged up by this script. The two objects created at the end of the script (z\_board\_totals and z\_hospital\_totals) will be empty if all the numbers match. If some numbers don’t match then the rows in these objects will specify what data to check. The rows are written to a CSV file for convenience, also in the quality\_checking folder.

### hb\_hosp\_qpi.R

This script uses the data in the returned data submissions to update the HB\_Hosp\_QPI.xlsx file. Firstly the script pulls in the data submissions from the /data\_submissions/ folder and formats them. Then it creates the Scotland totals by adding up the network figures for each QPI. The lookup is then joined onto the data by matching the year, TSG and QPI name. Finally the additional variables are derived and the new HB\_Hosp\_QPI.xlsx is saved out to excels\_for\_tableau/initial\_run/output/.

Running hb\_hosp\_qpi.R also creates a file containing just the rows where Scotland is the location, so that these can be used for the summary table (“/for\_summary\_table/Scotland\_rows\_no\_comments.xlsx”).

### /for\_summary\_table/generate\_summary\_in\_excel.R

This script takes an Excel file containing the performance data for Scotland for the current year. This can be produced using the hb\_hosp script (see above) or by simply manually copying the Scotland rows from the tail-end of the newly-generated HB\_hosp file, and saving them into a file called “Scotland\_rows\_no\_comments.xlsx”, and then deleting out the comments. Then run generate\_sumary\_in\_excel.R to pivot the table, to put the years in the columns, and add colour highlighting, resulting in a table that is suitable for copying and pasting into the PHS Word template for Official Statistics Summary publications.

### age\_gender.R

This script pulls in the data from the backgroundInfo tabs of the data submission and uses it to create an updated version of the background\_data\_age\_gender.xlsx file. Note: there is an existing script which does this already which is what Garry uses.

### case\_asc.R

This script has two parts. The first creates a template for adding in the new case ascertainment data. This template needs to be copied from the /templates/ folder in the /excels\_for\_tableau/input/ folder. The second half of this script will then combine the new data with the existing data and export the updated file to excels\_for\_tableau/output/. Like the age\_gender.R script there is an existing version that Garry uses instead of this one.

After the above scripts are run you will have a set of Tableau files in excels\_for\_tableau/initial\_run/output/ which are ready to be copied over to tabvol to populate the dashboard. However, there are often additional edits required and the following scripts are used to make those.

### change\_qpi\_names.R

Sometimes QPI names need to be changed. The most common change required is when QPIs are archived and need to have ‘years x:y’ added to the end of their name. This script edits HB\_Hosp\_QPI.xlsx to update the names to the new version.

The name changes need to be defined in an excel file in /lookup/qpi\_name\_changes.xlsx. This file contains the old names to match on, the new names to change to and the range of years for which this QPI name needs changed.

This script takes its input files from excels\_for\_tableau/qpi\_name\_changes/input/ and puts the updated file in the corresponding output folder. This is to prevent accidental overwrites.

### change\_hospital\_names.R

Similar concept as the script above. A lookup of old and new hospital names is in the /lookup/ folder and the input and output HB\_Hosp\_QPI.xlsx is in /excels\_for\_tableau/hospital\_name\_changes/.

### late\_lookup\_edits.R

Often we receive feedback from the networks asking for edits to the descriptions of the QPIs (numerator1, denominator1, exclusions1). This script pulls in the lookup (with the QPI description edits) and the hb\_hosp\_qpi.xlsx and makes the required edits. The input and output HB\_Hosp\_QPI.xlsx files are in /excels\_for\_tableau/late\_lookup\_edits/

### Additional scripts

The scripts functions.R and packages.R are accessed indirectly i.e. through code in other scripts. You therefore shouldn’t need to open them unless updating (or fixing) the process.

## Step 4: Update Tableau

The below steps are the standard Tableau updates which are required every time. Sometimes other fixes/changes are required.

### Move files to tabvol

The excel files which actually populate the Tableau dashboards have to be kept in a specific folder called [tabvol](file:///\\tabstore\tabvol\SCRIS_Data\SCRIS_Level_1\QPIs). Copy all updated files from the QPIs folder into tabvol. Making sure to keep the names the same as this is a requirement for Tableau to recognise the files.

### Refresh data from source

For every file which has been updated in tabvol you need to tell the Tableau workbook that there has been an edit. The workbook does not automatically refresh when a file is edited. Refreshing the data in the dashboard from source (the source in this case being the excels from tabvol) can be done by selecting the following options:

A screenshot of a computer

Description automatically generated

### Update landing page

The below section of the landing page of the dashboard should be updated to display the most up-to-date information. It should describe the nature of the most recent update and the next update.

A screenshot of a message

Description automatically generated

### Set default values

The dropdown/tickbox values that are selected at the time of publication are what determines the default values when viewing the dashboard as a user. When dashboard users log in they usually want to see the new data so in all tabs select the TSG and years which have just been updated. Also select QPI 1 (or the lowest active QPI) as the default QPI.

### Publish to preprod

Once you are ready to publish the workbook file should be saved. Then select “Server > Publish Workbook”. Make sure the Project field reads “SCRIS Dashboard Level 2 – Analytics” and the Name field reads “SCRIS\_Level\_2\_Cancer\_QPIs” then click Publish. [Level 2 preprod](https://viz-pre-prod.nhsnss.scot.nhs.uk/#/site/NSS/workbooks/7842/views) will now be published.

Do not save the workbook file during or after publishing preprod level 1.

To take out the comments and publish preprod level 1 go into the “Cancer QPI – Reporting View” tab, open the “Bar chart alternative” sheet and remove the ‘AVG(Comment check)’ field from the Columns list by dragging it away:

A screenshot of a computer

Description automatically generated

Repeat this process in the “Cancer QPI – Presentation View” tab (the equivalent sheet in this tab is called “QPI Bar Chart”) and publish to preprod level 1 in the same way as in level 2 just changing any 2’s to 1’s.

Again, do not save these changes. Close the workbook without saving once preprod level 1 is published.

### Migrate to live

## Step 5: Pull Request GitHub Changes and Create Release

In the release notes give the paths to the data files used for this release so that it can be reproduced if required.

# Additional Information:

|  |  |  |
| --- | --- | --- |
| **Cancer Networks:** | | |
| NCA | North Cancer Alliance | Grampian, Highland, Orkney, Shetland, Tayside, Western Isles |
| SCAN | South East Scotland Cancer Network | Borders, Dumfries & Galloway, Fife, Lothian |
| WoSCAN | West of Scotland Cancer Network | Ayrshire & Arran, Forth Valley, Greater Glasgow & Clyde, Lanarkshire |
|  |  |  |