

Data Protection Impact Assessment (DPIA)  
Ref: NRS-DPIA-2020-01  
Census Quality Assurance, Estimation and Adjustment – NHS Central  
Register (NHSCR) Dataset  
Version 2.2

## Document Control

Title	Quality Assurance for Census – NHS Central Register (NHSCR) Dataset
Prepared By	Head of Administrative Data
Approved By	Director of Statistical Services
Date of Approval	02/09/2022
Version Number	2.2
Review Frequency	Annual
Next Review Date	September 2023

## Document Owner

This document is owned by

Name	Role	Team / Branch	Directorate
Caroline Ellis	Head of Admin Data	Admin Data	Statistical Services

## Document Control

<b>Title</b>	Quality Assurance for Census – NHSCR Dataset
<b>Prepared by</b>	NRS: Head of Admin Data
<b>Approved by</b>	NRS: Head of Census, Statistics & Registration
<b>Date of approval</b>	As required or annually
<b>Review frequency</b>	As required or annually
<b>Next review date</b>	September 2023 – annual from date approved

## Status Control

<b>Version</b>	<b>Date</b>	<b>Status</b>	<b>Reason for Amendment</b>
0.1	12/07/2019	Draft	New project
0.2	27/04/2020	Draft	Updated after advice from Scottish Government Legal Department
0.3	26/06/2020	Draft	Updated after advice from Security
0.4	02/07/2020	Draft	Updated after advice from Information Risk team and Information Governance team
1.0	01/09/2020	Final	Reviewed by DPO and IAO
1.1	01/09/2020	Published	Creation of pdf for website
1.2	10/03/2022	Draft	Updated to reflect changes to census processing and census DPIA
2.1	29/08/2022	Draft	Updated to add Estimation and Adjustment to Census Processing
2.2	01/09/2022	Final	DPO advice provided.

## Part 1: Data protection impact assessment screening questions

These questions are intended to help you decide whether a DPIA is necessary. Answering 'yes' to any of these questions is an indication that a DPIA would be a useful exercise. You can expand on your answers as the project develops if you need to. You can adapt these questions to align more closely to project you are assessing.

### 1. Will the project involve the collection of new information about individuals?

No. The project re-uses administrative data collected by National Records of Scotland (NRS) and other public bodies for research and statistical purposes only.

### 2. Will the project compel individuals to provide information about themselves?

No.

### 3. Will information about individuals be disclosed to organisations or people who have not previously had routine access to the information?

Information will be provided to authorised NRS staff in the administrative data team to complete this project.

For the estimation and adjustment strand, an excerpt from the NHS Central Register (NHSCR) will be stored in the safe haven. An eDRIS research coordinator will be responsible for quality assuring any outputs.

**4. Are you using information about individuals for a purpose it is not currently used for, or in a way it is not currently used?**

Yes , we propose to use NHSCR data for a purpose it is not currently used for. Data from NHSCR for this project is split into two strands:

**Census Quality Assurance**

Data will be used to quality assure Scotland's Census Rehearsal 2019 and the Scotland's Census 2022 at individual level. The data will be used by NRS for statistics and research purposes only.

**Census Estimation and Adjustment**

Under the Census Estimation and Adjustment strand, NRS propose to use Administrative data obtained from the NHS Central Register (NHSCR) as part of the statistical processes for estimation (to estimate missing records) and adjustment (to impute skeleton records in the place of non-responding households).

For more information on legislation involved please refer to **Step 4: Necessity and Proportionality**.

**5. Does the project involve matching data or combining datasets from different sources?**

Yes.

**Census Quality Assurance strand**

The project involves matching Census data (both from Rehearsal for testing purposes and the main 2022 Census) with a defined extract of data from the statutory NHS Central Register (NHSCR).

**Census estimation and adjustment strand**

NRS proposes to link NHSCR Data to other administrative data sources, 2022 Census and 2022 census coverage survey.

**6. Does the project involve you using new technology that might be perceived as being privacy intrusive?**

No.

**7. Will the project result in you making decisions or taking action against individuals in ways that can have a significant impact on them? Will you profile individuals on a large scale?**

NRS will not use the data collected to make decisions or take action against individuals.

Location data from individuals on NHSCR will be used to quality assure individual census returns and to improve efficacy of our statistical linking process.

**8. Will you profile children or target marketing or online services at them?**

No. No contact will be made with individuals based on this project. Children will not be targeted for marketing or online services. Location data from NHSCR is used to quality assure individual census returns and to improve efficacy of our statistical linking.

**9. Is the information about individuals of a kind particularly likely to raise privacy concerns or expectations? For example, special category data such as health records or criminal records, or other information that people would consider to be private.**

**Census Quality Assurance strand**

No special category data is included. The use of the NHSCR data will be limited to identifying inconsistencies in the Census data set, rather than being merged with the Census data.

**Census Estimation and Adjustment strand**

No special category data is being requested from NHSCR. NHSCR may be linked with other data sources which contain special category data as part of this project.

**10. Will the project require you to contact individuals in ways that they may find intrusive?**

No. No contact will be made with individuals.

**11. Is the project collecting personal data from a source other than the individual without providing them with a privacy notice ('invisible processing')**

Yes, the project is collecting personal data from administrative sources and is not collecting personal information from the individual. The Census 2022 Privacy Notice will be amended to explain the use of Administrative datasets in the processing of the census records prior to any such processing occurring.

**12. Is the project tracking individuals' location or behavior?**

No behaviour information is included. Location information down to postcode level is included for a single point in time, so cannot track individuals.

**NRS maintains a record of answers to the screening questions in order to document that the decision on whether to carry out a DPIA was properly considered. If after completing the screening questions you decided a DPIA is not necessary you must send a record your answers to the [NRS Data Protection mailbox](#). The NRS Information Governance Team will review answers, and where appropriate ask the NRS Privacy Group for their opinion.**

**Decision of Information Governance Team**

**DPIA Required:** Yes  
**Reason for decision:**  
**Census Quality Assurance Strand**  
Any extension of use of the NHSCR data must be for a fair, lawful and for a necessary purpose. Although the data is being shared within the Registrar General's responsibility as Data Controller and is only being used for QA purposes, a DPIA is necessary because the project involves data linking for quality assurance purposes. A full DPIA will ensure that any privacy risks are fully addressed.

**Census Estimation and Adjustment**

Due to the low response rate NRS has proposed using Administrative data sources to complete the missing records by working out which records are missing and then imputing data from similar records for those missing individuals. A DPIA is necessary as the project involves data linking for Census estimation and adjustment purposes. The DPIA will ensure adherence to Data Protection by Design and Default and address Privacy Risks.

<b>Name:</b> NRS Head of Information Governance	<b>Date:</b> QA - 11/05/2020
<b>Name:</b> NRS Head of Information Governance	<b>Date:</b> Census Estimation and Adjustment – 06/06/22



## Part 2: Data protection impact assessment report

Use this report template to record the DPIA process and results. You can start to fill in details from the beginning of the project, after the screening questions have identified the need for a DPIA. The template follows the process that is used in the ICO code of practice. You can adapt the template to allow you to record additional information relevant to the DPIA you are conducting.

### Step 1: Describe the project and identify the need for a DPIA

Explain what the project aims to achieve, what the benefits will be to NRS, to individuals and to other parties. You may find it helpful to link to other relevant documents related to the project, for example a project proposal or business case.

It is important to include information about the benefits to be gained from the project in order to help balance any risk identified in the DPIA. This can help inform decisions on the level of risk to privacy that is acceptable, when balanced against the benefits or other justification for the project. Is there a benefit to the public? If a statutory duty exists provide details of this. Also summarise why the need for a DPIA was identified (this can draw on your answers to the screening questions) and identify the legal basis for processing.

The Registrar General (RG) decided that the 2022 Census programme would be conducted as an accurate enumeration based census aligned to the current decennial cycle. Additionally, this programme would also look at the future use of administrative data collected by public bodies and services to augment or replace NRS data collected by a traditional census. This approach was accepted by Scottish Ministers in March 2014<sup>1</sup>.

One of the Admin Data team main objectives is:

- To enable quality assurance and to improve the accuracy of the 2022 Census estimates

#### Census Quality Assurance Strand

Through the Admin Data project (eDRIS 1617-0195) we determined that there were only two datasets that provided large enough coverage to quality assure census 2022; namely the NHSCR dataset and the Health Activity dataset. Of these two datasets, only the NHSCR would be available within the time frame required to allow quality assurance in the short time period required.

Research approval was received by the Public Benefit and Privacy Panel for Health and Social Care (PBPP), to use this dataset to quality assure the census.

The project identified that the NHSCR dataset would be able to provide individual level quality assurance to improve the following statistical areas:

1. Stop removal of genuine persons due to incomplete census return.
2. Supply age or Date of Birth at census where date of birth is incomplete or missing for imputation purposes.
3. Highlight date of births on census that may be require further clerical checking due to age at census not matching NHSCR e.g. errors caused by transposition of figures, a person who is in full time education but year of birth has been keyed as 2001 instead of 2010.
4. An independent check to support the de-duplication of persons who have been entered twice.
5. Create a census coverage survey for the rehearsal from NHSCR dataset (approx. 3,000 records) to support testing of statistical methodology for census to census coverage survey data linkage for the 2022 Census.

The main benefits of using the NHSCR dataset are that it is an independent source from the census and the timeliness of the dataset. It can identify areas in census that may benefit from more in-depth clerical checking, provide assurance that the age characteristic is as expected and help create more accurate population estimates calculated from census returns. There were only two datasets from the Admin Data Team project (eDRIS 1617-0195) that had the range of population required for this task, the NHSCR dataset and the Health Activity dataset. When timelines were examined the Health Activity dataset was discounted as the processing time needed by the data supplier would be too long for this project. The NHSCR dataset could be delivered within a couple of weeks of the 2022 Census statistical dataset being created.

## **Additional information**

Any requests for changes to the data collection and methodology of NHSCR data are submitted to the Public Benefit and Privacy Panel for Health Panel and Social Care (PBPP) for approval. This Panel operates to fulfil three main aims:

- To provide a single, consistent, open and transparent scrutiny process allowing health and social care data to be used for a range of purposes including research
- To ensure the right balance is struck between safeguarding the privacy of all people in Scotland and the fiduciary duty of Scottish public bodies to make the best possible use of the health and social care data collected – it is important to note that each is in the public interest
- To provide leadership across a range of complex privacy and information governance issues, so that the people of Scotland are able to gain the benefits – ultimately better health and social care – from research and wider use of data, while ensuring compliance with legal privacy obligations, managing emerging information risks, addressing public concern around privacy, and promoting the protection of privacy as in the public interest.<sup>2</sup>

## **Census Estimation and Adjustment Strand**

Scotland's Census 2022 achieved lower than expected response rates. In light of this development, National Records of Scotland propose to introduce administrative data from a range of sources, including the NHSCR, to improve our statistical processes used to estimate the Scottish Population, including those who did not complete a census return, at a national and local level. This project is set out in eDRIS project 2223-0075.

Estimation refers to the process of calculating where our census dataset does not represent the whole population, then adjustment is the process of imputing skeleton records in the place of non-responding households. The intention of these processes is to ensure our population estimates are accurate at national (estimation) and local (adjustment) levels. A range of outputs, reports, and analysis will then be produced based on this census data.

We propose to link several administrative datasets together. The NHSCR dataset will only be used to remove, from the linked dataset, individuals who the NHSCR indicates are not in the Scotland population. The NHSCR has information for individuals who previously lived in Scotland but are now registered with a GP elsewhere in the UK, or have left the UK, or have died. Some

of these individuals may still appear on other administrative datasets. Such individuals need to be excluded to ensure that bias is not introduced to the census estimates. **NHSCR data itself will not leave the safe haven.**

At the beginning of this project, NRS considered the range of available data from across the public sector. NHSCR dataset was selected for inclusion because: it has coverage of a broad proportion of the population, and can identify individuals who are no longer part of Scotland's population. The census estimation process is designed to estimate the number of individuals who did not respond to the census, in order to estimate Scotland's total population. As more people than expected did not respond to the census, the accuracy of this estimation will be affected. Considering administrative data records alongside census responses during estimation can help achieve the target accuracy. However, it is essential to ensure that there are no records for individuals who are not in the Scotland population, as the estimation process assumes that every record represents someone in the population. The NHSCR will allow records for individuals who have died or have left Scotland to be removed from the list, before it is compared with the census responses. As the NHSCR covers almost everyone in the Scotland population, anyone who appears on other datasets but not NHSCR will also be excluded. This means that the NHSCR records for individuals who are still in Scotland are also required. The NHSCR will also be available in time to meet the milestones of this project.

Once the datasets have been linked, a minimised cut of variables (not including any information from NHSCR) will be removed from the safe haven and used in our estimation and adjustment statistical processes. This would be done by supplementing the number of people who appear on the census or the CCS with the number of extra people identified by the combined administrative dataset. NRS also employ a number of follow up procedures to fine tune population estimates, reducing bias. These procedures all involve using a secondary data source to identify and fix issues. The administrative data would be used to enhance these procedures, either improving the quality of the comparator data or acting as a comparator data set in its own right.

The aim of this project is to use administrative data to ensure the **accuracy and quality of the Scotland's Census 2022 is as high as possible.** The objectives are:

- **To use administrative data to support estimation** – the statistical process of identifying who is missing from our original census returns. This allows us to estimate the overall size of the Scottish population.

---

<sup>2</sup> <https://www.informationgovernance.scot.nhs.uk/pbpphsc/>

- **To use administrative data to support adjustment** – the statistical process of imputing skeleton records where data is missing. This allows our population estimates to be accurate to lower levels.
- **To harness administrative data** that include many people who are typically hard to count through census field collection, and make significant quality improvement over previous census methods
- **The data will only be used for statistical purposes** – i.e. it will not be used to contact people who have not filled in census returns.

Producing accurate estimates of Scotland's population will feed into the wider benefits that census provides (explained in more detail below). This includes providing UK Government, Scottish Government, local governments, health boards, and other service providers with the information they need to make decisions.

#### **Benefits of estimation and adjustment**

- The objective of this proposal is to use administrative data in the census statistical processes to **ensure the 2022 Census population estimates are as accurate as possible at a national and local level.**
- Census population estimates are used to allocate funding across the United Kingdom, through the Barnett formula. The Scottish Government in turn uses census data to allocate funding to local authorities, health boards and other public services. For funding to be allocated correctly, is essential that population estimates are as accurate as possible. **Therefore this project has a direct benefit to all people across the United Kingdom.**
- The population estimates from Scotland's Census are used in a multitude of other ways – to plan policies, deliver services and conduct research by national and international organisations. More detail of these benefits is set out below.
- The method proposed will allow us to produce accurate outputs within the time periods census users need

### **Global benefits of accurate population estimates**

Basic information on the population size, age, sex and location is fundamental to the work of Government and other service providers to allow them to:

- shape public policy and inform investment decisions, including fiscal policy and strategies to stimulate economic growth
- model future needs for services and infrastructure
- identify housing demand and create housing supply including information on household size and family make-up which are crucial to policies on local housing demand and planning, and poor housing and overcrowding
- identify areas of deprivation, enabling them to target services <sup>3</sup>

Previous research shows that people in deprived areas are less likely to respond to the census<sup>4</sup>. The use of admin data will specifically improve return rates in those parts of the country where return rates were the lowest, with a particular focus on supporting Scotland's poorest communities. Without this, there is potential for gaps in the understanding of issues such as deprivation.

### **Benefit of accurate population estimates to Scotland's funding**

The Barnett formula used to allocate funding across the United Kingdom will be based on census data for the next decade. If population estimates are inaccurate, Scotland's funding could be inaccurate over the long term. For every 16,000 people by which the census underestimates the Scotland population (0.3%), the Scottish Government could lose around £1 billion from the block grant over the next decade<sup>5</sup>.

---

<sup>3</sup> [GROS - Census - General Report Template.dot \(scotlandscensus.gov.uk\)](https://www.scotlandscensus.gov.uk/gros-census-general-report-template)

<sup>4</sup> [Scotland's Census 2022 Fairer Scotland Duty Impact Assessment \(Dec 20\) \(scotlandscensus.gov.uk\)](https://www.scotlandscensus.gov.uk/scotland-s-census-2022-fairer-scotland-duty-impact-assessment-dec-20)

<sup>5</sup> Detailed calculations can be provided on request. [Block Grant Transparency 2021 Explanatory Note .pdf \(publishing.service.gov.uk\)](https://www.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/94844/block-grant-transparency-2021-explanatory-note.pdf)

Census outputs were used to determine Health Board funding allocations, as part of the National Resource Allocation Formula (NRAC)<sup>6</sup>. Local authority funding is set by the Scottish Government, taking into account census population estimates for the area<sup>7</sup>.

### **Benefit of accurate census population estimates to health services**

The outputs from Scotland's Census 2011 were used in a range of health services within Scotland, including:

- Determining Health Board funding allocations, as part of the National Resource Allocation Formula (NRAC).
- Updating the Acute part of the Morbidity and Life Circumstances component of the Resource Allocation Formula
- Predicting the demands on health services
- Provide update rates of patients accessing various forms of treatment
- Information on carers under 18, and is the main source of information on informal caring, so gave information on unmet needs.<sup>8</sup>

NRS calculated how much health board funding allocations would have differed if the census had not been carried out in 2011. If census figures from 2001 had been used to make population estimates and allocate funding to health boards, then in 2014/15 there would have been misallocations of between £30m and £40m. Some health boards would have received more, some less, than their appropriate share<sup>9</sup>.

### **Benefits of having census population estimates accurate at a smaller level**

Low-level population and household estimates are used to plan services at local level, including schools, libraries, parks.

Furthermore, given the particular focus on supporting Scotland's poorest communities, accurate data zone level estimates would

---

<sup>6</sup> [Scotland's Census 2011 General Report](#) – Section 8.127

<sup>7</sup> [Early Learning & Childcare - Scottish Local Government Finance 'Green Book' 2022-23 - gov.scot \(www.gov.scot\)](#)

<sup>8</sup> [Scotland's Census 2011 General Report](#) – Section 8.127

<sup>9</sup> [Scotland's Census 2011 General Report](#)

be needed. Without this, there is potential for gaps in the understanding of issues such as deprivation, leading to missed opportunities to improve the lives of the most disadvantaged in society.

NRS has published detailed use and benefits of the 2011 Census Data in its General Report.<sup>10</sup>

## **Step 2: Describe the processing**

**Describe the nature of the processing:** how will you collect, use, store and delete data? What is the source of the data? Will you be sharing data with anyone? You might find it useful to refer to a flow diagram or other way of describing data flows. What types of processing identified as likely high risk are involved?

---

<sup>10</sup> [Scotland's Census 2011 General Report | Scotland's Census \(scotlandscensus.gov.uk\)](http://scotlandscensus.gov.uk)



## **How will the information be collected?**

### **All Strands (Census QA, Estimation and Adjustment )**

The project involves using data gathered by the administrative systems from the data provider. No new data will be collected as part of the project. However, the project will use administrative data and data collected primarily as part of Scotland's Census 2022 – this includes data from the 2019 Census Rehearsal, the 2022 Census and the 2022 Census Coverage Survey (CCS). The data will be delivered using a secure transfer mechanism as per data sharing agreement and then stored in a secure area that only the Admin Data team in NRS have access to. These individuals have to be named on the NHSCR data sharing agreement.

### **Census Quality Assurance Strand**

The NHSCR dataset will be collected for two specific snapshots: end of March 2019 and end of March 2022. Data required from NHSCR is explained in Annex A.

### **Census Estimation and Adjustment Strand**

The NHSCR dataset needed will only be at one snapshot: the end of March 2022. The data provider will provide personal data - names, dates of birth, sex and postcodes and additional data such as posting date and date of death.

Data required from NHSCR and justification as to why is explained in Annex A.

## **How will the data be collected and transmitted?**

### **Census Quality Assurance Strand**

The Admin Data team will need to link the NHSCR dataset to census. This will be identifiable linking of data on names, dates of birth, sex and postcodes.

The process to link the data is:

1. Data providers (NHSCR) send their data to the NRS Head of Admin data.
2. Head of Admin Data, will store in a secure IT location with only limited access by authorised personnel.
3. The Head of Admin Data transfers the dataset into secure location, (request is approved by senior manager).
4. A reduced output from census is produced. This output is used to check the following issues:
  - a. Stop removal of genuine persons due to incomplete census return.
  - b. Supply age or Date of Birth at census where date of birth is incomplete or missing for imputation purposes only.
  - c. Highlight date of births on census that may require further clerical checking due to age at census not matching NHSCR by providing age e.g. errors caused by transposition of figures
  - d. An independent check to support with the de-duplication of persons who have been entered twice.
5. The census output is linked to NHSCR. NHSCR data and outputs from this linking will remain in the Admin Area. Any clerical review (manual checking of records) is done by Admin Data staff as per the data sharing agreement.
6. A reduced output is returned to census to allow statistical changes to take place, these outputs would have all or some of the following fields:
  7. Census primary ID
  8. Pseudo ID - NHSCR
  9. Binary output confirmation if positive link has been discovered for partial records in census
  10. Date of Birth or age (only were DOB is missing, incomplete or different)
  11. Time stamp for when data is linked
  12. Clerical review timestamp and resolution
  13. Match key and strength
  14. Version control of dataset

15. Clerical review outcome

16. Name of Admin Data team member who performed clerical review

17. Data processing would use this information to create statistical outputs using edit and imputation purposes e.g. Donor Imputation where DOB was missing<sup>11</sup>, etc.

18. These outputs are then passed to Statistical Disclosure Control and Outputs who perform statistical disclosure control on the outputs before they are released to the public.

19. The original NHSCR data will be removed from the Admin Data secure area as per deletion and retention policy.

The data flow for the QA strand has been illustrated below in Figure 1:

---

<sup>11</sup> More information on Edit and Imputation especially Donor Imputation can be found in slides 52-74 on NRS website [https://www.scotlandscensus.gov.uk/documents/Statistical\\_Methods\\_Presentation.pdf](https://www.scotlandscensus.gov.uk/documents/Statistical_Methods_Presentation.pdf)

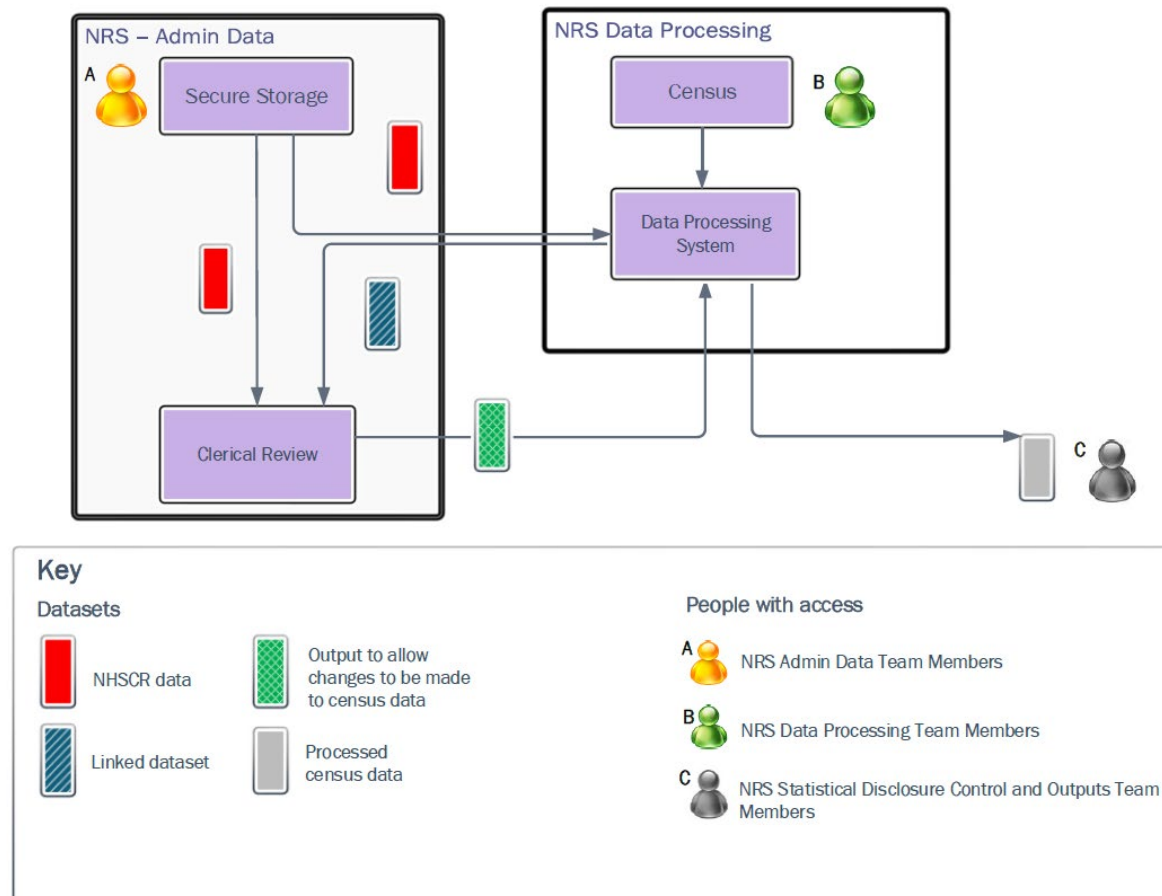


Figure 1: Data Flows for QA Strand of the Project

This would be an iterative process due to the Census 2022 being primarily digital and data would be continually being received for a period of time.

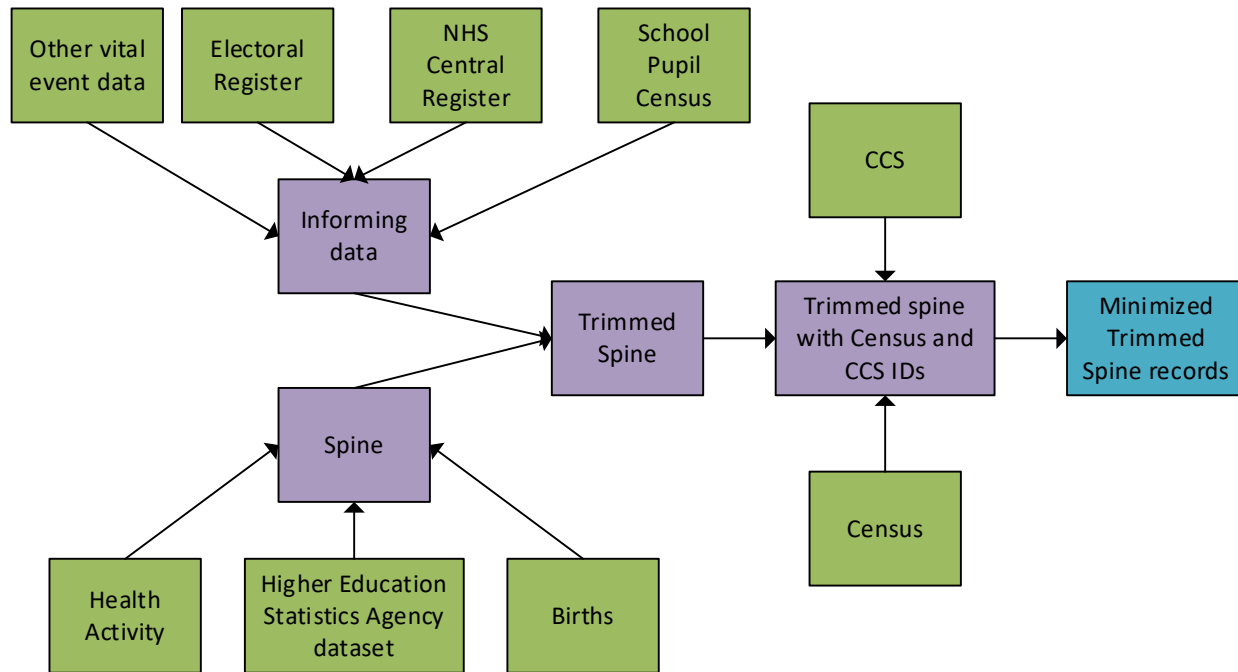
For 2019, the NHSCR was used to create a census coverage survey (approx. 3000 records) to test linking methodology to a reduced census rehearsal 2019 data. A full clerical review on this was undertaken by Admin Data Team only. A research paper on the findings was created to inform the linking methodology for the Census 2022 and Census Coverage Survey 2022 and the methodology was submitted to internal and external Peer Reviews and the Census Design Authority. The methodology papers were published on the census website ([link](#)).

### **Census Estimation and Adjustment Strand**

The data flow for the project is as follows:

1. Data is securely transferred from the provider to NRS Head of Admin data as per the data sharing agreement.
2. The data as provided is stored in NRS secure IT systems. Data is encrypted at rest and can only be accessed by limited named members of staff.
3. The data is quality assured and standardised by one named member of the admin data team. Quality assurance is important to allow any issues to be immediately flagged to data providers. Standardisation is important because it prepares the dataset to be linked with others. This team member can only access one dataset at a time.
4. Quality assured and standardised data is securely transferred to the National Safe Haven, following the safe haven's policies for transfer.
5. Datasets are linked in the Safe Haven. Access is limited to named members of the admin data team (a subset of staff named in Section 1.5) who are responsible for linking, and reviewing link quality. NHSCR is used in this process as an "Informing dataset". This means it's used to improve the quality of linking, but will not be brought out of the Safe Haven for use in estimation and adjustment
6. Information from other admin sources, **not including any data from NHSCR**, is then brought out of the safe haven for use in estimation and adjustment
7. The original NHSCR data will be removed from the Admin Data secure area and safe haven as per deletion and retention policy.

The datasets involved for linking in the safe haven are illustrated below in Figure 2:



**How will personal information collected be stored, and disposed of when no longer needed?**

**All strands -**

Personal information will be received and stored in the secure Admin Data area at NRS for a limited period of time (as stated within the data sharing agreement). This information will be limited to NRS Admin Data team only.

At the end of the project, the data will be deleted from the drives in accordance with Centre for the Protection of National Infrastructure (CPNI) destruction standards.

### **Census Estimation and Adjustment Strand:**

Additionally personal data will be stored in the National Safe Haven as per Figure 2.

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

#### **All strands**

On receipt of data NRS performed a Quality Assurance of Administration Data (QAAD) to assess the suitability of NHSCR as an independent quality assurance source to measure census. Two datasets were assessed, 2019 and 2022. The results are published on our [website](#). Once we receive data for the Estimation and Adjustment strand we will update the QAAD if necessary.

**Describe the scope of the processing:** what is the nature of the data, and does it include special category or criminal offence data? How much data will you be collecting and using? How often? How long will you keep it? How many individuals are affected? What geographical area does it cover?

#### **Nature of the data**

For NHSCR datasets, the data provider will supply personal information – names, dates of birth, sex, and postcodes. How we process this has been explained in the previous section.

#### **Data Collection**

#### **Census Quality and Assurance Strand**

The data collected will be two datasets, obtained from NHSCR, at the end of March for 2019 and 2022. These will hopefully be able to represent every person living in Scotland in order to use administrative data to quality assure the census. The NHSCR will cover the whole of Scotland but no direct geographical outputs will be produced from data, as it is supporting the quality assurance of census and the geographical outputs will be derived from the census statistical dataset.

#### **Census Estimation and Adjustment Strand**

The estimation and adjustment strand will only use the March 2022 dataset. This will be a one-off request (unless errors needing to be corrected are found). The dataset will cover the whole of Scotland – all individuals who are registered within the NHSCR are

requested. This will hopefully be able to represent every person living in Scotland in order to use administrative data to accurately identify the population who did not respond to the census. This will be used to create accurate national and local level population estimates and adjustments to the missing records.

No information from NHSCR will be taken out of the Safe Haven, instead it is used to select records for which there is sufficient confidence that they represent an individual in the population that they could be included in the census.

### **Retention of data**

#### **All strands**

The data held at NRS will be retained as per each individual data sharing agreement and shall be destroyed securely in accordance with CPNI destruction standards.

#### **Census Estimation and Adjustment Strand**

NHSCR Data will also be held in the National Safe Haven. The National Safe haven implements a period of secure retention after the end of the project, typically 5 years, in line with edris policies in their role as research coordinator.

Data stored in NRS will be securely deleted in accordance with HMG Infosec Standard 5 at the enhanced level and the National Cyber Security Centre (NCSC) guidance on secure sanitisation (<https://www.ncsc.gov.uk/guidance/secure-sanitisation-storage-media>). NRS will provide written confirmation to all providers when the data has been deleted.



**Describe the context of the processing:** what is the nature of your relationship with the individuals? How much control will they have? Would they expect you to use their data in this way? Do they include children or other vulnerable groups? Are there prior concerns over this type of processing or security flaws? Is it novel in any way? What is the current state of technology in this area? Are there any current issues of public concern that you should factor in? Are you signed up to any approved code of conduct or certification scheme (once any have been approved)?

NRS will not be contacting individuals on the basis of NHSCR data.

A limited amount of data will be extracted from NHSCR to cover the whole population of Scotland at a snapshot in time. Given the dataset includes all members of the population this would include children and vulnerable groups.

All staff within the Admin Data team have completed internal mandatory training courses on Data Protection and Information Governance. The Admin Data team have further completed the Medical Research Council online course on Information Governance, GDPR and Confidentiality. All members of the team are part of the Government Statistical Service and as such are bound by the Code of Practice For Statistics. Additionally all staff within the Admin team have signed the Census Confidentiality undertaking (this is to confirm that they understand the legal obligations of receiving census data and are aware of the penalties for unlawful disclosure or unlawful use of the information, or failure to keep it safe).

NHSCR data is governed by the Local Electoral Administration and Registration Services (Scotland) Act 2006. This needs to be taken into account when requesting this data even though NRS are the data owners of the NHSCR and Census datasets.

#### **Census Quality Assurance Strand**

The use of data in this project has also been taken to the Public Benefit and Privacy Panel for Health and Social Care (PBPP) and has been approved for QA purposes. This group looks at the use of health data within research/data linkage projects.

#### **Census Estimation and Adjustment**

The Registrar General for Scotland, has a statutory duty under section [2\(1\)](#) of the [Census Act 1920](#) to “make such arrangements and do all such things as are necessary for the taking of a census”. [Section 4 of](#) the Census Act gives the Registrar General the power to prepare statistical reports. Section 4(1) requires that the Registrar General “shall, as soon as may be after the taking of

a census, prepare reports on the census returns.” Inherent in section 4(1) is the power to do what is necessary to prepare such reports, including the acquisition and processing of necessary data.

The proposal, under the Census Estimation and Adjustment strand, is that the Administrative data obtained from the NHS Central Register (NHSCR) will be used for estimation (to estimate missing records) and adjustment purposes (to impute responses for those missing records). The values imputed using data from NHSCR will be disclosed to organisations and people such as researchers for statistical and research purposes, however these imputed values are not actual collated information about those individuals but information that has been imputed based on similar records.

The proposed use of data in the census estimation and adjustment strand is in the process of being prepared to be taken to the Public Benefit and Privacy Panel for Health and Social Care (PBPP) for approval. This group looks at the use of health data within research/data linkage projects. Additionally, work is currently underway to take the proposed use of data for the Census Estimation and Adjustment strand to the Statistical Public Benefit and Privacy Panel – the group which looks at the use of Scottish Government and Census data within Research and linkage projects.

**Describe the purposes of the processing:** what do you want to achieve? What is the intended effect on individuals? What are the benefits of the processing – for NRS, and more broadly?

The Census is key to the delivery of many public services. The size, age, sex and geographic distribution of the population are important statistics. This information also drives statistics about changes in the population and the factors driving these changes. These statistics have a wide range of uses. Central government, local government and the health sector use them for planning, resource allocation and managing the economy. They are also used by people such as actuaries for pricing pensions, market researchers and academics.

The processing will have no direct effect on any individual.

#### **Census Quality Assurance Strand**

The use of administrative datasets as an independent source to improve the quality of this statistical outputs can only benefit the quality of outputs for all census users.

### **Census Estimation and Adjustment Strand**

Scotland's Census 2022 achieved lower than expected response rates. In light of this development, National Records of Scotland propose to introduce administrative data from a range of sources, including the NHSCR, to improve our statistical processes used to estimate the Scottish Population, including those who did not complete a census return, at a national and local level.

### **Step 3: Consultation process**

**Consider how to consult with relevant stakeholders:** describe when and how you will seek individuals' views – or justify why it's not appropriate to do so. Who else do you need to involve within your organisation? Do you need to ask your processors to assist? Do you plan to consult information security experts, or any other experts? Describe the groups you will be consulting with and their interest in the project. Who should be consulted internally and externally? Explain the method you will use for consultation with any stakeholder groups and how you will communicate the outcomes of the DPIA back to them. How will you carry out the consultation? Explain what you learned from the consultation process and how they shaped your approach to the management of privacy risks. Explain what practical steps you will take to ensure that you identify and address privacy risks. You should link this to the relevant stages of your project management process. You can use consultation at any stage of the DPIA process.

NRS regularly consults with the International Census Forum (ICF), on use of administrative data in quality assuring the census. The ICF includes organisations from around the world which use administrative data to improve census results. The membership includes the Office for National Statistics (ONS, UK), Northern Ireland Statistics and Research Agency (NISRA), Central Statistics Office (CSO, Ireland), Australian Bureau of Statistics, Stats New Zealand, US Census Bureau and Stats Canada.

NRS are also consulting with ONS and NISRA on the development of the use of administrative data in population statistics. NISRA has published proposal papers on how the Health Card Register can help produce population estimates under a linked-record level administrative data.<sup>12</sup> ONS are further on in their Census Transformation programme, they are using a number of

---

<sup>12</sup> <https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/the-future-provision-of-census-of-population-information-for-northern-ireland.pdf>

administrative datasets to move towards an Administrative Data Census and they have created a number of administrative-data based population outputs.<sup>13</sup>

All these projects are examining how to link data at an individual level to create a statistical population dataset that can improve the quality of population statistics and looking at how this will inform and improve future censuses. Working with these other organisations has helped us understand how the use of administrative data can help improve the quality of census 2022 outputs for Scotland.

### **For Quality Assurance Strand**

The proposal to use NHSCR data to quality assure the 2022 census was approved by the Public Benefit and Privacy Panel (PBPP) for Health and Social Care. Within NRS the project is overseen by the Head of Profession for Statistics and the Census Information Asset Owner (IAO).

NRS are linking reduced datasets for quality assurance purposes. The NHSCR data is to help us make statistical decisions about census data. Every effort has been made to minimise the number of persons with access to the unprocessed datasets. All processing of the personal data at NRS is auditable and the data is held within a secure environment

Other related DPIAs have been carried out, including the overall DPIA for Scotland's Census 2022 which has been published on the website: <https://www.scotlandscensus.gov.uk/documents/data-protection-impact-assessment/>

As part of the Census Programme stakeholder engagement, in February 2020 the Admin Data Team has attended NRS stakeholder events on Census Methodology, discussing the use of administrative data to quality assure the Census 2022. <https://www.scotlandscensus.gov.uk/statistical-methodology-stakeholder-events>

Slides from this event are made available on Census 2022 website: [https://www.scotlandscensus.gov.uk/documents/Statistical Methods Presentation.pdf](https://www.scotlandscensus.gov.uk/documents/Statistical_Methods_Presentation.pdf)

---

<sup>13</sup> <https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusannualassessments>

### **For estimation and adjustment strand**

An [International steering group](#) was formed in May 2022. This is a group of administrative data and census experts, who have scrutinised and provided direction for this project. A statement from the group can be found on the Census website alongside minutes from all meetings.

The NHSCR Stakeholder reference group is also being approached to provide advice on use of NHSCR.

There are likely to be more stakeholder events covering the using of administrative data in the future.

### **Step 4: Assess necessity and proportionality**

**Describe compliance and proportionality measures, in particular:** what is your lawful basis for processing? Does the processing actually achieve your purpose? Is there another way to achieve the same outcome? How will you prevent function creep? How will you ensure data quality and data minimisation? What information will you give individuals? How will you help to support their rights? What measures do you take to ensure processors comply? How do you safeguard any international transfers?

#### **Legal Gateway**

##### **All Strands (QA, Census Estimation and Adjustment)**

The Registration of Births, Deaths and Marriages (Scotland) Act 1965 (1965 Act) section 1(3) provides that “The Registrar General shall exercise the functions pertaining to the said office by virtue of this or any other Act; and subject to that general duty, shall have power to do all such things as appear to him necessary or expedient for maintaining the utility and efficiency of the registration service in Scotland”.

The Local Electoral and Registration Services (Scotland) Act 1965 section 57 provides that the Registrar General may create and maintain a register of individuals [the NHSCR] for the purpose of “facilitating the carrying out of the functions of Health Boards, the Common Services Agency and local authorities”, without prejudice to section 1(3) of the 1965 Act.

The Census Act 1920 section 2(1) provides that “It shall be the duty of the Statistics Board in relation to England and Wales and the Registrar General for Scotland in relation to Scotland to make such arrangements and do all such things as are necessary for

the taking of a census in accordance with the provisions of this Act and of any Order in Council or regulations made thereunder, and for that purpose to make arrangements for the preparation and issue of the necessary forms and instructions and for the collection of the forms when filled up.”

The Census Act 1920 section 4(1) further provides that “The Statistics Board and the Registrar General for Scotland respectively shall, as soon as may be after the taking of a census, prepare reports on the census returns, and every such report shall be printed and laid before both Houses of Parliament.” Inherent in section 4(1) is the power to do what is necessary to prepare such reports, including the acquisition and processing of necessary data.

The Registrar General has determined that in order to improve the quality and accuracy of Scotland’s Census 2022 from which statistical reports shall be prepared it is necessary to use data from the NHSCR.

### **Lawful basis for processing**

#### **All Strands (QA, Census Estimation and Processing)**

The lawful basis for processing personal data is provided by Article 6(1)(e) of the UK General Data Protection Regulation (UK GDPR):

“processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;”

The parties are satisfied that for processing special category personal data condition 2(j) of Article 9 of the UK GDPR is met:

“processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on domestic law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.”

The basis in domestic law is provided by Data Protection Act 2018 Schedule 1 Part 1 Paragraph 4: “This condition is met if the processing –

(a) is necessary for archiving purposes, scientific or historical research purposes or statistical purposes,  
(b) is carried out in accordance with Article 89(1) of the UK GDPR (as supplemented by section 19), and  
(c) is in the public interest.”

### **Necessity and proportionality of processing**

The Registrar General for Scotland, whose function NRS fulfils, has a **statutory duty** under section 2(1) of the Census Act 1920 to “make such arrangements and do all such things as are necessary for the taking of a census”. Section 4 of the Census Act gives the Registrar General the power to prepare statistical reports. Section 4(1) requires that the Registrar General “shall, as soon as may be after the taking of a census, prepare reports on the census returns.” Inherent in section 4(1) is the power to do what is necessary to prepare such reports, **including the acquisition and processing of necessary data**. In order to improve the quality and accuracy of the 2022 Census the Registrar General is seeking to use information from the NHSCR.

### **Census Quality Assurance Strand**

Only the minimum data needed for this project is being requested from NHSCR.

The Admin Data Project analysed multiple datasets, but as has been explained in Step 1 only the NHSCR dataset was identified as being suitable for census quality assurance purposes. The processing of the NHSCR is necessary because there is no other reasonable means to achieve the same level of quality assurance and improve the accuracy of the 2022 Census estimates.

The data linkage proposal has been approved by the Public Benefit and Privacy Panel (PBPP) for Health and Social Care. The data requested has been restricted to a limited, set number of variables that would allow this project to proceed as per the application proposal. All data processing will take place in Scotland.

### **Census Estimation and Adjustment Strand**

Only the minimum data needed for this project is being requested from NHSCR.

The data from NHSCR which is needed and why is explained in Annex A below.

**Could the results obtained by using NHSCR be achieved by another method and/or dataset?**

- The NHSCR is used to identify individuals who appear on other administrative datasets but are no longer part of the Scotland population, and hence need to be removed from the dataset. The NHSCR is the only known Scottish dataset that covers everyone who has lived in Scotland, whilst identifying those who are not currently living in Scotland. NRS death registration data could provide information on people who have died, but this will only cover individuals who have died in Scotland. The NHSCR indicates a death date for individuals who have died elsewhere. Without NHSCR we could end up including individuals who previously lived in Scotland but died elsewhere. The Electoral Register can identify individuals who were previously registered to vote in Scotland but are now overseas voters. However, this does not include children, and does not include individuals who are now registered to vote elsewhere in the UK. HESA data can identify individuals who are domiciled in Scotland but study elsewhere in the UK. However, this does not include individuals who study outwith the UK, and only covers a comparatively small proportion of the population. The NHSCR itself indicates that most of the individuals who leave Scotland remain in the UK. Apart from students, the NHSCR is the only known source that can identify such individuals. The Registrar is data controller for the NHSCR and the information required is held within NRS – albeit a different area of NRS to the team performing this project. The minimum data needed from NHSCR information to meet the projects purpose (as detailed above) could be shared efficiently, securely and effectively within a shorter timeframe than requesting information from data controllers of other administrative data sources holding the same information.
- NHSCR data is available within the timescales of this project, allowing NRS to deliver census outputs within the timescales needed by data users. Missing this deadline would mean that the Barnett formula, Policies and decisions reliant on quality accurate census 2022 data would be using old or inaccurate data. As stated in the benefits section earlier the UK Government uses the Barnett formula to allocate funding across the United Kingdom and this will be based on census data for the next decade. If population estimates are inaccurate, Scotland’s funding could be inaccurate over the long term. For every 16,000 people by which the census underestimates the Scotland population (0.3%), the Scottish Government could lose around £1 billion from the block grant over the next decade<sup>14</sup>.

---

<sup>14</sup> Detailed calculations can be provided on request. [Block Grant Transparency 2021 Explanatory Note .pdf \(publishing.service.gov.uk\)](https://publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/101423/block-grant-transparency-2021-explanatory-note.pdf)



### **The extent of the improvement caused by using information from NHSCR as opposed to other Admin datasets**

- NRS has performed analysis based on 2019 data which shows without NHSCR data our population spine would be significantly too high, which would affect our estimation and adjustment processes.
- Errors of the magnitude explored in this analysis would likely bring the precision of our estimates above our Key Performance Indicator.
- Other available administrative datasets have been reviewed with the Census International Steering Group and it was decided there is not another administrative dataset that has the necessary population coverage and identifies individuals who are no longer in the Scotland population. The minutes of these meetings will be published on our website shortly.
- Analysis of 2019 data suggests that there are many individuals who on the other administrative datasets being considered, but who the NHSCR indicates as have died or left Scotland. Without access to the NHSCR data these individuals would likely end up on the administrative data spine, despite generally not being in the Scotland population. If such data were used in estimation this would cause bias in the census results. There are enough of these cases that the problems introduced by the administrative data would outweigh any benefits of using the administrative data in estimation.

**Step 5: Identify and assess risks**

**Describe source of risk and nature of potential impact on individuals.** Include associated compliance and corporate risks as necessary. Larger-scale DPIAs might record this information on a more formal risk register.

NRS has implemented a comprehensive Information Risk Management framework for Scotland’s Census 2022. Through this independently approved process it is managing a broad range of risks that have been identified as potentially affecting either the confidentiality, integrity or availability of the processes, systems or data collected via the census. Descriptions of the keys risk can be found in the Scotland’s Census 2022 DPIA. Specific risks relating to the Admin Data Population and Household Estimates Project are identified and assessed below.

**Likelihood and Severity of Harm Matrix:**

	<b>Minimal</b>	<b>Significant</b>	<b>Severe</b>
<b>Remote</b>	Low	Medium	Medium
<b>Possible</b>	Low	High	High
<b>Probable</b>	Medium	High	High

<b>No.</b>	<b>Risk and potential impact</b>  <b><u>Risks to individuals</u></b>  <ul style="list-style-type: none"> <li>harm or distress to an individual or group of individuals</li> </ul>	<b>Likelihood of harm</b>  (Remote, possible or probable)	<b>Severity of harm</b>  (minimal, significant or severe)	<b>Overall risk</b>  (low, medium or high)
<b>1a)</b>	<b>Incorrect Legal Gateway being used</b> – resulting in unfair and unlawful processing. This could cause distress to individuals to individuals.	<b>Possible</b>	<b>Severe</b>	<b>High</b>

1b)	<p><b>Unauthorised disclosure of information – IT</b> - There is a possibility that data processed at the National Safe Haven and on NRS systems could be compromised. This could result in personal data provided in census questionnaires being compromised or lost or subject to misuse or identity fraud.</p>	<b>Possible</b>	<b>Significant</b>	<b>High</b>
1c)	<p><b>Unauthorised Access of information - Personnel</b> -There is a risk of unauthorised access to identifiable data within the data processing environment to unauthorised staff. This could result in a possible breach of confidentiality causing potential distress and frustration to individual whose data has been accessed by unauthorised means.</p>	<b>Remote</b>	<b>Significant</b>	<b>Medium</b>
1d)	<p><b>Invasion of privacy</b> - There is a risk that robust profiles of individuals could be built up from combining NHSCR with administrative data sets provided by other public bodies. This could result in a greater invasion of privacy for individuals where more information is exposed than they would like. The risk here is that the individuals may not be aware of the extent of the data linkage from other provided sources which is increasing the breadth of width of the information held by NRS. This could potentially result in frustration and distress to individuals who have more information exposed than they would like due to data linkage.</p>	<b>Possible</b>	<b>Minimal</b>	<b>Low</b>
1e)	<p><b>Function creep</b> - There is a risk of function creep where the project is potentially used for unexpected or unintended future purposes. This could result in a breach of the law and potentially</p>	<b>Possible</b>	<b>Minimal</b>	<b>Low</b>

	cause distress and frustration to individuals who may not be aware of future data processing activities			
1f)	<p><b>Personal Risk – Storage/Transfer loss</b></p> <p>There is a risk that personal data are lost due to poor storage and transfer processes. This could result in potential harm, compromised data and identity fraud to those whose personal data is involved in the breach. This risk includes manual transfer of data.</p>	Possible	Severe	High
1g)	<p><b>Personal Risk (Insider Threat)</b></p> <p>There is a risk that employees, who are authorised to access data, may exploit their access to misuse or steal personal data. This could result in harm, identity fraud, financial loss and distress and upset to the relevant data subject or household involved.</p> <p>There is a risk of staff not having appropriate qualifications or training to manage the data in the appropriate way.</p>	Rare	Severe	Medium
1h)	<p><b><u>Personal Risk (Re-identification)</u></b></p> <p>There is a risk that individuals are identified or perceived to be identifiable through published tables. This could result in the disclosure of personal or sensitive data about a specific individual or household which could cause distress since this information would constitute as personal information.</p>	Possible	Significant	Medium
1i)	<p><b><u>Personal Risk - IAO/Director of Statistical Services potential perceived Conflict of interest</u></b></p> <p>The Information Asset Owner is also the Director of Statistical Services at NRS and responsible for the statistical component of Census delivery. It is their decision ultimately to accept any risks associated with the legal gateway and using NHSCR for this purpose as well as all</p>	Possible	Significant	High

	other risks. This may appear to be a potential conflict of interest by the Information Commissioners Office and/or the public.			
<b>2</b>	<p><b><u>Risk to NRS</u></b></p> <ul style="list-style-type: none"> <li>• reputational damage to NRS and/or the data providers;</li> <li>• possible enforcement action against NRS;</li> <li>• loss of confidence in NRS and/or the data providers; and,</li> <li>• loss of public finances to NRS and/or data providers.</li> </ul>			
<b>2a</b>	<p><b><u>Reputational Risk – NRS and NHSCR</u></b></p> <p><b><u>Reputational Risk due to unlawful use of NHSCR data</u></b></p> <p>NRS must be able to satisfy itself that using the NHSCR rather than other data sources or methods for both the Quality Assurance Strand and Census Estimation and Adjustment Strand is necessary for the purposes of improving the quality and accuracy of the Census 2022. The justifications for using NHSCR and limitations of other data sources have been discussed earlier .</p>	<b>Possible</b>	<b>Severe</b>	<b>High</b>
<b>2b)</b>	<p><b>Reputational Risk - Unauthorised disclosure of information by NRS.</b></p> <p>Examples include inadvertent compromise by a statistical process such as disclosure control, deliberate compromise by a member of staff or a targeted attack by cyber criminals</p>	<b>Possible</b>	<b>Significant</b>	<p><b>High</b></p> <p><b>Low</b></p>

2c)	<b>Vulnerability in or malfunction of security controls</b> - Data subjects may have privacy concerns relating to the security of information technology used to process their data and about the extent of organisational measures in place to protect their data.	<b>Possible</b>	<b>Significant</b>	<b>High</b>
2d)	<b>Unlawful/accidental Information sharing of data</b> -There is a risk that data is inappropriately shared due to the failure to apply statistical disclosure controls or follow robust information governance processes	<b>Remote</b>	<b>Significant</b>	<b>Medium</b>

**Step 6: Identify measures to reduce risk**

Identify additional measures you could take to reduce or eliminate risks identified as medium or high risk in step 5.

<b>No.</b>	<b>Options to reduce or eliminate risk</b>	<b>Effect of risk</b> (eliminated, reduced or accepted)	<b>Residual risk</b> (low, medium or high)	<b>Measure approved</b> (yes, no)
1a	<p><b>Measures - Incorrect Legal Gateway being used</b> – resulting in unfair and unlawful processing</p> <ul style="list-style-type: none"> <li>• Basis in law for processing has been confirmed (see Step 4).</li> <li>• Registrar General for Scotland, whose function NRS fulfils, has duty to take the census and prepare reports on census returns.</li> <li>• This DPIA assesses fairness and lawfulness of processing.</li> <li>• The Data Protection Officer will provide advice and guidance.</li> <li>• The Information Asset Owner will consider the advice of the DPO to assess risks.</li> </ul>	<b>Reduced</b>	<b>Low</b>	<b>Yes</b>
1b	<p><b>Measures - Unauthorised disclosure of information – IT</b></p> <ul style="list-style-type: none"> <li>• National Safe Haven provides a secure analytic environment providing with secure file transfer on IT infrastructure provided by EPCC at University of Edinburgh.</li> <li>• A comprehensive security programme of policies and procedures have been implemented which are aligned to current regulatory</li> </ul>	<b>Reduced</b>	<b>Medium</b>	<b>Yes</b>

	<p>legislation and industry standards e.g. UK and EU GDPR, Data Protection Legislation, National Cyber Security Centre (NCSC) etc.</p> <ul style="list-style-type: none"> <li>• Strong, auditable security controls between the Census Data Processing Environment, the rest of NRScotland and the SCOTS network have been put in place.</li> <li>• Frequent audits, penetration tests, vulnerability scanning and monitoring is of NRS IT infrastructure have been implemented whilst assurance of Scottish Government network will be requested and validated frequently.</li> </ul>			
<b>1c</b>	<p><b>Measures - Unauthorised Access of information – personnel</b></p> <ul style="list-style-type: none"> <li>• Robust security controls are in place at the National Safe Haven to ensure that only approved researchers can access research data.</li> <li>• Access controls policies have been put in place at NRS to ensure only relevant key staff have access to NRS: Admin Area.</li> <li>• Named individuals are only allowed to work on the project. They are required to be approved by and meet the training requirements of the PBPPs.</li> <li>• NRS staff have passed as a minimum the Baseline Personnel Security Standard (BPSS) and follow the code of official statistics.</li> <li>• Regular IT audit access reviews are carried out to monitor access privileges and joiner, movers and leavers.</li> <li>• All NRS staff involved in processing Census data sign the Census Confidentiality Undertaking. which is underpinned by the Census Act 1920, prohibiting the sharing or unauthorised use of census data. This Act makes it a criminal offence, punishable by imprisonment, a fine or both, for any person to disclose any personal census information to another person without lawful authority. Data Provider and NRS will complete and sign a Data Sharing Agreement which will explain and agree to the minimum</li> </ul>	<b>Reduced</b>	<b>Low</b>	<b>Yes</b>



	data requested, purpose, and roles and responsibilities (including dissemination restrictions) of both parties.			
1d	<p><b><u>Measures (Extensive User-Profiles)</u></b></p> <p>NRS has taken steps to inform users that census data will be linked and merged with public admin data sets to improve the quality of the statistical processing requirements. This was highlighted in census privacy information and online census guidance before processing begins.</p>	Reduced	Low	Yes
1e	<p><b><u>Measures (Function Creep)</u></b></p> <ul style="list-style-type: none"> <li>• Approvals, evaluations and policy reviews for increased use of census data will be required.</li> <li>• Any future uses will be advised in the census privacy notices prior to processing so all respondents are fully informed of how their data will used.</li> </ul>	Reduced	Low	Yes
1f	<p><b><u>Measures Personal Risk – Storage/Transfer loss</u></b></p> <ul style="list-style-type: none"> <li>• The administrative datasets will be transferred into the secure admin data processing area, to be quality assured and standardised by a member of the admin data team. This team member can only access one dataset at a time. Quality assured and standardised data is then transferred to the National Safe Haven.</li> <li>• Appropriate storage policies/procedures that outline specific physical security controls are in place within NRS to manage admin data received.</li> <li>•</li> </ul>	Reduced	Medium	Yes

	<ul style="list-style-type: none"> <li>• Frequent physical security assessments will be conducted by NRS to ensure NRS sites processing data are safe and secure. Security Improvement Plans will be formed to track mitigation actions.</li> <li>• Appropriate security and Information Governance training is in place and has been provided to NRS staff.</li> <li>• All personal data will only be stored in the UK.</li> <li>• Access to personal data will be based on job role requirements only. All staff will have up to date security checks in place for security clearance purposes.</li> <li>• All security controls for the storage, transfer and destruction of data will be aligned to UK and EU GDPR, the ISO 27001 Security standard, the ISO 15489 Records Management Standard and guidance from the National Cyber Security Centre (NCSC).</li> <li>• All Census Information Asset Registers will be up to date to reflect that a relevant Information Asset Owner (IAO) has ownership of a physical asset for each service area whilst key information risks for each asset will be highlighted in the Corporate Census Information Risk Register.</li> </ul>			
1g	<p><b><u>Measures - Personal Risk (Insider Threat)</u></b></p> <ul style="list-style-type: none"> <li>• All employees have a minimum level of security clearance to Baseline Personnel Security Standard (BPSS).</li> <li>• All employees with access to census data will receive data protection training and are appropriately qualified for their roles.</li> <li>• Employees will only have access to data required to perform their role. <ul style="list-style-type: none"> <li>• All NRS staff involved in processing Census data sign the Census Confidentiality Undertaking. which is underpinned by the Census Act 1920, prohibiting the sharing or unauthorised use of census data. This Act makes it a criminal offence, punishable by imprisonment, a fine or both, for any person to disclose any</li> </ul> </li> </ul>	Reduced	Medium	Yes

	<p>personal census information to another person without lawful authority.</p> <ul style="list-style-type: none"> <li>• Security incident and event monitoring tools will be implemented.</li> </ul>			
1h	<p><b><u>Measures - Personal Risk (Re-identification)</u></b></p> <ul style="list-style-type: none"> <li>• No information from NHSCR will be made public in any form.</li> <li>• All statistical disclosure control documentation will follow industry standards and best practices to avoid the re-identification of individuals in published statistical tables and research data. These methodologies have been assessed through External Methodology Assurance Panels (EMAPs) and aligned to other UK census offices where they have been reviewed the UK Census Committee before being approved.</li> <li>• System statistical processes will be validated and assured before the census goes live to ensure special characteristics of individuals, households or groups remain protected.</li> </ul> <p>NRS has many years' experience of deidentifying data and ensuring safe research practices are deployed in our census operations.</p>	Reduced	Low	Yes
1i	<p><b><u>Measures - IAO/Director of Statistical Services potential perceived Conflict of interest</u></b></p> <ul style="list-style-type: none"> <li>• This DPIA assesses the risks of processing for all approvers of the DPIA and DPO to consider not just the IAO</li> <li>• The Data Protection Officer will provide advice and guidance on Data Protection issues/concerns.</li> <li>• The Director of Statistical Services (who is also the IAO) will consider the advice of the DPO to assess risks and authorise the DPIA, and when the steering group has given their advice will re-review this DPIA.</li> </ul>	Reduced	Medium	Yes

	<ul style="list-style-type: none"> <li>• The NHSCR stakeholder reference group (steering group with external stakeholders) will be approached to review and advise upon use of NHSCR for this project.</li> <li>• The QA strand has been approved by PBPP</li> <li>• The Census Estimation and Adjustment Strand is going to be reviewed by PBPP shortly.</li> </ul>			
2a	<p><b>Measures - Incorrect Legal Gateway being used</b></p> <ul style="list-style-type: none"> <li>• Basis in law for processing has been confirmed (see Step 4).</li> <li>• Registrar General for Scotland, whose function NRS fulfils, has duty to take the census and prepare reports on census returns.</li> <li>• This DPIA assesses fairness and lawfulness of processing.</li> <li>• The Data Protection Officer will provide advice and guidance.</li> <li>• The Information Asset Owner will consider the advice of the DPO to assess risks.</li> </ul>	<b>Reduced</b>	<b>Low</b>	<b>Yes</b>
2b	<p><b>Unauthorised disclosure of information.</b></p> <ul style="list-style-type: none"> <li>• All persons who may come into contact with census information will be required to sign the Census Confidentiality Undertaking which is underpinned by the Census Act 1920, prohibiting the sharing or unauthorised use of census data. This Act makes it a criminal offence, punishable by imprisonment, a fine or both, for</li> </ul>	<b>Reduced</b>	<b>Low</b>	<b>Yes</b>

	<p>any person to disclose any personal census information to another person without lawful authority.</p> <ul style="list-style-type: none"> <li>• NRS is committed to ensuring that privacy of every individual whose data will be collected and processed as part of this programme will be protected. All statistical outputs produced by NRS fully comply with the Code of Practice for Official Statistics: Under section T6 on Data governance: Organisations should look after people's information securely and manage data in ways that are consistent with relevant legislation and serve the public good</li> <li>• A member of Admin Data team will de-identify personal identifiable data used in the study to minimise the amount of personal identifiable data used in the study.</li> </ul>			
<b>2c</b>	<p><b>Measure - Vulnerability in or malfunction of security controls</b></p> <ul style="list-style-type: none"> <li>• Appropriate security controls are on NRS servers, data will be processed in secure IT area with access limited to named staff or securely in the National Safe Haven.</li> <li>• Administrative datasets will be stored in the secure admin data area, with access restricted to authorized personnel.</li> <li>• Datasets will be encrypted at rest, and only decrypted when it is needed for data processing. The unencrypted file will be deleted when processing is complete.</li> <li>• NRS has a Security Breach process in place in the event that a personal data breach should occur.</li> </ul> <p>All staff receive mandatory security training.</p>	<b>Reduced</b>	<b>Medium</b>	<b>Yes</b>

2d	<p><b>Measure - Inappropriate sharing of data.</b></p> <ul style="list-style-type: none"> <li>• Datasets received from providers will not be shared outwith the named individuals from the admin data team.</li> <li>• Raw data used for these projects will not be made publically available</li> <li>• Robust, transparent, consistent, and proportionate information governance will be applied to all external data requests.</li> </ul>	Reduced	Low	Yes
----	--	---------	-----	-----

<b>Step 7: Sign off and record outcomes</b>		
<b>Item</b>	<b>Name/date</b>	<b>Notes</b>
Measures approved by:	NRS Director of Statistical Services	Integrate actions back into project plan, with date and responsibility for completion
QA DPO advice provided:	NRS Head of Information Governance 29/04/2022	DPO should advise on compliance, step 6 measures and whether processing can proceed
Census Estimation and Adjustment DPO advice provided:	NRS Data Protection Officer 01/09/2022	
<p><b>QA Summary of DPO advice:</b> The necessity of using the NHSCR dataset for quality assurance of Scotland's Census 2022 has been evidenced and the basis in law for processing the data has been established. Extensive consultation and peer review has taken place. Appropriate security controls and other measures have been identified under step 6 to reduce the risks associated with the processing identified under step 5.</p> <p><b>Census Estimation and Adjustment Summary of DPO Advice:</b> The necessity of using the NHSCR dataset for Estimation and Adjustment to Census Processing has been evidenced and the basis in law for processing the data has been established. Consultation has taken place, including with independent experts. Appropriate security controls and other measures have been identified under step 6 to reduce the risks associated with the processing identified under step 5, and the seeking of DPO advice has mitigated any potential/perceived conflict of interest on the part of the IAO.</p>		
DPO advice accepted or overruled by:	NRS: Director of Statistical Services 02/09/2020	If overruled, you must explain your reasons

Comments: Advice accepted.		
Consultation responses reviewed by:	N/A	If your decision departs from individuals' views, you must explain your reasons
Comments:		
This DPIA will kept under review by:	NRS: Head Of Admin Data	The DPO should also review ongoing compliance with DPIA



### Part 3: Linking the DPIA to the UK GDPR and EU GDPR data protection principles

Answering these questions during the DPIA process will help you to identify where there is a risk that the project will fail to comply with the UK GDPR and EU GDPR or other relevant privacy legislation, including the Human Rights Act.

#### UK GDPR and EU GDPR Principle (a) (Article 5(1)(a))

Personal data shall be processed fairly and lawfully and, in particular, shall not be processed unless:

- a) at least one of the conditions in Article 6 is met, and
- b) in the case of special category personal data, at least one of the conditions in Article 9 is also met.

Have you identified the purpose of the project?

#### **Census Quality Assurance Strand**

The purpose of the project is to enable quality assurance and to improve the accuracy of the 2022 Census estimates.

#### **Census Estimation & Adjustment Strand**

Administrative data obtained from the NHS Central Register (NHSCR) will be used for estimation (to estimate missing records) and adjustment purposes (to impute responses for those missing records). This will improve the quality and accuracy of the 2022 Census.

How will you tell individuals about the use of their personal data?

This has been achieved by the publication of the overarching Census DPIA (found here <https://www.scotlandscensus.gov.uk/documents/data-protection-impact-assessment>) and by the other privacy information we do and will make available on our websites and through other communication channels.

The Census Privacy Notice found here <https://www.scotlandscensus.gov.uk/privacy/> will also inform individuals about the use of their personal data.

Do you need to amend your privacy notices?

NRS continuously reviews its Privacy Notices to ensure that they reflect the current position. We make clear that NRS receives data from other public bodies and uses this for statistical and research purposes.

We will update the Census Privacy Notice and the NRS Privacy Notice to reflect using Administrative data for Census QA, estimation and adjustment. This will be done before any processing using Administrative data for these purposes starts so that the public are informed.

Census Privacy Notice <https://www.scotlandscensus.gov.uk/privacy/>

NRS Privacy Notice Privacy | National Records of Scotland ([nrscotland.gov.uk](http://nrscotland.gov.uk))

We will also amend the 'Who has access to your data' page on the NHSCR section of the NRS website [National Health Service Central Register | National Records of Scotland \(nrscotland.gov.uk\)](http://National Health Service Central Register | National Records of Scotland (nrscotland.gov.uk)) to explain that NHSCR data is being used for improving the quality and accuracy of the Census.

NRS will publish statement on the website detailing Census estimation and adjustment strand. Link to this will be added once published.

Minutes from the International Steering Group will be made available on our website. Link to this will be added once published.

Have you established which conditions for processing apply?

The lawful basis for processing is provided by Article 6(1)(e) of the General Data Protection Regulation (GDPR): "processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller".

For the processing of special categories of personal data condition 2(j) of Article 9 of the GDPR is met: "processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject."

If you are relying on consent to process personal data, how will this be collected and what will you do if it is withheld or withdrawn?

Consent is not being relied on or sought.

If your organisation is subject to the Human Rights Act, you also need to consider:  
Will your actions interfere with the right to privacy under Article 8?

Processing of personal data by professional statisticians to produce aggregate statistics or census estimates and adjustments is not envisaged to present any additional interference with the privacy rights of individuals. The provisions of Article 8 as incorporated in the Human Rights Act 1998 allow public authorities to enquire into a person's private life where they have a legal authority to do so and where such an enquiry is necessary in a democratic society for one of the aims stated in the Article. Lawful authority is given by the Census Act 1920 and the processing is necessary for the economic well-being of the country and for the purposes of the protection of health and the rights and freedoms of others.

Have you identified the social need and aims of the project?

The use of accurate Census data guides significant Government expenditure and provides material benefit to society.

Are your actions a proportionate response to the social need?

Yes. The approach identified enables the use of administrative data to meet the social needs and aims of the project, whilst balancing the need for individual privacy.

Census QA, estimation and adjustment will all improve the accuracy and quality of the 2022 Census data.

**Benefits**

- The objective of this proposal is to use administrative data in the census statistical processes to **ensure the 2022 Census population estimates are as accurate as possible at a national and local level.**
- Census population estimates are used to allocate funding across the United Kingdom, through the Barnett formula. The Scottish Government in turn uses census data to allocate funding to local authorities, health boards and other public services. For funding to be allocated correctly, is essential that population estimates are as accurate as possible. **Therefore this project has a direct benefit to all people across the United Kingdom.**
- The population estimates from Scotland's Census are used in a multitude of other ways – to plan policies, deliver services and conduct research by national and international organisations. More detail of these benefits is set out below.
- The method proposed will allow us to produce accurate outputs within the time periods census users need

#### **UK GDPR and EU GDPR Principle (b) (Article 5(1)(b))**

Personal data shall be obtained only for one or more specified and lawful purposes, and shall not be further processed in any manner incompatible with that purpose or those purposes.

Does your project plan cover all of the purposes for processing personal data?

Safeguards are in place to ensure that the data collected are only used for lawful purposes.

Have you identified potential new purposes as the scope of the project expands?

These projects are for the purposes of Census 2022 quality assurance and estimation. Any expansion in the use of data for these projects will need to be reviewed and approved by the Public Benefit and Privacy Panels, data owners and Information Asset owner.

The use of NHSCR for this project is new and is to be discussed at the NHSCR stakeholder reference group for their review and advice. This DPIA will be reviewed again once the NHSCR Stakeholder reference group has provided their advice.

**UK GDPR and EU GDPR Principle (c) (Article 5(1)(c))**

Personal data shall be adequate, relevant and not excessive in relation to the purpose or purposes for which they are processed.

Is the quality of the information good enough for the purposes it is used?

We are only requesting the minimum personal data we need to achieve the purpose of the project.

Quality assurance processes will be put in place to ensure that the statistics produced using the information collected are fit for purpose and best meet the needs of data users. The Quality Assurance of Administrative Data (QAAD) relating to the NHSCR dataset has been published as part of the Census Statistical Quality Assurance Report.

It can be found here

[Quality Assurance of Administrative Datasets | Scotland's Census \(scotlandscensus.gov.uk\)](https://scotlandscensus.gov.uk/quality-assurance-of-administrative-datasets)

Which personal data could you not use, without compromising the needs of the project?

The nature of the project dictates that we need to know basic demographic information about people – their age and sex – and where they live.

**UK GDPR and EU GDPR Principle (d) (Article 5(1)(d))– accurate, kept up to date, deletion**

Personal data shall be accurate and, where necessary, kept up to date.

If you are procuring new software does it allow you to amend data when necessary?

No. This is a statistical project. We will be using SAS, Microsoft Excel and Microsoft Access to process the data.

How are you ensuring that personal data obtained from individuals or other organisations is accurate?

The Quality Assurance of Administrative Data (QAAD) relating to the NHSCR dataset has been published as part of the Census Statistical Quality Assurance Report and s report on quality assurance measures taken. It can be found at: [Quality Assurance of Administrative Datasets | Scotland's Census \(scotlandscensus.gov.uk\)](https://scotlandscensus.gov.uk/quality-assurance-of-administrative-datasets)

### **UK GDPR and EU GDPR Principle (e) (Article 5(1)(e))**

**Personal data processed for any purpose or purposes shall not be kept for longer than necessary for that purpose or those purposes.**

What retention periods are suitable for the personal data you will be processing?

Data will be retained as per the Data Sharing agreement.

Data is regularly backed up as backups of essential business information must be taken to ensure that NRS can recover from a disaster, media failure or other form of error. After the data is deleted the data will stay in the backup as per NRS Back up Policy until the backup is overwritten.

The project has an end date of December 2024. Data stored in NRS will be securely deleted in accordance with HMG Infosec Standard 5 at the enhanced level and the National Cyber Security Centre (NCSC) guidance on secure sanitisation (<https://www.ncsc.gov.uk/guidance/secure-sanitisation-storage-media>). NRS will provide written confirmation to all providers when the data has been deleted.

NHSCR Data will also be held in the National Safe Haven. The project has a set end date for data access for analysis of December 2024, and then a period of secure retention. On this end date for access for analysis no-one will be able to access the data on the Safe Haven and all requests for access for the purposes of analysis via the National Safe Haven will be denied. At the end of the

retention period, typically 5 years, the data used for this project held on the Safe Haven will be securely and confidentially destroyed.

Are you procuring software that will allow you to delete information in line with your retention periods?

No specialist software is being procured as part of the Admin Data project.

### **UK GDPR and EU GDPR Articles 12-22**

**Personal data shall be processed in accordance with the rights of data subjects under this Act.**

Will the systems you are putting in place allow you to respond to subject access requests more easily?

Existing arrangements allow us to respond to data subject requests. Because the data is being processed for statistics and research purposes only, the rights of data subjects are restricted by Article 89(2) of the GDPR and Schedule 2, Part 6 Paragraph 27 of the Data Protection Act 2018.

If the project involves marketing, have you got a procedure for individuals to opt out of their information being used for that purpose?

Not applicable

### **UK GDPR and EU GDPR Principle (f) (Article 5 (1)(f))**

**Appropriate technical and organisational measures shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data.**

Do any new systems provide protection against the security risks you have identified?

No new systems are being procured for this project.

**Technical and Organisational measures to protect data** include:

NRS currently have Cyber Essentials and Cyber Essentials Plus accreditation

<https://www.ncsc.gov.uk/cyberessentials/search>

National Records of Scotland

Cyber Essentials - Certificate number: IASME-CE-030754

Cyber Essentials Plus - Certificate number: IASME-CEP-006839

Ahead of Scotland's Census 2022, National Records of Scotland asked Bridewell Consulting to undertake an independent review of the strength of our security measures. Bridewell Consulting are a specialist cyber security and data privacy consultancy and are certified by the National Cyber Security Centre (NCSC). The aim of the review was to identify any risks to census systems, services and information. We also wanted to provide an independent view of our security for stakeholders.

The review looked at several parts of our system and how it will be delivered, including:

- the people involved
- processes and technology used
- our supply chain

The review concluded that overall National Records of Scotland has a comprehensive security programme in place. It has been designed to reduce the risk of compromise to the delivery of the census, and to citizen data. The review found that strong controls were also in place to detect and respond to threats that may impact the census when it is in live operation. Additionally, it found that security controls in place have built upon and improved on those in place during the 2019 census rehearsal.



The NRS Technical Security Standards - <https://erdm.scotland.gov.uk:8443/documents/A32279387/details> - covers the mandatory baseline technical security standards of NRS storage of data and access within the physical and technical environments, both for our internal systems and for any 3rd parties.

In order to prevent the infection of NRS computers and networks and avoid the potentially dire consequences of such infection, there are several key controls that will be adopted as policy. These can be found in the NRS end point protection policy - <https://erdm.scotland.gov.uk:8443/documents/A33687913/details> .

Access to data is limited to individual NRS user accounts that need access.

All users must sign and agree to abide to the Census Confidentiality undertaking.

The NRS – IT Access Control Policy section 2.5 - <https://erdm.scotland.gov.uk:8443/documents/A33752030/details> and the NRS logging and monitoring policy - <https://erdm.scotland.gov.uk:8443/documents/A33584358/details> will be applied.

Additionally use of the Safe Haven for the Census Estimation and Adjustment Strand will help protect against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data .

What training and instructions are necessary to ensure that staff know how to operate a new system securely?

No new systems are being procured for this project.

#### **UK GDPR and EU GDPR Article 24**

**Personal data shall not be transferred to a country or territory outside the European Economic Area unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data.**

Will the project require you to transfer data outside of the European Economic Area (EEA)?

No. All data will remain within UK.

If you will be making transfers, how will you ensure that the data is adequately protected?

**All Strands**

Data will be securely transferred within NRS IT systems as it is the data owner. All transfers of data are auditable.

**Census Estimation and Adjustment Strand**

Data will be transferred in and out of the Safe Haven following eDRIS policies for secure transfer.

**Annex A - Table to show variables requested from NHSCR and their justification.**

<b>Data File</b>	<b>Variable Name within NHSCR</b>	<b>Description</b>	<b>Justification</b>
NAME	n_createdate	Date cut of data is prepared	Reference date
	n_modifydate	Date modified	Modified reference date
	Nameid	ID from NHSCR	Alternative linking variable
	nmenddate	Name end date (Death)	Death date for linking
	nmstartdate	Name start date (Birth)	Date of birth for linking
	personid	Linking variable	Linking variable
NAMEPARTS	nameid	ID from NHSCR	
	nametext	Personal name data (text)	Previous names for linking
	nametype	Code to signify previous names where: =0 (Last) =1 (First) =2 (Middle)	Separated out to get all name parts in one record
	personid	Linking variable	Linking
PERSON	birthcountry	Country code if not in Scotland	Birth country if not on vital events (births) for linking
	birthdate	Date of birth	Date of birth for linking
	deathdate	Death date	Death date for linking
	firstname	First name	Name for linking
	gender	Gender	Gender for linking
	last name	Last Name	Name for linking
	last updated		
	middlename	Middle Name (or initial)	Name for linking
nhsnumber	National Health Service number (10 characters)	Payload identifier	

	personid	Linking variable	Linking variable
	posting	Latest posting code (earlier postings are recorded on Postings File but are not used)	Posting for linking and business rules
	ucrn	Unique Citizen Reference Number	Alternative unique identifier (character string)
POSTCODE	pcenddate	End date for postcode	
	pcstartdate	Start date for postcode	
	personid	Linking variable	Linking variable
	postcode	Postcode of property	Location data for linking
	uprn	Unique Property Reference Number	Location data for linking