

**census**  
2021

# 2021 Census Quality Assurance Strategy

January 2021

# Contents

1. Introduction .....	4
2. Background.....	6
3. Quality Assurance End-to-End Journey .....	9
3.1 Question Development.....	9
3.1.1 Consultation.....	10
3.1.2 Testing and Research.....	11
3.1.3 2017 Census Test.....	11
3.1.4 Topic Expert Meetings .....	11
3.1.5 2019 Census Rehearsal .....	12
3.2 Questionnaire Design.....	12
3.3 Census Address Register .....	16
3.3.1 2011 Experience .....	16
3.3.2 2017 Address Check.....	17
3.3.3 2019 Census Rehearsal .....	17
3.3.4 Development of the 2021 Census Address Register .....	17
3.4 Data Collection.....	19
3.4.1 Initial Contact .....	19
3.4.2 Publicity and Support.....	21
3.4.3 Field Force Strategy.....	22
3.5 Data Capture, Coding and Transfer .....	23
3.6 Data Processing.....	24
3.6.1 Removal of False Persons and Resolution of Multiple Responses.....	25
3.6.2 Addition of Census-Under-Enumeration Records .....	25
3.6.3 Edit and Imputation.....	26
3.6.4 Coverage Estimation and Adjustment.....	26
3.6.5 Validation of Population and Household Estimates .....	27
3.6.6 Topic Analysis.....	27
3.6.7 Assignment of Output Geographies .....	28
3.7 Statistical Disclosure Control .....	28

3.8	Output Production and Dissemination.....	29
3.9	Other Important Aspects of the Quality Assurance Work.....	29
4.	Developing the Quality Assurance Approach.....	30

## 1. Introduction

Once every ten years the census provides an opportunity to build a detailed and comprehensive picture of the population, reflecting everyone in our society. The results are invaluable for policy formation, planning services and the effective targeting of resources. The census is comprised of a large number of interlinked processes and the quality of the results is driven by quality assurance procedures being applied at every stage of the census.

Censuses in Northern Ireland are governed by the [Census Act \(Northern Ireland\) 1969](#) and it is intended that a census in Northern Ireland shall be conducted on 21 March 2021.

The UK National Statistician and the Registrars General for Scotland and Northern Ireland released a Statement of Agreement<sup>1</sup> in October 2015 on the conduct of the 2021 Census across the UK to ensure consistent and comparable outputs. In particular, the National Statistician and the Registrars General agreed that the three Census Offices and the Welsh Assembly Government would work together and facilitate harmonisation where it is in the interest of census users. This agreement is unchanged with the decision of Scottish Government to postpone a census of Scotland for 12 months until 2022.

There also exists a statement of agreement between the Director General of Central Statistics Office (CSO), Ireland and the Registrar General for Northern Ireland<sup>2</sup>. The Government of Ireland decided in September 2020 to postpone the Census of Population of Ireland by one year to April 2022. In delivering joint census outputs, CSO and NISRA will continue to work together as outlined in the agreement. This will include working together on an Ireland - Northern Ireland census report to be published after the completion of the censuses by both offices.

The [Registrar General's 2021 Census Proposals Document](#) sets out plans for the next census and provides details of the key strategic objectives which are:

---

<sup>1</sup> For more information please see [The Conduct of the 2021 Censuses in the UK](#)

<sup>2</sup> For more information please see [2021 Census Consultative arrangements](#)

- to provide high quality, value-for-money, fit-for-purpose statistics that meet user needs, and which are consistent, comparable and accessible across the UK;
- to protect, and be seen to protect confidential personal census information;
- to deliver a high quality 2021 online census data collection operation;
- to maximise overall response rates and minimise differences in response rates in specific areas and among particular population groups;
- to secure public and user confidence in the final results and deliver them in a timely manner; and
- to encourage wider use and exploration of census results to facilitate greater benefits from census outputs.

In keeping with the high-level strategic aims referenced above, and in line with corresponding quality documents produced by the other Census Offices across the UK<sup>3</sup>, the purpose of this paper is to outline the Quality Assurance (QA) work that will be undertaken by NISRA in order to ensure that the outputs released from the 2021 Census in Northern Ireland are robust, reliable and meet the needs of users. This document only relates to the quality assurance of the Northern Ireland data.

An important element of this work will be assessing the extent to which any Census estimate varies from those derived from alternative sources and ensuring that the difference is understood, plausible and can be explained to users.

As part of the planning process for the 2021 Census, a number of statutory impact assessments have been undertaken. These include the 2021 Census Regulatory Impact Assessment, 2021 Census Data Protection Impact Assessment, and the 2021 Census Equality Impact Assessment Screening Template<sup>4</sup>.

---

<sup>3</sup> The Office for National Statistics published [Approach and processes for assuring the quality of the 2021 Census data](#). National Records Scotland (NRS) published [Scotland's Census 2021 Statistical Quality Assurance Strategy](#).

<sup>4</sup> These can be found on the [2021 Census – Impact Assessments](#) webpage

If you have any questions or comments in relation to this strategy document, including proposals for other analyses, data sources or methods that could be adopted which would provide reassurance on the reliability of the census results, please send these to [census@nisra.gov.uk](mailto:census@nisra.gov.uk).

## 2. Background

As producers of government statistics, NISRA have a responsibility to manage and report on the quality of statistics, clearly and concisely, to support everybody who needs to use them.

The Government Statistical Service (GSS) defines quality assurance as identifying, anticipating and avoiding problems that can arise from data inputs or the methods and processes used to calculate statistics<sup>5</sup>.

The quality assurance of the previous census of Northern Ireland in 2011 was rigorous and comprehensive<sup>6</sup>. NISRA plan to build on this success for the 2021 Census whilst also taking into account any lessons learned.

The 2011 Census was assessed by the Office for Statistics Regulations (OSR), the regulatory arm of the UK Statistics Authority, as part of a three-phase accreditation process. The statistics produced were found to be fully compliant with the Code of Practice for Statistics and therefore designated as National Statistics. NISRA are again undertaking this accreditation process for the 2021 Census to ensure that National Statistics designation is retained<sup>7</sup>.

The Code of Practice provides a framework for producing statistics that is based on three pillars: Trustworthiness, Quality and Value. Value requires that the statistics and data are

---

<sup>5</sup> [Quality Statistics in Government](#) guidance was published by the GSS in October 2019.

<sup>6</sup> The [Northern Ireland Census 2011 General Report](#) reviews the entire 2011 Census operation from the consultation and planning stages through to the production and dissemination of outputs and evaluation.

<sup>7</sup> Details of OSR assessment of the 2021 Censuses in the UK, including preliminary findings and NISRA responses can be found on the [National Statistics Accreditation](#) webpage.

useful, easy to access, remain relevant, and support understanding of important issues. Trustworthiness comes from the organisation that produces statistics and data being well led, managed and open, and the people who work there being impartial. Lastly, compliance with the Quality pillar ensures statistics fit their intended use, are based on appropriate data and methods, and are not materially misleading. Quality requires skilled professional judgement about collecting, preparing, analysing and publishing statistics in ways that meet the needs of people who want to use the statistics, with producers being able to demonstrate how they assure themselves that their statistics are robust and reliable. The three pillars are distinct but they support each other to deliver public confidence that statistics are of public value, high quality and produced by departments and public bodies that can be trusted.

To assure the quality of statistics the following principles must be adhered to:

1. Statistics should be produced to a level of quality that meets users' needs. The strengths and limitations of the statistics and data should be considered in relation to different uses, and clearly explained alongside the statistics.
2. Quality assurance arrangements should be proportionate to the nature of the quality issues and the importance of the statistics in serving the public good. Statistics producers should be transparent about the quality assurance approach taken throughout the preparation of the statistics. The risk and impact of quality issues on statistics and data should be minimised to an acceptable level for the intended uses.
3. The quality of the statistics and data, including their accuracy and reliability, coherence and comparability, and timeliness and punctuality, should be monitored and reported regularly. Statistics should be validated through comparison with other relevant statistics and data sources. The extent and nature of any uncertainty in the estimates should be clearly explained.
4. Scheduled revisions, or unscheduled corrections that result from errors, should be explained alongside the statistics, being clear on the scale, nature, cause and impact.

5. Systematic and periodic reviews on the strengths and limitations in the data and methods should be undertaken. Statistics producers should be open in addressing the issues identified and be transparent about their decisions on whether to act.

The quality assurance of the 2021 Census aims to:

- ensure that the census results provide a reliable basis for decision-making;
- give data-users confidence that the census results are fit-for-purpose;
- deliver census results as soon as possible;
- minimise the risk of errors in the census estimates, which might undermine the overall credibility of the statistics; and
- leave a legacy of methods, tools and skills for the quality assurance of future population statistics.

In achieving this NISRA will:

- run quality assurance processes to ensure that census processes have worked correctly;
- validate estimates to check that the final results are credible and consistent with other evidence;
- approach quality assurance with the "eyes of a user" and use checks and comparator data sources that a user would be likely to use;
- build trust in our methods by inviting external review of our proposed approach;
- build trust in the census results by being open about their strengths and limitations;
- work in partnership with data-users to understand what evidence would give them confidence in the census estimates;
- quantify accuracy wherever possible - this will allow us to prioritise our checks, help make evidence-based decisions on any interventions required, and help users understand the reliability of the results;
- use flexible methods and tools to react quickly to unexpected anomalies in the census data;



- use expertise across NISRA, making sure that topic experts on subjects such as demography, housing, labour market and health are assessing the stories shown by the provisional census results in the context of other evidence and trends;
- be outward looking, sharing experiences and learning from other census-taking organisations and fostering harmonisation and efficiency by collaborating with teams working on the censuses in England, Scotland and Wales;
- conduct quality assurance in parallel to census data processing, starting quality assurance checks as soon as the first data comes in order to identify issues as soon as possible; and
- take explicit account of all lessons learnt from 2011.

The principles above have been shared and agreed with other Census Offices across the UK.

### **3. Quality Assurance End-to-End Journey**

The publication of census results is the culmination of a series of processes, from developing the census questions, collecting and processing the data, to the production and dissemination of statistical outputs. Each of these processes could introduce some degree of error and therefore, in ensuring quality of the results, it is important to consider quality assurance at every stage of the census process.

Our approach to quality assuring each of these processes is to:

- understand each process and the nature of any error or uncertainty that might be introduced;
- estimate the expected size and impact of potential errors within key processes in order to prioritise our checks effectively; and
- where possible, define metrics that allow us to understand the size and nature of these errors during 2021 Census.

#### **3.1 Question Development**

The design and content of the 2021 Census has been shaped by:

- successes and lessons learned from the 2011 Census and 2011 Census Quality Survey;
- developments and lessons learned in international census taking;
- requirements from the user community about the types, quality, frequency and detail of outputs required;
- user testing of proposed new questions and those with significant changes from 2011;
- changes in technology, in particular, the opportunities offered by the internet;
- changes in the way the public interacts with government digitally, enabling a move away from a traditional paper-based census;
- improvements in administrative data sources, giving the potential for their increased use in the production of official statistics; and
- the continued and ongoing need to make the most effective use of public money.

### 3.1.1 Consultation

The 2021 Census Topic Consultation for Northern Ireland was carried out between September and December 2015. This consultation was NISRA's first step towards understanding what information users need from the 2021 Census.

The consultation paper presented NISRA's initial view on the topic content of the 2021 Census questionnaire and included some new topics for consideration by users. The aim of the consultation was to promote discussion and encourage the development of strong cases for topics to be included. The consultation was issued to over 1,400 key users who were identified by NISRA and promoted on the NISRA website and through social media channels to ensure that the general public were aware of the consultation.

Public information events were also held in both Omagh and Belfast. Feedback was received from the consultation on topics such as income, knowledge and proficiency in languages other than English, vacant properties, use of sign language and past membership of the UK armed forces.

A common approach across the UK was used to assess the responses to the Northern Ireland consultation exercise and NISRA published the findings of the Topic Consultation in a report entitled '2021 Census: Topic Consultation for Northern Ireland – Assessment

of the responses received' in August 2016. In addition, NISRA also published a further paper providing evidence for the exclusion of a question on UK Armed Forces veterans.<sup>8</sup>

### 3.1.2 Testing and Research

Following on from the topic consultation, NISRA carried out a programme of testing and research to better understand the public's perception of the census, question wording, questionnaire design and response channels (online and on paper). This testing included public acceptability testing, cognitive testing of question wording and response categories, and testing of questionnaire design. In addition, NISRA conducted qualitative testing in 2017 to develop knowledge of the best local terminology around various census topics. The primary aim of this important work were to ensure that the Census questions would be readily understood by the general public and would be easy to complete. As the questionnaire was developed, later versions of specific questions were tested with members of the public.

### 3.1.3 2017 Census Test

As part of this round of testing NISRA conducted a large-scale, voluntary census test in Autumn 2017. The test involved sending a census-type questionnaire to a sample of 15,000 households, who were invited to respond online or on paper. The main objective of the Census Test was to inform decisions about the possible inclusion of a sexual orientation question in the 2021 Census. A secondary objective was to test the structure of the knowledge of Irish and Ulster-Scots questions. The test was also used to trial other potential new questions and test operational procedures.

An evaluation question was included on the form which asked, at household level, if any questions were difficult to complete. The majority of households who answered this question replied 'None' confirming that most respondents understood the questions being asked. The results were evaluated and the [2017 Census Test Report](#) was produced and published.

### 3.1.4 Topic Expert Meetings

NISRA established a range of topic expert groups throughout 2016 and 2017. A total of six groups were set up to cover the themes of:

---

<sup>8</sup> To review these documents please refer to [2021 Census - Topic Consultation Documents](#)

- Housing;
- Demography and Migration;
- Ethnic group, Identity, Language and Religion;
- Health;
- Education; and
- Labour market and Travel to place of work/study.

Topic expert groups were comprised of customers for census information from a range of Government Departments (including NISRA statisticians and policy colleagues) together with representatives from public bodies such as the Northern Ireland Housing Executive, Equality Commission and the Public Health Agency, together with leading academics.

### **3.1.5 2019 Census Rehearsal**

As part of the preparations for the 2021 Census, NISRA held a census rehearsal of 19,000 households located in selected areas of Northern Ireland. A full end to end rehearsal was conducted to ensure that all processes, systems and services were effective and working as they should before the 2021 Census. The primary objectives of the 2019 Census Rehearsal included assessing the effectiveness of proposed questions, particularly new questions and those that have changed significantly from the 2011 Census. The results of the rehearsal were published in the [2019 Census Rehearsal Evaluation Report](#).

## **3.2 Questionnaire Design**

The majority of households responded to the 2011 Census by completing a paper questionnaire, however, it is anticipated that at least 60% of those who respond to the 2021 Census will do so online.

The aims of an online-first approach are to:

- improve data quality;
- ease respondent burden;
- enable responses to be processed faster; and
- reduce costs.

The electronic questionnaire has been developed according to Government Digital Service (GDS) standards<sup>9</sup> and as part of the work to ensure the required standards were met, NISRA assessed the electronic questionnaire used for the 2017 Census Test.

User experience testing was also conducted with participants purposively selected to cover a wide range of ages and digital abilities, and included participants with physical and mental health conditions or illnesses. The research took place using a range of devices and assistive technologies and focused on:

- problems preventing successful completion of the online questionnaire;
- respondent ability to navigate around the questionnaire;
- respondent understanding of the questions; and
- overall respondent burden.

Design of the electronic questionnaire has included:

- responsive design feature that will adapt the layout of questions and web pages for different devices accordingly;
- section and information pages to describe what the respondent will be asked and what information they will need in order to answer the questions to follow;
- hub pages which show a list of questionnaire sections and their completion status;
- the main question will appear first and may be accompanied by additional information such as a question description, include panel or accordion guidance.

Where possible, NISRA have sought to minimise the amount of additional information shown on screen for each question. However, testing has shown that some additional information can help respondents provide more accurate answers, leading to better-quality data:

- the content of the question can change based on previous responses. For example, the first line of the household address might appear in questions about the household;

---

<sup>9</sup> Further information can be found on the GOV.UK [Service Standard](#) webpage.

- response options will appear as interactive elements and fall into one of three categories
  - Tick-box: respondents can select one or more responses;
  - Radio button: respondents can only select one possible response; or
  - Write-in: respondents can provide a written response
- summary pages which allow respondents to review their responses and make any necessary changes;
- clear navigation which allows respondents to easily progress to the next question, review previous answers and return to the hub page;
- online help web page will provide a comprehensive source of census information and guidance;
- respondents will have the option to switch between Irish, Ulster-Scots and English language forms at any stage of the electronic questionnaire;
- automatic routing will allow us to present respondents with only the questions that are relevant to them depending on the answers that were previously given;
- automatic text fill is the process of using a respondent's previous responses, such as a name or address, elsewhere in the questionnaire. This reduces respondent burden by making questions more clear and reducing the number of times that respondents need to enter the same information;
- search-as-you-type functionality will be available for a number of questions. When a respondent starts typing a write-in response, a list of suggested answers will appear. There will remain the option to write in a response that does not appear as a suggested answer, if the respondent chooses to do so;
- where two or more responses to a single question cannot both be true at the same time, the electronic questionnaire includes functionality to de-select previously ticked responses. This functionality removes the possibility of a multi-tick error when processing the online returns;
- responses are saved after every question. If a respondent closes or logs out of a partially completed questionnaire to complete it at another time, their responses will not be lost; and

- validating responses through built-in checks to improve the quality of the data collected. This is done using error messages and additional questions.

Where necessary, a paper questionnaire will be provided to the public. The paper questionnaire has been designed to be easy to fill out with space for up to six persons in each household.

The starting point for the design of the paper questionnaires was the design used in the 2011 Census and will again take into account good practice standards and guidance. A major challenge in designing the paper questionnaire was the limitation in space on each page, however the 2021 Census questionnaire will be no longer, in terms of questions, than the 2011 Census questionnaire.

The design of the paper questionnaire has included:

- incorporation of pre-addressed information and pre-paid return envelopes to support the data collection process;
- changes to the colours on the paper questionnaire to improve readability and accessibility of questions, guidance notes and routing;
- where appropriate, the order of the questions aligns with that used for the 2011 Census to ensure consistency;
- optimal sizes for free-text write-in boxes to ensure quality data is captured;
- availability of the paper questionnaire and explanatory material in Braille and large print; and
- a number of translation booklets available to those respondents whose first language is not English, to assist with the completion of their paper questionnaire.

The design of the questionnaires was tested during the 2019 Census Rehearsal and it was found that the quality of the data collected through the online questionnaire was a significant improvement from the 2011 Census. As a result of the extensive research and testing detailed above, NISRA published the [2021 Census Household Questionnaire](#) in August 2020.

### 3.3 Census Address Register

The development of a robust Census Address Register (CAR) is essential in counting every household and communal establishment in the 2021 Census. A fit-for-purpose CAR is needed to:

- provide the basis for contacting the public to invite them to participate in the census;
- be the source for the system to monitor responses made to the census; and
- inform the field operation during the follow-up of non-responding addresses.

As done previously for the 2011 Census, the main source for the 2021 CAR is the Land & Property Services (LPS) POINTER<sup>10</sup> database, supplemented with additional address level datasets available through government or within administrative/utilities data (e.g. Royal Mail Postal Address File, LPS Domestic Valuation listing and the Northern Ireland Housing Executive Social Housing list).

The development of the address register for the 2021 Census builds on lessons learned from the 2011 Census and recent tests/rehearsals undertaken in the interim period. The aim for any address register is to maximise coverage and accuracy as much as possible. Having an accurate register minimises the risk of sending letters and/or questionnaires to addresses that don't exist, or sending them inadvertently to the wrong address. Having a register that has high coverage reduces the risk of people not being invited to take part in the census.

#### 3.3.1 2011 Experience

In 2011 the overarching aim of the CAR was to maximise accuracy in the initial desk-based development of the register, with the coverage being maximised with a 100 per cent on-the-ground address check just prior to the delivery phase of the census. Results from the 2011 Census suggested that under this approach, there was still 3 per cent of addresses found in the operation that were previously not known to NISRA. In terms of

---

<sup>10</sup> POINTER is the address database for Northern Ireland which is maintained by Land and Property Services with help from local councils and Royal Mail. Information on this dataset can be found on the [nidirect](http://nidirect) website.



accuracy, circa 4 per cent of addresses in 2011 were not delivered by the postal service provider.

### **3.3.2 2017 Address Check**

In September 2017 NISRA performed an on-the-ground address check of approximately 16,000 addresses to test the address register at that time. The results showed that there were around 1.5 per cent of addresses that were previously unknown to NISRA. This sample showed that when compared to the experience in the 2011 Census, addressing coverage seemed to be getting better – but it was still not good enough.

### **3.3.3 2019 Census Rehearsal**

During the 2019 Census Rehearsal, NISRA carried out a full on-the-ground check of approximately 19,000 addresses. Overall, the coverage of addresses was excellent and the address register was 97% accurate (an improvement from the overall level in 2011). During the rehearsal, field staff found some additional addresses that were previously not known to NISRA (about 0.6%). In addition, a test was performed in which a number of addresses were deliberately left out of the address register prior to live operations – only half of those addresses were found by field enumerators, casting doubt on the benefit of having field staff look for addresses.

### **3.3.4 Development of the 2021 Census Address Register**

As described, the development of the 2021 CAR was based on the LPS POINTER product supplemented by other address centric datasets. NISRA has invested significant resource into improving understanding of the POINTER address database and has worked closely with LPS to improve data quality. Regular meetings are conducted between the organisations and a monthly POINTER quality report is produced by NISRA and shared with LPS.

In addition, NISRA has done a lot of work to link POINTER address information to administrative datasets to get a better indication of which addresses in POINTER might be occupiable and domestic – and therefore valid for a census enumeration. This approach has allowed NISRA to make huge inroads into the selection process for addresses to be included in the CAR. Furthermore, this increased access to administrative datasets, coupled with the availability of robust GIS data and software,

meant that NISRA could perform much more verification and validation of address records using a desk-based approach rather than having to rely on expensive on-the-ground operations.

Taking into consideration the improvements to data and systems outlined, and the experience gained in the 2019 Rehearsal which cast doubt on the on-the-ground address checking, NISRA was already considering the merits of a 100 per cent field address check for the 2021 Census, but when the COVID-19 pandemic struck it became much more risky to plan for such an exercise. As a result, a decision was taken to not perform an on-the-ground address check in advance of the 2021 Census in favour of a fully desk-based approach.

The revised plans focussed on checking addresses using Geographic Information Systems (GIS) as a vehicle for providing checkers with both address and spatial information to check the validity of addresses. This was completed in two stages. Firstly, checkers were provided with a list of around 55,000 addresses which the CAR developers had insufficient information to verify whether they should be included in the census or not. Secondly, every enumeration district in Northern Ireland was checked to make sure that coverage had been maximised and that there were no pockets of addresses where NISRA didn't have records.

A cautious approach to desk-based checking was adopted, in recognition of the fact that the data and maps could potentially be out of date in some areas. As such, a decision was taken to maximise coverage over accuracy to ensure nobody was missed from the census – any issues with accuracy could be “fixed” by field enumerators in the field during the planned follow-up phase of the census enumeration using the money saved by not conducting a full on-the-ground address check.

After address checking there were still some areas where the address coverage was not clear. These related mostly to areas where development had been taking place or was planned to take place but the address datasets had not yet caught up. In these areas NISRA utilised the Census Area Managers (CAMs) to work with local estate agents to try

to ascertain what addresses were valid from a census perspective – these were fed through to the CAR in advance of finalising the dataset.

The development work above resulted in 843,000 household addresses and 1,500 Communal Establishment addresses being identified for inclusion in the 2021 Census. During live operations a Response Management (RM) system will operate to track responses and manage fulfilment requests. Any addresses not known to NISRA will be added to the register if deemed valid. Likewise, any addresses that are not valid will be removed. All address validation (whether it be additions, changes or removals) will come through a central Address Resolution service which will operate to ensure that each change is validated and that each address is postally accurate. To facilitate this Address Resolution service, the CAR development team will continue to update their addressing databases using the approaches adopted for the development of the initial CAR to ensure the most up to date information is available and reflected in the RM. This will mitigate against the risk of people not being included in the census due to the time lag between CAR delivery and live Census operations (December 2020 to March/April 2020).

### **3.4 Data Collection**

The challenge for the data collection phase will be to secure information on a population which is less tolerant than ever of official surveys. Indeed, experience from the 2011 Census and in official surveys since then has indicated that it is getting more difficult to make contact and engage householders across Northern Ireland.

To help manage some of these issues NISRA plan to make greater use of digital technology to improve the quality and efficiency of the census data collection operation.

The effectiveness of data collection in 2021 will be measured by the overall response rate which will be monitored throughout the census and evaluated.

#### **3.4.1 Initial Contact**

International research suggests that a postal first contact is the best and most cost-effective solution to ensuring maximum coverage, therefore the primary innovation for the

2021 Census is that most households will initially be sent a letter containing a Unique Code rather than a paper questionnaire. This code will allow the recipient to complete their census questionnaire online. A further postcard will be issued prior to Census Day to all householders reminding them of the need to complete their census questionnaire.

Those unable or unwilling to complete the questionnaire online can request a paper version which will contain a unique barcode associated with that particular address to facilitate the delivery and tracking of the questionnaire. A pre-paid return envelope will be included to allow safe return of the completed form.

Initial contact will also be tailored to maximise participation. Around 20% of all households will be issued a paper questionnaire instead of a letter, for example, those areas with poor broadband who may have difficulty completing the online form.

During the 2017 Census Test, a census-type questionnaire was sent to a sample of 15,000 households, who were asked to respond online or on paper. An initial invite letter, which included a Unique Online Access Code, was sent to each household. The Access Code gave respondents access to their online questionnaire. Two weeks later, all non-responding households were sent a reminder letter and then a further two weeks later, those households who had still not responded were sent a final reminder letter. This final reminder contained a copy of the paper questionnaire. An additional question was included on the paper questionnaire which asked why respondents decided to use a paper questionnaire rather than online. The most popular reason was “I prefer paper” (50%) followed by “do not have a computer/internet connection” (29%) and “I am not confident in using the internet” (20%).

The method of initial contact was also tested during the 2019 Census Rehearsal. Follow-up interviews with some residents highlighted that some of those who received initial invitations to complete their questionnaire online did not know that they could request a paper questionnaire. Also, the response rate amongst students remained low and further consideration is needed on how best to enumerate this group. Overall, the rehearsal

showed that the method of initial contact worked well but some issues need addressed for the 2021 Census.

To ensure maximum participation in the 2021 Census a dedicated team will be established to capture the characteristics of those living in approximately 1,500 communal establishments such as hotels, hospitals, boarding schools or prisons. This team will work with the managers of the communal establishments to ensure a response is made. Additional procedures will also adopted for some special populations. This will include people with no settled place of residence and persons sleeping rough. These special enumeration measures will be informed through consultation with representative organisations.

### **3.4.2 Publicity and Support**

The 2019 Census Rehearsal found that stakeholder engagement was more successful when it was face to face and some networking events took place. Therefore in 2021, a more detailed plan for contacting and communicating with stakeholders and community groups is needed and where possible, the communications will be aimed at named individuals rather than generic email addresses. A suite of products will be designed with messages tailored to specific 'hard to reach' groups which address specific concerns, language or accessibility needs. Early engagement will be key to ensure that as many potential relationships as possible can be built.

The census rehearsal media campaign included radio advertising, print in regional newspapers in the rehearsal areas, outdoor advertising and digital programmatic advertisements on Google search, Belfast Live, Facebook, Instagram, etc... It was estimated that 53% of the NI adult population were exposed to the campaign at least once.

During the rehearsal a Contact Centre was also in operation. The operators answered calls, held webchats and answered social media queries. In addition Interactive Voice Response (IVR) was available continually (24/7) with an automated paper questionnaire request service. In total, over 1,000 contacts were received. The contact centre was found to have worked well and the quality of the training material provided meant that

very few calls were escalated to census HQ. The online help was found to be a useful source of information both for the contact centre staff and the public and many of the social media posts were shared by stakeholders with a high level of positivity.

NISRA plan to expand upon the success of the rehearsal and run a large-scale publicity campaign to allay any concerns and motivate householders to take part in the 2021 Census, whilst also giving the public access to information on all options for support. NISRA will also work with local Councils and key organisations such as Age NI, the Royal National Institute for the Blind and the Simon Community in order to improve awareness and coverage amongst communities, minority populations and special interest groups. Assistance will again be available online and analytics will be used to help review website traffic and gain a better understanding of how best to engage the public.

### **3.4.3 Field Force Strategy**

During the census rehearsal, a wide variety of activities were carried out to enable a team of field staff to follow-up with those in the rehearsal areas who had not yet responded. The data collected showed a relatively low response rate to door knocks from enumerators (25%). There were also some minor signal/coverage issues, however, the application used allowed enumerators to access their worklist of non-responding households on their mobile device even when offline. The data on the application was stored securely at all times and once signal coverage resumed the field staff were able to return their updates to response management. This enabled the follow-up to continue without requiring continuous signal coverage.

For the 2021 Census, given the poor contact rate on the door step, different initiatives will be investigated to determine the best time to carry out field visits, ways of identifying field staff more clearly, better use of community and social media groups to let people know enumerators are in their area. Student areas in particular had a low response rate, so Census Office are investigating bringing forward follow-up in 2021 to try to contact students before the Easter break. Field staff will also be fully trained to assist and encourage the public to complete their census questionnaire.

### 3.5 Data Capture, Coding and Transfer

It is anticipated that most respondents will wish to complete their census questionnaire online, although those respondents preferring to complete a paper questionnaire will be able to do so. Completed paper questionnaires will be scanned and the data captured, coded and amalgamated with the online questionnaires into a single dataset. This will be outsourced to a supplier with the final dataset being transferred by the Office for National Statistics (ONS) to NISRA.

These elements were tested ahead of the 2019 Census Rehearsal. Given the number of questions and various browsers and devices that could be used to access the online questionnaire, this work took a significant amount of time and staff to test. Unfortunately the timetable of deliverables for online questionnaire testing was delayed and therefore not all the minor issues discovered in testing were resolved prior to 'go live'. Operational pressures also meant that data types and data dictionaries were not fully specified before the data was sent and this caused issues with data being in the wrong format. This was solvable in live but again more testing is required. Also, given time constraints, several areas including receipt of incremental updates, image files, response management feeds and diagnostic reports were not fully rehearsed.

For the 2021 Census there will be a number of approaches used to ensure the captured and coded data from questionnaires is of the highest quality:

- the supplier will be required to meet specified data security protocols and quality targets;
- samples of data will be checked for accuracy by Census staff;
- specific checks on identified risks, for example, 'major' in job title being interpreted as military officer incorrectly; and
- further and rigorous testing using test questionnaires ahead of the 2021 Census.

There will also be an established process for checking the data, both prior to transfer from ONS and following receipt of data by NISRA. This will ensure that the data satisfies the specified requirements and will include checks such as:

- variable range checks;
- variable format checks;
- analysis of variable distributions to identify unusual patterns;
- number of records being transferred;
- distribution of records by response mode; and
- geographical distribution of responses.

### 3.6 Data Processing

For the 2021 Census NISRA will conduct processing of all Northern Ireland census returns locally for the first time. Data processing performs a number of actions on the census response data focussing on making it fit-for-purpose for census outputs. As has been the case since 2001, the 2021 Census processing system will consolidate the data, identify and remove duplicates, impute missing data and insert person and household responses to account for under-coverage in the census. Fundamentally NISRA plans to run a processing system that applies methodologies tried and tested in the 2011 Census. The sections below highlight some of the main processes and NISRA's proposed plans for how to measure/ensure quality in each of them.

Invariably, the census is affected by under coverage and erroneous responses which presents challenges in terms of getting the enumeration as high and complete as possible, implementing statistically robust methodologies to estimate and adjust for undercount, and generating a complete and consistent final dataset. It is therefore necessary to have processes to measure and manage quality during data processing, which help inform the communication of uncertainty and limitations to users of census data.

Therefore, each step of data processing will require bespoke quality assurance checks which will include validating that the process performed correctly and documenting the results. NISRA are still developing the plans for the quality assurance of these processes and as such they are subject to change.



### 3.6.1 Removal of False Persons and Resolution of Multiple Responses

The Removal of False Persons (RFP) process targets census returns which do not contain sufficient information to be treated as a valid response. The Resolution of Multiple Responses (RMR) process seeks to identify incorrect census returns which relate to a person already correctly covered on another return from the same or very close location and either remove those responses or consolidate them into relevant households.

Quality assurance activities may include:

- use of record level matching to administrative data records to effectively identify and examine potential false persons and multiple responses;
- assessment of the false person responses for validity or otherwise;
- clerical checking of those flagged as multiple responses; and
- analysis of the effects of merging records.

A process will be applied after RFP and RMR to validate the database changes. This post-RM validation will ensure database consistency by looking at person counts within households against expectations to check that the processes have run as planned.

### 3.6.2 Addition of Census-Under-Enumeration Records

As was done in the 2011 Census, the 2021 Census will use administrative data records if required to insert missing people into the census dataset to account for non-response. Again, as with the 2011 Census, this will only be done for completely non-responding households where NISRA believes that a response should have been received (known as the CUE project in 2011). This time the CUE process will be facilitated by the availability of a greater range of administrative data, linked to provide a strong source of information on the size and demographics of non-responding households.

Quality assurance of the CUE Project in 2021 shall include;

- verifying the CUE process and validity of the administrative data by checking a sample of responding households to compare the information between census returns and the administrative data; and
- analysis of the CUE records to ensure that the addition of households and person records does not introduce bias.

### 3.6.3 Edit and Imputation

Edit and imputation is the process of imputing for a missing or clearly incorrect value on a census return, and is conducted in the census using a donor-imputation methodology.

Quality assurance activities may include:

- distribution checks before and after every step to ensure distributions are preserved and bias has not been introduced;
- checks to ensure overuse of donor records, and resulting 'spikes' in the data, have not taken place; and
- comparing imputed values against those collected in other administrative data sources.

### 3.6.4 Coverage Estimation and Adjustment

Coverage Estimation and Adjustment are the processes by which census and Census Coverage Survey (CCS) information is used to estimate and adjust for the people and households missed by the census. This is over and above the CUE process. Once estimated, the records required to adjust for under-enumeration are then added to the census data with minimal socio-demographic information. Records are subsequently fully populated using the Edit and Imputation process.

Owing to the complexity of Coverage Estimation and Adjustment and the number of different elements in the process, it will be essential to have an extensive and effective quality assurance mechanism in place. Proposed quality assurance measures include:

- comparison of age-sex estimates at different levels of geography with those pre-estimation and also with comparator sources such as the mid-year population estimates and the administrative data sources;
- analysis of sex ratios in the generated estimate to identify any anomalies;
- iterative re-sampling (Bootstrapping) which will inform the calculation of confidence intervals to be reported alongside the final estimates; and
- examination of the distribution of generated households by size, geographical areas and key variables, for example, age, sex and religion.

### 3.6.5 Validation of Population and Household Estimates

The main aim of these quality assurance checks will be to ensure that the census estimates are plausible at both the Northern Ireland and smaller levels of geography.

The primary focus will be the comparison of population and household estimates with a variety of comparator datasets and will include examination of:

- households with usual residents in terms of the overall number, the distribution by household size and the average household size;
- the number of usual residents living in households, communal establishments and in each of the main population sub-groups (for example, children, working age, pension age, students);
- the distribution of usual residents by age and sex;
- the inferred sex ratio distribution for usual residents by age (that is, the number of males per 100 females); and
- the inferred fertility and mortality rates for usual residents when these are calculated using the 2021 Census population estimates.

In view of the fact that the 2021 Census results will not align precisely with the comparator datasets, appropriate tolerances will be developed for each comparison to specify where the results should fall in order for the census information to be considered an acceptable quality. The evidence to support our population and household estimates will also be reviewed by internal and external Assurance Panels who will advise on whether estimates are fit-for-purpose or require further work or adjustment.

### 3.6.6 Topic Analysis

Topic analysis examines the census results relating to each topic covered in the census (for example, housing or health) to check for any evidence of systematic errors in the results for that or related topics, and to understand, as far as possible, the reasons for any difference between the census results and other sources of information on that topic. NISRA plan to use comparator data sources and topic experts to quality assure each topic in order to assess the stories shown by the provisional census results in the context of other evidence and trends. Comparison to the results from the 2011 Census will be

made for those questions which are broadly comparable. In respect of the new questions being introduced, work will also be undertaken to ensure that the distributions look plausible when compared against relevant information from sample surveys and/or administrative data sources.

### **3.6.7 Assignment of Output Geographies**

There is an anticipated demand to produce 2021 Census outputs for the current Electoral Wards. However, a review of local government boundaries is impending, which will potentially result in a change to the Ward boundaries. NISRA therefore require a flexible method for producing 2021 Census statistics at a range of different geographical levels, while adhering to disclosure control requirements.

A number of options that are currently being considered by NISRA which include:

- assessing the areas most affected by population growth since 2011 and identifying options for amending their boundaries; and
- investigating the most effective method for aggregating census data at grid-square level to produce statistics for different administrative geographies.

NISRA is currently conducting research into both the options listed above with the aim to publish a paper outlining the proposed approach in due course. Any method adopted will be subject to rigorous quality assurance measures.

## **3.7 Statistical Disclosure Control**

Statistical Disclosure Control (SDC) techniques are applied in order to prevent the identification of individuals and households, and their attributes, in published census outputs. In line with other Census Offices across the UK, NISRA are considering the application of two forms of SDC<sup>11</sup>:

1. Targeted record swapping which involves swapping the geographical locations of records and which was successfully applied to the 2011 Census data; and
2. Cell Key Perturbation, a new technique which introduces small adjustments to cells within output tables.

---

<sup>11</sup> Further information can be found in [Statistical Disclosure Control Methodology for 2021 Census](#).

Quality assurance measures will ensure these methods are applied effectively without altering the representativeness of the data and are likely to include:

- estimating the expected scale of error introduced by record swapping, by looking at the evidence from 2011 in conjunction with any change to the parameters and rules being applied in this process;
- estimating the expected scale of error introduced by cell key perturbation from first principles – that is, derived mathematically from the properties of the perturbation algorithm; and
- comparison of adjusted figures with the unadjusted figures.

### 3.8 Output Production and Dissemination

The 2021 Census Outputs Strategy Consultation for Northern Ireland was conducted between October 2018 and January 2019 and sought views from users on the proposed strategy for producing and disseminating 2021 Census results<sup>12</sup>.

Quality assurance of 2021 Census outputs will be thorough and shall include:

- checks of tabulations against the census database to ensure they are consistent and have been correctly specified;
- comparison of tabulations where the populations are the same to ensure they are consistent; and
- ensuring the tabulations are non-disclosive and do not identify any person or household.

### 3.9 Other Important Aspects of the Quality Assurance Work

NISRA plan to start the Northern Ireland Census Coverage Survey (CCS) on 12<sup>th</sup> May 2021. The CCS sample will include approximately 15,000 households and use a subset of census questions to collect basic demographic characteristics (such as age, sex, marital status, religion and economic activity). As response to the CCS is voluntary, maximising response rate has been a strategic aim during planning and development of the 2021 CCS. Evaluation of the CCS in 2011 found that the quality assurance processes

---

<sup>12</sup> Details of the consultation can be found on the [Outputs Strategy Consultation](#) webpage

applicable to the census are equally applicable to the CCS. In light of this, NISRA will be providing more detail on the quality assurance processes for the CCS in due course.

NISRA are currently considering additional methods for quality assuring the 2021 Census including a voluntary survey, similar to the [2011 Census Quality Survey in Northern Ireland](#).

## 4. Developing the Quality Assurance Approach

NISRA will seek the views of stakeholders in the development of our proposals for the quality assurance of the 2021 Census. This will ensure that our approach reflects best practice and includes, where possible, checks that stakeholders would plan to conduct. It will also help us understand what evidence users require to understand the reliability of the census results.

In particular NISRA will seek feedback through:

- Census Oversight Board (COB) – which is chaired by the Director of Census and Population Statistics and attended by the NISRA Census Office Senior Management Team to consider operational, statistical and project management issues;
- Demographic Statistics Advisory Group (DSAG) - census users and experts from academia and business which has a formal advisory role in the content and delivery of the census in Northern Ireland;
- Census Assurance Panel (CAP) - expert users of census information and specialists from other parts of the NI Civil Service provide advice and guidance to Census Office in key areas such as human resources, finance and procurement;
- Census Research Assurance Group (CRAG) – a group of methodological and demographic specialists from across the UK statistical agencies which provide advice and assurance on methodology and evaluate items in advance of external review;
- Using methodological expertise within the NISRA Statistical Support Branch;
- External expert peer review; and

- drawing on experiences and approaches of other census-taking organisations and participating in a UK Census Quality Assurance Harmonisation Working Group along with ONS and NRS.

The quality assurance strategy will be informed by the comments and suggestions NISRA receive from users.