



# Data Protection Impact Assessment – Scottish Health Survey

Information Assurance and Risk



Scottish Government  
Riaghaltas na h-Alba

This template was developed by the SG Data Protection and Information Assets team.

This template was last updated in June 2018.

Before conducting the [Data Protection Impact Assessment](#), please refer to the guidance that accompanies this template.

## 1. Introduction

The purpose of this document is to report on and assess against any potential Privacy Impacts as a result of undertaking the Scottish Health Survey.

## 2. Document metadata

2.1 Name of Project: Scottish Health Survey

2.2 Author of report: Julie Landsberg, Health & Social Care Analysis

Date of report: Latest version, April 2023. First version April 2018.

2.3 Name of Information Asset Owner (IAO) of relevant business unit: Nicola Edge and Anita Morrison, Heads of Health & Social Care Analysis

2.4 Date for review of DPIA

### Review date

November 2019

### Details of update

Reviewed and updated for GDPR compliance

### Completion date

November 2019

### Approval Date

November 2019

### Review date

May 2021

### Details of update

Updated to reflect:

- telephone interviewing
- home working
- online self-completions
- the inclusion of Intake24
- development of the SHeS dashboard
- linkage with NHS data

### Completion date

May 2021

### Approval Date



May 2021

**Review date**

April 2023

**Details of update**

Updated to reflect:

- the switch back to face-to-face interviewing
- change in process for the child boost sample
- inclusion of Intake24

**Completion date**

April 2023

**Approval Date**

April 2023

**3. Description of the project**

3.1 Description of the work:

The Scottish Health Survey (SHeS) is a large-scale household survey commissioned by the Scottish Government Health Directorates to provide reliable information on the health, and factors related to health, of people living in Scotland.

Prior to 17 March 2020, interviewers interviewed respondents in their own homes via a computer assisted personal interviewing (CAPI) questionnaire. Some of the more sensitive questions were asked via paper or computer-assisted self-interviewing (CASI) self-completion. The survey also included physical measurements – height and weight, and for a sub-sample waist circumference, saliva sample (for measurement of cotinine) and blood pressure.

A urine sample was included until 2017 and may be reinstated at some stage. In 2021, the online dietary intake tool, Intake24, was included in the survey, with respondents asked to complete the tool after the telephone interview and on one other day.

On 17 March 2020, in-house interviewing was suspended for all SG surveys due to the COVID-19 pandemic. In August and September 2020, a shortened version of the survey was conducted by telephone. In 2021, interviews were also conducted by telephone but over a longer period between April 2021 and March 2022. Online versions of the self-completion elements of the survey were developed and worked well. Part way through the year, interviewers were able to start visiting sampled households to encourage response (an approach known as knock-to-nudge). This significantly increased response levels. The 2022 survey started with telephone interviewing with a knock-to-nudge approach and returned to the usual survey format of in-home interviewing from June, following approval from the Chief Medical Office that this could recommence. An opt-in process continued to be followed for the Child Boost until in-home interviewing using a sample linked to the Community Health Index to help identify addresses with children was introduced in September 2022.

The sample for the survey is drawn from the Royal Mail's Postcode Address file. Date of birth and names of all respondents are collected within the survey interview.

Respondents are asked if they would be willing to be contacted for the purpose of follow-up research. Survey responses are also linked to NHS health records unless the respondent advises the interviewer or contacts the contractor to say that they do not want to be included in the linkage.



Survey responses may be linked to other datasets subject to approval by the relevant governance and NHS research ethics groups.

## **Ownership**

Scottish Ministers are the data controllers of the survey data. Food Standards Scotland are joint data controllers of the data collected via Intake24. Survey data is shared with the wider research community through the UK Data Archive once appropriate disclosure control measures have been applied (see section 7 for details). The disclosure control methods pseudonymise the data and protect respondents' identities. Details of the disclosure controls used are published alongside the datasets.

The current contractor, the Scottish Centre for Social Research (Scotcen) (the Scottish arm of the National Centre for Social Research (NatCen)) and the Office for National Statistics (ONS), are data processors. The University of Cambridge are data processors (along with ScotCen) for the Intake24 data.

## **Governance**

Governance of the Scottish Health Survey is broadly carried out by the following:

1. Project and Contract Manager – manages the project on a day-to-day basis and directs longer term strategic work.
2. Health & Social Care Analysis Division – the Project Manager, C2 (Senior Statistician) and Division Heads (DDs) have overall responsibility for the survey.
3. Project Board – The Project Board includes senior policy users, representatives of key external users and the contractor. The Board is responsible for making or agreeing strategic decisions about the survey. This includes:
  - Discussing potential changes to the questionnaire and agreeing major changes such as new topics for inclusion or topics to be removed.
  - Agreeing the dissemination strategy for the survey.
  - Agreeing any major changes to the methodology of the survey or new modes of data collection.
  - Supporting the Project Manager and contractor in efforts to improve survey response.
  - Agreeing appropriate action in response to fieldwork performance issues (non-contractor members).
  - Making decisions in relation to the procurement or extension of the survey contract (non-contractor members).
  - Input into the content of publications.
  - Providing a link to users including health boards and local areas.

In addition, the Office of the Chief Statistician & Data Officer are involved in some survey decision-making around sampling, weighting and a small number of questions that are also included in the two other main Scottish Government household surveys (the Scottish Household Survey and the Scottish Crime and Justice Survey) and collated to form the Scottish Survey Core Questions.

## **Benefits of the project**

The Scottish Health Survey (SHeS) provides information about the health, and factors related to health, of people living in Scotland. The survey has been running continuously since 2008 (with some break in 2020 due to the Covid-19 pandemic) and prior to that ran in 1995, 2003 and 2008. The



survey provides information at national level annually and at Health Board and Local Authority level every four years.

The survey aims to:

- estimate the occurrence of particular health conditions
- estimate the prevalence of certain risk factors associated with health
- look at differences between regions and between subgroups of the population
- monitor trends in the population's health over time
- make a major contribution to monitoring progress towards health targets.

It provides robust data on a wide range of different topics, including smoking, alcohol consumption, drug use, obesity, physical activity, diet, general health, mental health, dental health, cardiovascular disease, respiratory conditions, caring, accidents/injuries and gambling. Much of the evidence provided by the survey cannot be obtained from other sources.

Data collected through the survey provides information to monitor a number of National Indicators within the National Performance Framework and a wide range of other targets, indicators and guidelines. The survey is used extensively across the Scottish Government, Health Boards, other health bodies, the research community and the third sector.

### **Planning mechanisms**

Planning is undertaken by the survey project manager and the survey Project Board (which has both internal and external members).

### **Reporting mechanisms**

The main release of SHeS results is through the annual report and a short infographic based summary report. In addition, the SHeS Dashboard provides data on a number of the survey variables by age, sex, deprivation, income, disability, health board and local authority. Some additional tables providing breakdowns for further variables are also published. These breakdowns will be brought in to the survey dashboard going forward. A technical report is published each year, outlining the survey methodology.

Reporting is undertaken by the survey contractor with input and quality assurance by the survey project manager and team within Scottish Government. The Scottish Government developed and update the SHeS dashboard and conduct a wide range of further analysis in response to internal and external user needs.

Indicators sourced from the SHeS are also published in a number of other web resources, including the following:

- National Performance Framework
- Equalities Evidence Finder
- Active Scotland Outcomes Framework
- Local Health Profiles (Scottish Public Health Observatory)

### **Risk management**



Managing risk, including risks in terms of a leak of personal data or the possibility of an individual being identified within data outputs, is essential to the successful operation of the project.

The measures set out in section 7 are monitored to minimise possible risk in terms of a leak of personal data and to ensure that personal data is handled in accordance with data protection legislation. Disclosure controls protecting individual anonymity are applied before data is placed in the UK Data Archive. Three Statistical Disclosure Control methods are applied to SHeS data: variable removal, top-coding and re-coding.

**Variable removal** is used when survey questions are deemed too sensitive to be included in microdata. This is also used in cases where a number of component survey questions are used to derive a summary variable for analysis, if the components are individually not recommended for analysis but might present some disclosure risk.

**Top-coding** is used on ordered variables for which values above or below a threshold value are grouped together, lowering the disclosure risk for individuals or households with uncommon characteristics.

**Re-coding** is used on variables to group responses into a smaller number of broader response categories, grouping together response categories which individually represent a small number of survey participants.

Data is available to researchers for analytical purposes upon their registration.

### 3.2 Personal data to be processed.

Variable	Data Source
Serial numbers	Randomly allocated by contractor
Age	Scottish Health Survey interview
Gender	Scottish Health Survey interview
Sex	Scottish Health Survey interview
Trans status	Scottish Health Survey interview
Education	Scottish Health Survey interview
Marital status	Scottish Health Survey interview
Household income level	Scottish Health Survey interview
Employment details	Scottish Health Survey interview
Sexual orientation	Scottish Health Survey interview
Ethnicity	Scottish Health Survey interview
Religion	Scottish Health Survey interview
Country of birth	Scottish Health Survey interview
Relationship to other household members	Scottish Health Survey interview
Physical and mental health	Scottish Health Survey interview
Suicide attempts and self-harm	Scottish Health Survey interview
Adverse Childhood Experience	Scottish Health Survey interview
Food insecurity	Scottish Health Survey interview
Gambling	Scottish Health Survey interview
Physical measurements – height, weight, waist circumference, blood pressure, saliva sample (and sometimes a urine	Scottish Health Survey, biological interview



sample). A blood sample may be introduced.	
Name	Scottish Health Survey interview (stored separately to survey answers)
Address	Postcode Address File (used for drawing the survey sample)
Postcode	Postcode Address File (used for drawing the survey sample)
Telephone number	Scottish Health Survey interview (stored separately to survey answers)
Email address	Scottish Health Survey interview (stored separately to survey answers)

### 3.3 Describe how this data will be processed:

#### **How it will be gathered**

The sample of the survey is drawn from the Royal Mail's Postcode Address File. The sample for the child boost is matched to the Community Health Index to identify households with a child resident. This improves fieldwork efficiency and reduces costs.

Personal details and answers to interview questions are provided by the interviewee to the survey interviewer either in person or, if the respondent requests, by telephone (only telephone interviewing was allowed between August 2020 and May 2022) who records within the CAPI questionnaire on their secure laptop computer. Response to some of the more sensitive questions are provided by the interviewee via online or, if the respondent prefers, paper self-completion. The paper self-completion questionnaires are scanned by a third party company, Stor-a-file) The Stor-a-file employees involved in this process will see the question answers and the respondent's first name only (which is collected solely to ensure the correct questionnaire is given to each respondent in the household). No other personal details of the respondent are collected and each form has a serial number to enable linking with other SHeS survey answers.

The online questionnaires are hosted on servers provided by Rackspace. A company called Artax have access to the servers for the purpose of providing software support only. There are Third party agreements between these companies and NatCen regarding data security. All data processing is done by NatCen on a secure network. The only personal information entered by the respondent in the online survey is their first name and the first name of the child (for parents' 4-12 questionnaire). The CAWI sample record contains an access code, but not the survey serial number.

In 2021, respondents were asked to complete an online dietary recall (Intake24) after the interview and on one other day. Intake24 is managed by the University of Cambridge. Intake24 does not collect personal details such as name or address, instead respondents were provided with an access code by NatCen. The University of Cambridge process the Intake24 data and return to NatCen who then link each respondents Intake24 responses to their responses to the rest of the SHeS via the





access code. In the event that participant contact details are passed to the University of Cambridge, consent is obtained and recorded as part of the initial online interview and is only for express purposes including telephone assistance for access issues or to complete the Intake24 recall by telephone. This information is stored separately to survey answers. Cambridge University staff do not have access to any other personal identifiable information other than is required for this specific purpose.

Biological measurements and saliva samples (and urine samples when included) are taken by specially trained interviewers, they are labelled using the serial number and the respondent's date of birth. These measurements and samples are special category biometric data.

### **Who will have access?**

Within SG, the survey manager and a small number of other analysts working on the survey within Health & Social Care Analysis Division have access to the pseudonymised data for the purpose of analysing the results. A small number of analysts within the Office of the Chief Statistician & Data Officer also have access to the pseudonymised survey results and the address of respondents for the purpose of allocating the sample, calculating the survey weights, fieldwork quality monitoring and for enhancing analysis through the Scottish Surveys Core Questions dataset. Access is granted by the SG survey manager. The access list is reviewed annually.

Food Standards Scotland (FSS) have access to the data from Intake24 as joint data controllers of this part of the data.

Within Natcen, those working in the logistics, statistics and programming departments and the research team as well as interviewers and fieldwork managers have access to address details of those included in the sample files. The data manager and programmer have access to all of the personal data, with the programmer being responsible for separating all personal data from the survey results (pseudonymisation), researchers have access only to the pseudonymised survey results. Interviewers will have access to the personal details, survey question answers that they collect verbally from respondents and height and weight measurements (when these are able to be taken) during the interview and prior to transmission.

For the purposes of field quality monitoring, the Natcen recall team have access to recall data files containing the name, address and telephone numbers of respondents which they use to enable them to telephone respondents to ask about the interview process as part of their quality assurance systems. They will not have access to any of the survey data. The recall data is collected in Natcen's secure Sample Management System for all projects, then around 35% is extracted for telephone recalls which the Senior Technical Data Director, Senior Survey Programmer and the telephone interviewers working on recalls have access to. Recalls data is securely deleted 15 months after the end of fieldwork for each project year.

Within ONS, those working in the Research Team, Survey Operations team (including interviewer managers and regional managers) and Interviewers will have access to address details of those included in the sample files. Interviewers will have access to the personal details, survey question answers that they collect verbally from respondents and height and weight measurements during the interview and prior to transmission. For the purposes of field quality monitoring, the ONS Research Team and Survey Operations Team will have access to recall data files containing the name, address and telephone numbers of respondents which they use to enable them to telephone respondents to





ask about the interview process as part of their quality assurance systems. They will not have access to any of the survey data.

If translation is required, on occasion, translators from third party companies are brought into interviews to translate for the interviewer and respondent.

When saliva or urine samples are included in the survey, those working in the Royal Victoria Infirmary laboratory in Newcastle will have access to the samples and ACM global laboratories in York will have access to saliva samples. These are labelled with the respondent serial number and date of birth.

Those working in the third party scanning company that Natcen employ to scan the paper self-complete questionnaires will have access to the paper self-completed questionnaires and electronic files that result from the scanning.

If a respondent has considerably raised blood pressure, the interviewer will provide the Survey Doctor with the respondent's name and telephone number to enable the Doctor to contact them. The survey program prompts interviewers when the respondent has considerably raised BP and in addition, a weekly report is sent highlighting cases of considerably raised and also low blood pressure. Interviewers can also contact the Survey Doctor regarding low blood pressure if they or the respondent are concerned. This is under the GDPR lawful basis of vital interests.

The printing company who print all of the survey documentation will have access to the sample address file and household serial numbers to enable them to merge sample household addresses onto the advance letters.

### **How it will be transmitted and how frequently**

Where personal data are transmitted outside Scotcen (to those involved in the research team in Natcen, ONS, FSS, third parties, the Scottish Government, PHS for data linkage and potentially for other future data linkage subject to Public Benefit and Privacy Panel approval) if transmitted electronically, all those working on SHeS follow strict guidelines, using a PGP encryption, protected using the 256-bit AES-encryption feature in WinZip or a secure FTP server. The FTP website offers the facility to transfer data securely over a FIPS 140-2 compliant SSL connection, without the need for third party organisations to install specialist software on their local PC. It has been verified by DigiCert and the system used for SHeS is Enhanced File Transfer (EFT) Globalscape. Personal data sent in hard copy form are sent by courier involving a tracking number and requiring a signature on receipt. The survey advance letters and leaflet explain how the respondent's data is handled and refer respondents to additional privacy information published on the survey website at [Scottish Health Survey: interviewee FAQs - gov.scot \(www.gov.scot\)](http://www.gov.scot). The interviewers' (both Scotcen and ONS interviewers) laptops contain the questionnaire responses prior to them being uploaded remotely to the office of the contractors Natcen, this happens throughout the data collection period as soon after the interview as is possible. The files are encrypted when transmitted. Those self-complete questionnaires completed on paper are sent to the Natcen office using SAEs as soon after the interview as possible. The data for the self-completions is securely transferred to NatCen and all processing is done at NatCen on a secure network.



The trained interviewer's label and despatch saliva and urine samples (when these are included in the survey) via Royal Mail to the Royal Victoria Infirmary laboratory in Newcastle (RVI). Saliva samples received at the RVI are assigned barcodes and dispatched bi-weekly in polythene bags (20 samples per bag) by secure courier for overnight delivery to ACM global laboratories in York.

Information sent between Scottish Government and the contractor is done so via secure FTP server. The contractor provides the pseudonymised survey datasets to the Scottish Government each year.

Sample files are transmitted from Natcen to ONS via an on-line secure transfer system.

When self-completion paper questionnaires arrive in the Natcen office, they are kept in locked cabinets until the start of the data capture process. The questionnaires are then batched, and sent via secure courier to whichever agency has been allocated the work. The data capture agency scan the questionnaires and return image files and data files to Natcen via a secure FTP site.

For the purposes of field quality monitoring, recall data files are transmitted from Natcen to ONS via an on-line secure transfer system.

The reporting of personal details (name and telephone number, interview time and serial number) of respondents who have considerably raised blood pressure by the interviewer to the Survey Doctor is very rare. This is always done verbally via the telephone.

The printing company who print all of the survey documentation will have access to the sample address file to enable them to merge sample household addresses onto the advance letters.

### **How it will be stored, and disposed of when no longer needed**

Within Scottish Government, the data is stored on the Government's secure datashare server with restricted access to named analysts within Health & Social Care Analysis Division and Office of the Chief Statistician & Data Officer.

Within Natcen, the Scottish Health Survey has its own specific security requirements and a Data Security plan detailing the security procedures to be followed. Details of all third parties working on a project, such as collaborators and subcontractors including printers and mailing houses, are recorded on the Data Security Plan. All third party individuals working on SHeS are asked to sign appropriate Third Party Information Security and disclosure agreements to ensure that they comply with data protection legislation. Additionally, Natcen have made site-visits to data scanning agencies to assure that any questionnaires that they do hold are stored in a secure environment.

Datasets are securely stored on Natcen's network with each project having a secure sub-folder for respondent confidential data which has restricted staff access. All paper documents containing data about a respondent are identifiable only by a serial number; such documents are returned in the post separately from any document(s) containing the respondent's name and address or other personal details. All interviewer laptops are protected with full disc encryption, to the FIPS 140-2 standard using PGP. Any information that could be used to identify individual respondents is stored separately, and once verified by the interviewer, is concealed should anyone else attempt to access the questionnaire file. CASI data (collected prior to the introduction of online self-completions in 2021) was also 'locked' after it was entered and could not be accessed by the interviewer. Following



transmission of SHeS interview data on their laptops or tablets, interviewers initiate the deletion of this data on their devices when they complete their monthly assignment. The deletion is normal deletion on the laptop, dod7 secure deletion is not required as the laptops and tablets have FIPS 140-2 standard full disc encryption.

Any information transmitted to ONS (for example sample files and recall information) is stored on an ONS secure network, with restricted staff access in place.

As a National Statistics product, all data files and materials relating to reporting outputs for SHeS are kept within secure folders with access limited to only those staff working directly on the study. All authors receive specific guidance on their responsibilities in relation to National Statistics, including the need to lock away or shred all draft text and tables. Report files are transferred using PGP or via the FTP secure server. Data sticks and unauthorised hardware are not permitted to be used in any NatCen computer (the USB ports have been disabled). SafeConsole USB storage devices, which are encrypted to the FIPS 140-2 standard, are used if files need to be transferred to a third party computer system). Level 3 (respondent confidential) data is never stored on an Iron Key. All our files, including data, are backed up daily on off-site servers.

The nature of SHeS as a National Statistic product means that Natcen adhere to the National Statistics principles and comply with the National Statistics Code of equal access, ensuring that only the immediate research team and report authors have access to the data before publication.

Names and addresses of informants are not linked to the data obtained. Strict guidelines are followed on separating Personal Identifiable Data from the survey data - before the survey dataset is delivered to the Data Manager personal identifiers such as name and address are removed. Individuals are identified on all stored computer data and written forms by means of a serial number only.

The hard copies of self-complete questionnaires are stored securely at the scanning agency until Natcen have checked the quality of the image files – if they are deemed acceptable, Natcen requests the agency to securely destroy, including on back up servers the electronic questionnaires. At the end of the project year, after the last batch has been sent to the agency and the last set of images received back and confirmed as up to standard by Natcen, the agency is then instructed by Natcen to destroy all hard copies. Natcen are then supplied by the agency with certificates of destruction stating that the work was carried out securely. When scanned self-complete questionnaire files are sent for data capture within Natcen, the respondent first names are not recorded in the datafile. The only place that the name is captured is on the image of the scanned questionnaire, which is deleted along with the other questionnaire data when the project is deleted as a whole.

The printing agency stores the sample address files securely and as soon as Natcen have confirmed they have received the advance letters and they are correct, they instruct the printing company to securely delete the sample files.

Saliva samples (and urine samples when included in the survey) are checked when they arrive at RVI for correct identification then assigned a barcoded laboratory accession number. Saliva samples are stored securely at 4°C before being batched and sent by secure courier overnight to ACM. Saliva samples are analysed on arrival at the ACM lab and are then stored until the end of the year of fieldwork in which they are collected. Urine samples are analysed on the day they arrive at RVI labs and then immediately destroyed.



Names, addresses and telephone numbers of respondents who have provided consent are held in a separate file to the survey answers by the survey contractor. These are used for follow-up research (respondents are asked if they are willing to be contacted for this purpose) and for data linkage (the SHeS data are linked to some NHS health record data). This linkage increases the value of the information collected as it allows research looking at the relationship between health behaviours and health outcomes/hospital stays. This retention is reviewed regularly to ensure that it is still appropriate. Any researchers can make applications to resample from this information, with endorsement from the relevant Scottish Government policy team. Applications for access to this data are assessed by the SHeS project manager and, if necessary, by the Scottish Government's Public Benefit and Privacy Panel.

Data collected via Intake24 is stored by the University of Cambridge. The IT infrastructure at the University of Cambridge that supports research across the MRC Epidemiology Unit is hosted in data centres located in various University of Cambridge buildings to mitigate risk. Physical and remote access to the server infrastructure is very tightly controlled and there are strict network and access controls in place around all aspects of the Unit's IT network and storage volumes. All Unit-managed desktops and laptops are encrypted using Microsoft BitLocker using the AES-128 method. Data backups are performed in a number of ways. The primary storage is automatically snapshotted regularly to disk providing the first layer of redundancy and offering user-driven file/data recovery. Key volumes are additionally automatically mirrored to a replica storage appliance elsewhere within the University for local disaster recovery purposes. Virtual servers containing NatCen data outside of the "file/folder" storage volumes are replicated daily to dedicated infrastructure hosted in a geographically separate data centre under University of Cambridge Clinical School control.

Data are transferred between NatCen and Cambridge University using the NatCen's secure file transfer service and storage location access is restricted to only those staff who need it for the project activity purposes. Where there is need to share participant personal data (e.g. to support Assisted Dietary recalls), such data is held on a Secure Research Drive (SRD) with enhanced and tightly controlled access arrangements. Staff provided with access are required to undergo clearances and access to the SRD is formally reviewed a minimum of three times a year. The University of Cambridge securely delete all SHeS Intake24 data they hold one year after the date of data archiving.

If the study were to end, and all data were to be transferred to the Scottish Government from the contractors, the Scottish Government would specify the retention and disposal requirements. Natcen have detailed data protection, retention and secure disposal procedures and facilities.

The pseudonymised data is held indefinitely by the SG for research and statistics purposes – i.e. to allow continued analysis across the full time series of SHeS.

The UK Data Archive and National Safe Haven hold data securely in perpetuity but this data has been fully disclosure checked and does not contain personal identifiers

### **Who owns and manages the data**

The Information Asset Owners are Nicola Edge and Anita Morrison, Heads of SG Health & Social Care Analysis Division. The survey manager is Julie Landsberg, Health & Social Care Analysis Division.



### How the data will be checked for accuracy and kept up to date

Scotcen undertake quality assurance checks on the SHeS data in accordance with the procedures in their Quality Management System (QMS), which outlines the minimum checks required for any survey data. Given the complexities of SHeS, additional checks are also run on the data. The checks are undertaken at an early stage in fieldwork, at the half way stage and when fieldwork is complete to ensure that any problems are identified and rectified as soon as possible. The majority of checks are run within SPSS. The computer-assisted personal interviewing (CAPI) questionnaire also has some validation checks built in to ensure that the data is as accurate as possible at the point of data entry.

The Scottish Government undertake additional summary checks on the data once it has been processed and a draft dataset has been provided by Scotcen.

The data does not need to be kept up to date. A new sample is drawn each year.

### 3.4 Explain the legal basis for the sharing with internal or external partners:

The Scottish Health Survey provides information to monitor how well Scottish Ministers are fulfilling their duty under the National Health Service (Scotland) Act 1978 to promote the improvement of the physical and mental health of people in Scotland.

The information gathered through this survey contributes to:

- Monitoring the population's health.
- Monitoring progress made towards health targets.

See section 6 for further details.

## 4. Stakeholder analysis and consultation

### 4.1 List all the groups involved in the project, and state their interest.

**Group** - Members of the public

#### **Interest**

Members of the population are contacted and asked if they will answer the survey questions. Currently around 40% of those contacted take part in the survey when in-home interviewing is allowed. Response for telephone interviewing is lower at around 15%.

**Group** - Scottish Government, Health & Social Care Analysis, Population Health team

#### **Interest**

The SG SHeS team consists of three statisticians who quality assure and analyse the data, including the SG survey/contract manager. This team handles all requests for analysis, special data sets and access to the follow-up sample, referring higher risk requests to the SG Public Benefit and Privacy Panel.



The Heads of Health and Social Care Analysis Division are data controllers of the survey data and sign off key decisions and this DPIA.

**Group** - Scottish Government, Office for the Chief Statistician & Data Officer

**Interest**

Conduct the sampling and weighting of the survey and co-ordinate the SG data access panel.

**Group** - Scottish Centre for Social Research (Scotcen) (the Scottish arm of Natcen)

**Interest**

The contractor is the data processor and conducts the survey on behalf of the Scottish Government.

**Group** - Office for National Statistics

**Interest**

Sub-contractor and data processor.

**Group** - Food Standards Scotland

**Interest**

Joint data controller of the data from Intake24.

**Group** - University of Cambridge

**Interest**

Sub-contractor responsible for running the online dietary intake tool, Intake24.

**Group** - Public Health Scotland (PHS)

**Interest**

Involved in the linkage of the survey data with health record data.

**Group** - UK Data Service

**Interest**

Stores pseudonymised datasets.

4.2 Method used to consult with these groups when making the DPIA.

The survey contractors have been fully involved in this assessment. The Office of the Chief Statistician, ONS, the University of Cambridge and PHS have contributed to relevant sections.

4.3 Method used to communicate the outcomes of the DPIA.

This Data Protection Impact Assessment will be published on the SHeS survey website. A Scotstat email will be sent to alert stakeholders that a new version has been published.





## 5. Questions to identify privacy issues

### 5.1 Involvement of multiple organisations

The Scottish Government commission the survey.

Scotcen run the survey on behalf of the Scottish Government and process the data. Scotcen is ISO ISO27001 compliant and does not transfer data out-with the UK.

The Office for National Statistics (ONS) are sub-contracted (by Scotcen) to undertake around a third of the survey interviews when the survey can take place within households, the rest are conducted by Scotcen.

The University of Cambridge run the online dietary intake tool Intake24 which respondents were asked to complete after the SHeS interview in the 2021 survey. The university processed the Intake24 data and then transferred to ScotCen for inclusion in the SHeS data sets.

Public Health Scotland link the SHeS data to health record data. Scotcen transfer the contact details of participants to PHS via secure file transfer. The FTP website used offers the facility to transfer data securely over a FIPS 140-2 compliant SSL connection, without the need for third party organisations to install specialist software on their local PC. It has been verified by DigiCert and the system used is Enhanced File Transfer (EFT) Globalscape.

### 5.2 Anonymity and pseudonymity

A random serial number is assigned to each respondent to the Scottish Health Survey by the contractor. The survey datasets include the survey answers alongside this serial number. The contractor retains (on behalf of SG) the personal identifiers of respondents (name, address and date of birth) in a separate file which is also referenced by this serial number.

A similar approach is taken for the linkage of the health survey data to health records. The personal identifiers of respondents together with a different randomly assigned serial number created by the contractor are sent securely to PHS. No survey answers are included alongside the personal identifiers. PHS then use the personal identifiers to match to Community Health Index numbers in order to pull off the relevant health records of respondents. A separate DPIA has been undertaken specifically in relation to this data linkage.

Any other linkage agreed by the relevant Public Benefit and Privacy panel/s would involve indexing by National Record of Scotland and linking to the population spine.

The samples sent to the labs by interviewers are labelled by the interviewer with the respondent's serial number and date of birth and sent with a dispatch containing the serial number, sex, date of birth and smoking status. These details are entered onto the labs secure computer system and the samples are assigned a unique barcode. The use of this minimal personal information is necessary to ensure that the samples are correctly assigned to respondents. No lab results are recorded alongside personal identifiers, the serial number is used to link respondent's lab results with their other survey results.

### 5.3 Technology





Within Scottish Government, personal data is held electronically on restricted areas of the Scottish Government secure server as described in Section 3.3. Section 3.3 also describes the Scotcen procedures. Each new or additional information technology is assessed for privacy intrusion before it is used in the project.

The Scottish Government survey team only uses technology cleared by Scottish Government IT experts.

#### 5.4 Identification methods

Each respondent to SHeS is identified within the survey datasets by a unique serial number. An additional identifier is used to identify each household.

#### 5.5 Sensitive/Special Category personal data

Sexual orientation

Trans status

Ethnicity

Religion

Physical and mental health

Suicide attempts and self-harm

Physical measurements – height, weight, waist circumference, blood pressure, saliva sample, urine sample (a urine sample has not been included since the 2017 survey).

#### 5.6 Changes to data handling procedures

Names, addresses and telephone numbers of those respondents who agree to be contacted for the purpose of follow-up research will only be used for this purpose for a maximum of five years following the survey year. So respondents interviewed in 2018, for example, could be contacted up until the end of 2023. In practice, only respondents from the most recent year or two of published data will be used unless the research is restricted to a relatively small group of the population requiring more years to provide a big enough sample.

A new process is in place going forward for linking the survey data with health record data. The SG Public Benefit and Privacy Panel and the Health & Social Care Public Benefit & Privacy Panels have agreed this process.

The handling of any other data linkage agreed by the relevant Public Benefit and Privacy panel/s would be agreed with the panel/s.

#### 5.7 Statutory exemptions/protection

GDPR Article 89(1) exemption for statistical purposes applies subject to appropriate safeguards such as pseudonymisation.

#### 5.8 Justification



The information collected in the survey provides reliable information to the government and others on the health, and factors related to health, of people living in Scotland. This information is used to inform policy, monitor changes over time and assess health inequalities

#### 5.9 Other risks

N/A

## 6. General Data Protection Regulation (GDPR) Principles

### 6.1 Principle 1 – fair and lawful, and meeting the conditions for processing

Compliant – Yes

Processing is necessary for compliance with a legal obligation - Article 6 (1) (c) UK GDPR.  
In the case of the sensitive data collected, Article 9 (2) (j) applies – “processing is necessary for ..... statistical purposes in accordance with Article 89(1) ....”.

The survey advance letters and leaflet explain how the respondent’s data is handled and refer respondents to additional privacy information published on the survey website.

Respondents are asked to agree for their personal details (name, address and date of birth) to be used by the Scottish Government or passed on to research agencies for the purpose of follow-up research endorsed by the government.

the respondent’s data is handled and refer respondents to additional privacy information published on the survey website.

Respondents are asked to agree for their personal details (name, address and date of birth) to be used by the Scottish Government or passed on to research agencies for the purpose of follow-up research endorsed by the government.

### 6.2 Principle 2 – purpose limitation

Compliant – Yes

GDPR Article 89(1) exemption for statistical purposes applies – see section 5.7.

### 6.3 Principle 3 – adequacy, relevance and data minimisation

Compliant – Yes



The content of the survey is regularly reviewed to ensure that there is a continued need for the data collected.

#### 6.4 Principle 4 – accurate, kept up to date, deletion

Compliant – Yes

The information is obtained directly from respondents. The computer-assisted personal interviewing (CAPI) questionnaire includes some built-in logic checks and further quality assurance checks are performed by the contractor.

The data does not require to be kept up-to-date as it is representative of the survey year in which the respondent was interviewed. A new sample of respondents is drawn each year.

#### 6.5 Principle 5 – kept for no longer than necessary, anonymization

Compliant – Yes

The pseudonymised survey datasets are held indefinitely by the Scottish Government to enable analysis looking at changes over time. The contractor holds the survey datasets for as long as it holds the survey contract, after which the data will be permanently deleted.

The personal identifiers of respondents are held in a separate file by the contractor (on behalf of the Scottish Government) for the purpose of follow-up research and data linkage. These files will be transferred to the Scottish Government when the contract ceases (or at any time at the Scottish Government's request). When the contract ceases, following transfer to the Scottish Government, the contractor will permanently delete the personal identifiers.

#### 6.6 GDPR Articles 12-22 – data subject rights

Compliant – Yes

Respondents are advised in the survey leaflet that the survey is not compulsory and that they do not have to answer all the questions.

#### 6.7 Principle 6 – security

Compliant – Yes

Within Scottish Government, the survey data is held on a restricted access area of the government's secure server. Data is transferred from Scotcen to the Scottish Government via password controlled secure file transfer.

Within Scotcen, survey data has its own specific data security plan. Datasets are securely stored on Natcen's network with a secure sub-folder for respondent confidential data which has restricted staff access. All paper documents containing data about a respondent are identifiable only by a serial number; such documents are returned in the post separately from any document(s) containing the respondent's name and address or other personal details. All interviewer laptops are protected with full disc encryption, to the FIPS 140-2 standard using PGP. Any information that could be used to



identify individual respondents is stored separately, and once verified by the interviewer, is concealed should anyone else attempt to access the questionnaire file. CASI data is also 'locked' after it is entered and cannot be accessed by the interviewer. As a National Statistics product, all data files and materials relating to reporting outputs for SHeS are kept within secure folders with access limited to only those staff working directly on the study. All report authors receive specific guidance on their responsibilities in relation to National Statistics, including the need to lock away or shred all draft text and tables. Report files are transferred using PGP or via the FTP secure server. Data sticks and unauthorised hardware are not permitted to be used in any NatCen computer (the USB ports have been disabled). Where personal data are transmitted outside Scotcen electronically, strict guidelines are followed, using a PGP encryption, protected using the 256-bit AES-encryption feature in WinZip or a secure FTP server. The FTP website offers the facility to transfer data securely over a FIPS 140-2 compliant SSL connection, without the need for third party organisations to install specialist software on their local PC. It has been verified by DigiCert and the system we use is Enhanced File Transfer (EFT) Globalscape. Respondent confidential data is never stored on an Iron Key. All files, including data, are backed up daily on off-site servers.

The IT infrastructure at the University of Cambridge that supports research across the MRC Epidemiology Unit is hosted in data centres located in various University of Cambridge buildings to mitigate risk. Physical and remote access to the server infrastructure is very tightly controlled and there are strict network and access controls in place around all aspects of the Unit's IT network and storage volumes. All Unit-managed desktops and laptops are encrypted using Microsoft BitLocker using the AES-128 method. Data backups are performed in a number of ways. The primary storage is automatically snapshotted regularly to disk providing the first layer of redundancy and offering user-driven file/data recovery. Key volumes are additionally automatically mirrored to a replica storage appliance elsewhere within the University for local disaster recovery purposes. Virtual servers containing NatCen data outside of the "file/folder" storage volumes are replicated daily to dedicated infrastructure hosted in a geographically separate data centre under University of Cambridge Clinical School control.

As a research institution, MRC Epidemiology Unit servers, including servers for Intake24, are centrally maintained for security and updates. Critical updates are implemented in an advertised window \*at least\* once every four weeks, sometimes more frequently as appropriate. Disruption is extremely minimal (seconds/minutes if anything) and it is very rare that problems would be detected. Server and data backups are run routinely for business continuity purposes, including encryption and rotating of all disks or tapes created for business continuity purposes and stored in secure locations. Access to backups is strictly controlled. The University of Cambridge has a detailed and regularly reviewed Business Continuity Plan.

6.8 GDPR Article 24 - Personal data shall not be transferred to a country or territory outside the European Economic Area.

Compliant – Yes

The survey personal data is not shared with any organisation that operates or sub-contracts operations outside the EEA

## 7. Risks identified and appropriate solutions or mitigation actions proposed



Is the risk eliminated, reduced or accepted?

**Risk** - Personal information about an individual is lost/leaked during fieldwork.

### **Solution or mitigation**

Interviewers sign confidentiality agreements before they start work and receive information security training when joining NatCen and ONS and annual refresher courses.

Scotcen has ISO27001 information security accreditation. Environment and procedures are externally audited twice a year as part of maintaining accreditation and assuring compliance.

Any third parties contracted by Scotcen, such as document fulfilment or mailing companies and labs also have to conform to Scotcen's data security standards and the best practice defined in ISO27001. Third party individuals such as the Survey Doctor have to sign a Third Party Information Security agreement and a non-disclosure agreement stating that they will conform to Natcen data security standards.

Interviews conducted by the ONS are collected using Scotcen encrypted laptops and uploaded directly to Scotcen systems.

For Intake24, all respondent confidential data is saved in secure locations only. Access is controlled via a data security plan which specifies who has been granted access to the data, with access being revoked when no longer needed. All staff who access data are required to sign a confidentiality agreement. Data is restored from secure backup

Saliva samples and urine samples (when included in the survey) are labelled only with respondent date of birth and serial number; they are also accompanied by a dispatch note that contains the above information plus the sex of the respondent. This ensures that the samples could not be linked to any individual if they were to be lost. The samples are sent by interviewers via business class (1<sup>st</sup>) Royal Mail and in accordance with the posting regulations for hazardous samples. Saliva samples are also then batched by the RVI laboratory and sent to another laboratory (ABS) via secure courier for processing.

Result - Risk reduced

**Risk** - Personal information about an individual is accidentally leaked/released during or after processing.

### **Solution or mitigation**

Personal identifiers (names, addresses and dates of birth) are stored securely and separately to the survey answers of respondents. Access to personal identifiers within Scotcen is limited to all interviewers, field managers, the programmer, the data manager, the immediate research team, logistics and statistics and field monitoring staff. Access to personal identifiers within ONS is limited to the research team, survey operations team (including interviewer managers and regional managers) and all interviewers. Other third parties that have access to personal identifiers include printing and data capture agencies (scanning) and the Survey Doctor.



The survey datasets are pseudonymised and stored securely on the ScotCen server with access limited to the Scotcen research team, data manager, programmer and Natcen and external chapter authors (e.g. academics). The datasets are transferred to the Scottish Government via secure file transfer and stored on the Scottish Government's secure server with access limited to the survey team.

All Scottish Government and Scotcen personnel with access to the survey datasets are trained on data protection requirements at least annually and are clear on the processes for protecting personal information, including sensitive personal information.

Result - Risk reduced

**Risk** - A person is identified from the pseudonymised survey datasets or survey analysis.

### **Solution or mitigation**

Statistical disclosure control procedures are performed on the data before the dataset is made available to users via the UK Data Service. These controls are in line with those used by the other two large SG household surveys (the Scottish Household Survey and the Scottish Crime and Justice Survey) and have been approved by the Office of the Chief Statistician & Data Officer.

Special Dataset requests are assessed by the SG survey team in the first instance and scored against the SG data access risk matrix. Any request with a score of over 12 are referred to the SG Public Benefit and Privacy Panel.

Result - Risk reduced

## **8. Incorporating Privacy Risks into planning**

Explain how the risks and solutions or mitigation actions will be incorporated into the project/business plan, and how they will be monitored. There must be a named official responsible for addressing and monitoring each risk.

**Risk** - Personal information about an individual is lost/leaked during fieldwork

### **How risk will be incorporated into planning**

Interviewers must report any data loss during fieldwork as information security incidents to Natcens Head of Project Management and to the Scotcen survey lead as soon as they become aware of it. ONS interviewers will report any data loss to the Head of ONS Social Survey Operations, who in turn will notify NatCens Head of Project Management and Survey lead team. All third parties are also required to report actual or potential data security breaches to the Scotcen survey lead. The Scotcen survey lead will then raise all of the above formally as a critical incident within Natcen and report this to the Scottish Government survey manager as soon as possible. Procedures will be reviewed formally as part of the critical incident process to minimise any future risk.

Owner - Nicola Edge/Anita Morrison//Julie Landsberg



**Risk** - Personal information about an individual is accidentally leaked/released during or after processing.

**How risk will be incorporated into planning**

The Scotcen survey lead will report any leak of personal data to the Scottish Government survey manager as soon as possible and raise this formally as a critical incident within Natcen. Procedures will be reviewed formally as part of the critical incident process to minimise any future risk.

Owner - Nicola Edge/Anita Morrison//Julie Landsberg

**Risk** - A person is identified from the pseudonymised survey datasets or survey analysis.

**How risk will be incorporated into planning**

Should the survey manager be advised of any individual being identified from the survey datasets or analysis, the survey disclosure control procedures will be amended to minimise the possibility of any future risk.

Owner - Nicola Edge/Anita Morrison//Julie Landsberg

**9. Data Protection Officer (DPO)**

9.1 The DPO may give additional advice, please indicate how this has been actioned.

**Advice from DPO** - Add further information about the linkage of the survey results with health record data.

Action - Further details added.

**Advice from DPO** - Clarify the position in relation to third party organisations involved.

Action - Discussed with ScotCen and clarified references to third parties.

**Advice from DPO** - Confirm that there is a privacy notice for the survey.

Action - Referred to the survey privacy notice in section 3.3 and provided a web link.

**10. Authorisation and publication**

The DPIA report should be signed by your Information Asset Owner (IAO). The IAO will be the Deputy Director or Head of Division.





Before signing the DPIA report, an IAO should ensure that she/he is satisfied that the impact assessment is robust, has addressed all the relevant issues and that appropriate actions have been taken.

By signing the DPIA report, the IAO is confirming that the impact of applying the policy has been sufficiently assessed against the individuals' right to privacy.

The results of the impact assessment must be published in the eRDM with the phrase "DPIA report" and the name of the project or initiative in the title.

Details of any relevant information asset must be added to the Information Asset Register, with a note that a DPIA has been conducted.

**I confirm that the impact of (undertaking the project/applying the policy – add appropriate wording) has been sufficiently assessed against the needs of the privacy duty:**

Name and job title of a IAO or equivalent	Date each version authorised
Angela Campbell, Head of Health & Social Care Analytical Services	<b>April 2018</b>
Nicola Edge/Anita Morrison, Heads of Health & Social Care Analytical Service	<b>November 2019</b>
Nicola Edge/Anita Morrison, Heads of Health & Social Care Analytical Service	<b>May 2021</b>
Nicola Edge/Anita Morrison, Heads of Health & Social Care Analytical Service	<b>April 2023</b>





© Crown copyright 2023



This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit [nationalarchives.gov.uk/doc/open-government-licence/version/3](https://nationalarchives.gov.uk/doc/open-government-licence/version/3) or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk).

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

This publication is available at [www.gov.scot](http://www.gov.scot)

Any enquiries regarding this publication should be sent to us at

The Scottish Government  
St Andrew's House  
Edinburgh  
EH1 3DG

ISBN: 978-1-80525-945-9 (web only)

Published by The Scottish Government, June 2023

Produced for The Scottish Government by APS Group Scotland, 21 Tennant Street, Edinburgh EH6 5NA  
PPDAS1304322 (06/23)

W W W . g o v . s c o t