National Records of Scotland

Beyond 2011

Stakeholder Engagement workshops
Combined write-ups from events
November 2012 – March 2013

Published on 4 March 2014

Executive summary

This report provides an analysis of the discussions held with our stakeholders between November 2012 and March 2013 on the Beyond 2011 programme. The Beyond 2011 Programme in Scotland was established by National Records of Scotland (NRS) in September 2011 to explore future options for producing population and socio-demographic statistics that best meet user needs in Scotland. The programme is investigating a range of possible solutions, including the possibility of using administrative data sources or developing a more efficient census design.

Events were held across Scotland through a combination of facilitated events, one-to-one discussions and videoconferencing. There was also a series of equality focused events in March 2013.

These discussions were held before the formal consultation on user requirements opened at the end of March 2013 and were used as a ‘warm-up’ exercise to the consultation.

Uses of the Census

The following sections highlight the views expressed by stakeholders at the events.

It is clear that census data is hugely important to many aspects of delegates work. Census filters through to all sectors of local and central government through areas such as service planning, transport, housing, equalities monitoring and Single Outcome Agreements.

Census is also used widely across the country by academics, voluntary groups, charities and religious communities. It is also used as a way to bring revenue to the country through Historical tourism.

Census is seen as the benchmark for other sources of information and is seen as the ‘gold standard’ for statistics.

However, the 10 year frequency and the long release schedule are seen as issue that had more impact as the time elapsed from a census increased.

Alternative data sources used by stakeholders

A wide range of other data sources are used in delegates areas, including surveys, administrative data and management information systems.

Delegates use a range of government based surveys, including the Scottish Household Survey and the Scottish Health Survey. Delegates also mentioned the use of local or ad-hoc surveys where required but noted the expense in carrying out small scale specific surveys. Delegates explained that these are generally used to supplement the census, rather than necessarily as an alternative.

Delegates also used administrative data sources in their work. Administrative data is derived from information collected and maintained as part of an administration system, such as health records, vehicle licensing and tax systems.

The merits of a number of alternative data sources were discussed in more detail, particularly in providing information on hard to reach groups.

Some delegates were also concerned that many of the alternative and supplementary sources that they currently use have components that are derived from the census (for example, Scottish Neighbourhood Statistics and the Scottish Index of Multiple Deprivation) and were keen to emphasise how useful these sources were.

Management Information Systems in use were related to delegates specific areas including, Education management information software (SEEMiS), Social work information system (SWIFT and local authority Customer Relationship Management (CRM).

Implications of not having a Census

Delegates discussed what the main implications would be if the census outputs were not available. For many delegates the census is essential to service provision and to understanding their performance. Some delegates expressed concerns that if they did not have the census they would, ultimately, need to conduct or commission their own surveys to ensure adequate coverage at lower geographic levels. This could have significant financial implications.
Views on Future Models

Delegates were given an opportunity to discuss the different possible options for producing population and demographic statistics in future. These included a modernised online census and using administrative data combined with data linkage. There was a general view that we should be learning from international methods for producing population and demographic statistics.

Administrative data

At the events delegates discussed the merits and problems of using administrative data sources and data linkage. For some delegates this was based on their own experience of using administrative data and data linkage projects. Opinions on using alternative sources were generally positive, with delegates agreeing that there is huge potential for them to be used in future.

However, delegates noted that there were a range of issues with administrative data including legal, ethical and privacy concerns. There was also the issue that administrative data, by its definition, is collected for purposes other than statistical collection and therefore could be dependent on changes to collection processes, e.g. a policy change. There could also be issues with comparability across different administrative data sources as well as issues of data quality.

Online Census

Delegates were, in general, extremely positive about the potential of the option to develop the online completion of the census. However there was also concern over the impact of changing to a web-based format and the need to ensure returns are accurate and response rates are not adversely affected, especially in specific sub-groups of the population.

Future Requirements

Delegates were asked to discuss the following aspects of the information they use and their relative importance:

- Quality and accuracy
- Geographic coverage
- Frequency
- Aggregation

At the events the common view on the relative importance of the key aspects was that this depended on what use of the census or variables were being discussed. For example, the view for housing may be very different for the view on education. It was also stressed that this depended on local outcomes and needs.

However many delegates noted that an ideal solution would be to produce more frequent statistics at the current levels of accuracy and geographic coverage.

Equalities

Views from the equalities specific events were generally the same as from the main events but delegates stressed the importance of complying with the specific duties within the Public Sector Equality Duty (PDF document (242 Kb) available on the Press for Change website).

Delegates expressed concerns that an administrative based model with a sample would not allow the current level of reporting on equalities indicators and could mean that service providers and charity groups would not be able to fulfil their legal obligations

Conclusions

The evidence collected from the stakeholder engagement events will be added to the body of evidence about what users require for producing population and demographic statistics

The views collected here compliment the findings from the formal user consultation event held after these events.

Contents

1. Introduction

This report provides an analysis of the discussions held with our stakeholders between November 2012 and March 2013 on the Beyond 2011 programme and contains a cross section of their views. The Beyond 2011 Programme in Scotland was established by National Records of Scotland (NRS) in September 2011 to explore future options for producing population and socio-demographic statistics that best meet user needs in Scotland. The programme is investigating a range of possible solutions, including the possibility of using administrative data sources or developing a more efficient census design.

Events were held across Scotland at the following events:

- Edinburgh 5, 8 and 11 March 2013
- Edinburgh 19 and 21 February 2013
- Glasgow 7 and 8 February 2013
- Aberdeen 30 January 2013
- Inverness 29 January 2013
- The Islands 29 January and 11 February 2013
- Perth 24 January 2013
- Edinburgh 15 January 2013

Each event included a presentation about the background to the Beyond 2011 Programme, which explained the main reasons for the programme and the main options for the future. Delegates at events were invited to discuss how they currently use census information and how any changes could affect their work.

The discussions also gathered users' views on their requirements for population and socio-demographic statistics including:

- the frequency of data provided;
- the geographical level at which data is available;
- the accuracy of the data; and
- the level of disaggregation required.

A series of equality specific events were held in Edinburgh on the 5th, 8th and 11th March 2013 and are included as a separate chapter in this report.

These discussions were held before the formal consultation on user requirements opened at the end of March 2013.

The formal consultation on user requirements was held mid 2013 and the report can be found in the Consulations section of the NRS website.

2. Current uses of Census data

The following section details how the census is currently used by the delegates. It was clear that census data is hugely important to all aspects of many delegates work.

> 'Almost everything I do relies on the census.'

2.1 Planning and Resource Allocation

Delegates use a range of aggregate statistics derived from the census to effectively plan services. Aggregate statistics are a collection of summary counts of the numbers of people, families or households resident in specific geographical areas possessing particular characteristics, or combinations of characteristics. These are drawn from the themes of population, including ethnicity and religion, health, work, and housing. They are derived from analysis of the information provided in returned census forms. Adjustments are made in the outputs for people and households that may have been missed out and not been recorded on a form, or that may have been counted more than once by being recorded on more than one form.

A number of the particular characteristics and combinations of characteristics drawn from the census, such as health and ethnicity, were mentioned by the delegates as being important to the planning of a range of services.

Considerations of equalities characteristics were seen as important for Local Authorities when producing and maintaining Single Outcome Agreements (SOAs). SOAs set out how Community Planning departments in councils will contribute to the Scottish Government's National Outcomes. Information for both of these is available on the Scottish Government website. The purpose of a SOA is to identify areas for improvement and deliver better outcomes for the people of Scotland as well as individual local authorities.

> 'Most data requests go to me and they are currently focused on assessing Single Output Agreements. Other alternative data sources such as SIMD[1] are not as useful as the census. For instance we run a multiple deprivation analysis with the census data and we could identify deprived areas within our communities at very low geographical level. One street within a supposedly affluent area can be very deprived and this is something the SIMD survey doesn't detect.'

Delegates stressed the importance of knowing the population profile of an area and how services and provisions are suited to the population. As part of this a number of delegates emphasised the importance of cross tabulations to planning services.

Cross tabulations summarise data based on two or more of the characteristics and can show the inter-relationship between two or more

Footnote

1) Scottish Index of Multiple Deprivation (SIMD) available on the Scottish Government website.

characteristics. For example, gender and ethnicity can be cross tabulated with employment information to produce figures on how many unemployed Pakistani women there were within a specified geographical area. Delegates agreed that reliable information on these issues can only be found in the census.

> ‘I also use [the census] to help local authorities understand their ageing population.’

> ‘We rely on the census heavily for information on ethnic groups to understand service needs and initiatives.’

> ‘I run the Community Planning Programme (CPP) and for us accurate data is key.’

> ‘I can’t think of a better source for things like religion, ethnicity, over-crowding and multiple-occupancy. You could not get any of this information from the private sector. It is unique.’

Delegates also mentioned that the census is currently the only source of information to provide data on the number of offshore workers, which is an important issue for the islands.

> ‘For offshore workers there is no other way to capture that information other than the census.’

2.2 Resource Allocation

The wide demographic coverage of the census was seen as an important feature by delegates as the census underpins many resource allocations.

> ‘The census covers everything. Other research and study gives an indication but not the full picture. The census can never be substituted. Resources could be provided and decisions made but wrongly if they are not as accurate as the census.’

> ‘The resource allocation [for religious services] is based on the population count and the religion question.’

Some delegates also discussed the financial implications of the census:

> ‘The census as a tool is very important. The user should be able to extract information from the census. If there is a tool for information it saves a lot of money. I think it may turn out that we have an outcome that says ‘even on the basis of cost the census is the best option’. I think money invested in the census saves money elsewhere’.

2.3 Housing

The delegates discussed how census data is cross tabulated to profile areas in order to understand and predict housing needs for service provisions. Delegates also felt that small geographical levels are integral to housing planning.

> 'We also use data on housing needs across the LA and the population distribution. We like to be able to use Isle specific population data.'

> 'We need population and migration statistics to understand community profiles and forecast the housing market.'

> 'We use the census to understand housing needs. It is particularly useful to understand the age profile to plan for housing needs such as care homes and to break down particular needs, such as disability. It also helps to provide measures to meet needs.'

> 'It is essential to carrying out needs and demands strategy. We rely on the census for information on tenure and the number of rooms within households.'

> 'I examine affordable housing, the private sector and fuel poverty among other areas. These all use information from the census.'

2.4 Transport

Some delegates used the census for transport planning and emphasised that this is currently the only source from which they can derive information on land use and transport.

> 'The data is used to model the interaction between land use and transport. We supply the outputs to local authorities on travel to work and local transport models. For example, we use census data to create models for the A9. It is essential to the modelling for us and local authorities.'

2.5 Language

A number of the delegates indicated that they were looking forward to the release of the 2011 Census data and that they would be using data from the question on language, especially Gaelic. Census data on language and ethnicity is also used in the provision of language classes and community groups. The census was seen by delegates as currently the only reliable source of information on language.

> 'The Gaelic question is unique to the census and is the only data source that asks the whole population their language ability. It gives you age breakdown such as 5 to 12 year old Gaelic speakers'.

> 'Sign language users in Scotland still need the census as a benchmark'.

‘The information on language helps us to understand the demand for language classes’.

‘I use population estimates for things like helping education departments design new schools. Knowing the demand for Gaelic education is important. Travel to work areas are very useful too’.

2.6 Health and Disability

Some delegates use General Health questions which are ‘fundamental to providing healthy life predictions’ to plan services. General Health questions ask the population how their health is in general and if they have any disabilities which have lasted or are expected to last over 12 months[2]. Health can be cross tabulated with characteristics on ethnicity, for example, to understand health inequalities in an area.

‘We need census data to understand how many people will have needs in terms of learning disabilities and carers. We are very interested in seeing the results from the 2011 Census on this’.

However, other delegates noted that the accuracy of personal assessments of health is questionable and there may be better ways of assessing public health.

2.7 Equalities Duties

The public sector equalities duty (PDF document (242 Kb) is available on the Press for Change website) arising from the Equalities Act 2010 (legislation.gov.uk website) also requires all public bodies to make assessments to ensure they are not discriminating against anyone with a protected characteristic (i.e. age, gender, disability, ethnicity, sexual orientation, religion and belief, etc). As discussed above, the census provides local authorities and NHS Boards with baseline information on these characteristics to monitor discrimination. All delegates stressed the importance of the census as being the primary source of equality information, particularly relating to ethnicity, disability (self-diagnosed), and religion.

‘There is no other source for equalities data. I have been relying on census 2001 data even though it is now well out of date.’

‘We use the census for population based analysis to identify specific groups, in particular ethnicity, with the aim to target services. For instance, we've got a big Polish population and we'd like them to access the services they need. We use a lot the census output area data.’

Footnote

2) These questions were asked in the 2001 Census questionnaire (PDF document available on the NRS website) and then again in the 2011 Census one (PDF document available on the Scotland’s Census website) along with an additional question about having a disability expected to last over 12 months with disability categories, including mental health, physical disability and learning disability.

> ‘An Arab question was included for the first time in 2011. This will have a huge impact for the Arab community. Up until now the desire for this was variable.’

Aggregate statistics on religion are also used in the planning of religious services:

> ‘We use nation-wide figures to make sure we have the correct number of [Church of Scotland] ministers. We feed information to congregations, for example, what type of people, the numbers and how affluent or deprived an area is.’

2.8 Geographical Coverage

Aggregate statistics derived from the census are available from data zones to a national level. This flexibility to the statistics was considered very important to the effective planning and targeting of resources. The delegates emphasised the importance of small area data to their work.

SOAs require data below Local Authority level to help identify areas for improvement within local authorities. Low geographical levels of data can be used to aggregate up to the level that is required to build data zones. Delegates stated that small geographies were very useful to Local Authorities when they receive information requests from the public and businesses for data, especially when the requested geographical areas might not match the official boundaries of Local Authorities. The detailed geographical coverage of the census allows for the data to be aggregated depending on the users’ needs.

> ‘All boundaries for different services are different in our area. Having data at a low level of geography is how we can determine our information for the different areas. We don’t need data zone level all the time but we do need to split our council.’

> ‘Having data at a small area level is valuable, especially economic data. Being able to compare employment with areas of deprivation is good.’

> ‘We need to look at small data zones by age in years and gender for strategic reasons to understand where there are lots and less people going in to school.’

Some delegates noted that population estimates at the highest geographical levels (up to the national level), derived from the census and the subsequent Mid-Year Estimates (MYE)[3], are important in providing the denominators to calculate service provision. Other delegates noted that population estimates at lower levels are also essential to a range of other calculations used for service provisions in providing the denominator for calculations. This was particularly important to the delegates working in

Footnote

3) The mid-year population estimates provide an annual estimate for the population as of 30 June. They are based on census data and records of births, deaths and migration patterns.

public health, where rates use population estimates as their denominator with the numerator provided by other relevant data sources.

The increased demand for detailed information at low level data zones, and how important data at low levels is for effective resource allocation was discussed by some delegates.

> ‘There is an increased demand for access at lower level data zones. It was a Scottish Government initiative to get lower level statistics.’

> ‘Chief Executives want to know what is happening at low geography areas for resource allocation.’

2.8.1 Islands

Some delegates emphasised the importance of census data available at low geographical levels for effective service planning when working with low population density and variability between islands:

> ‘Each individual island is different.’

> ‘The census is the only source for information on the islands. All other sources amalgamate the areas and this is not even done consistently.’

For this reason, delegates felt that it was important to be able to analyse data at postcode level, especially as some local authorities cover multiple islands. However, one delegate noted that postcode level analysis can be problematic for analysis in the Islands:

> ‘Postcode level information is good to receive but as postcodes can be split across islands this can also pose a problem for analysing data.’

One of the most important aspects of the census for the Islands was that the census is unique in that it is the only survey that is compulsory and covers the entire population. Delegates felt that this was important in terms of response rates and ensuring extremely high coverage to provide accurate population estimates. Many of the delegates in the Islands use the population estimates from the census as a benchmark.

> ‘The census is a benchmark for any other statistics, which are either not available or not accurate because the numbers are very small.’

> ‘Although census has a 10 year gap, it allows checking of other data.’

2.9 Benchmarking

The census is used by a number of delegates to check the accuracy of their own data and to make sure that they are not ‘out of sync’ in the long term.

> ‘The census gets us back to a sound footing.’

A number of delegates emphasised that the census is widely considered a reliable, authoritative data source. The census was seen as being politically neutral and had no 'sponsorship' behind it.

> 'It is seen as a 'gold star standard' and is valued above all other surveys.'

> 'The census provides a benchmark against which other sources can be checked.'

A number of delegates also said that the census is important in providing a benchmark and to check the quality of their survey data:

> 'We also use it to check the quality of our survey. You adjust the survey to fit with the census. It is used as a baseline measure.'

> 'We use the census to provide health based information and weighted values for the local population samples for health studies. We also use the census to corroborate our health databases.'

> 'The census is the benchmark for future social research. Mid-Year Estimates (MYE) would not have a benchmark and could be wildly under or over counted.'

2.10 Trust in the Census

Across the discussions a range of delegates emphasised that the census is perceived as the most trusted and authoritative source of information. For example:

> 'The census is the single most comprehensive source that we have.'

> 'One could question a local survey but not census statistics.'

A number of delegates across the groups also raised the issue that it is the perceived authority and trust in the accuracy of the census that is important to their work. This perceived authority of the census was discussed in the groups as being important to legitimising planning and policy decisions.

> 'When plans and policies are developed people object. If we don't have the data we cannot defend and justify what we are doing.'

> 'We also have our own local sources of data but they don't have the same clout.'

> 'Everyone trusts the census – you just have to say it's the census and people never question it. It's as good as you get.'

> 'I would like to have the reliability of the census for every kind of information. The census is a baseline. It would be a lot of work to build up as much trust in other sources.'

‘Even people who are not interested in research are aware of the census and its reliability. People look forward to it for the whole decade.’

‘We still use 2001 Census data in equalities outcomes as we believe in the quality of the data.’

2.11 Information Requests

The delegates also discussed their role in providing socio-demographic information to external bodies such as community groups and local charities. They stated that these bodies require this information to plan their own funding bids and services.

‘Religious groups use the census for parish information – this cannot be found anywhere else.’

‘Local charities need data for funding (often Lottery funding) so they can reach the hard to reach groups they are there to help.’

Some of the delegates regularly received information requests about the census. They were expected to disclose and explain a wide range of outputs to members of the public, which includes academics doing research projects and small businesses looking for the best place to invest.

‘We are asked for things like small level population statistics and projections.’

‘We get requests from community planning and councils about settlement areas.’

Delegates also discussed the importance of the census in providing information to businesses. Demographic and geographical information can be effectively used when targeting investment.

‘Prospective businesses ask the council for ward information to see what areas are the best to start up in’.

‘The businesses come for planning purposes and profile of the ages to which they will provide services.’

‘There are private enquiries from small businesses. For example, we might get dental practices looking for socio-demographic data.’

‘Census information is used along with the Scottish Household Survey and Scottish Health Survey to find out social demographics of different areas. This is used for a market segmentation approach for detailed spots in the population.’

2.12 Social Survey Sampling and Weighting

Delegates from social research organisations primarily use the census to produce stratified samples[4]. Delegates from these organisations were worried about their ability to stratify if no census data was available.

> 'We have a team of statisticians that deal with sample design and weighting. It's used to design subgroup samples to find out where the subgroups are most likely to be. It's used for sample stratification methods.'

> 'If we need to target a specific subgroup then we will target that group from information from the census.'

It was also noted that census data is then used to calculate weights for their survey data after it has been collected. These weights adjust for the stratification sampling method used and non-response bias. In most surveys it will be the case that some groups are over-represented in the raw data and others under-represented. These over and under-representations are usually dealt with by weighting the data.

2.13 Research

The aggregate statistics derived from the census and subsequent Mid-Year Estimates (MYE) are used with other data sources as denominators in research. Some delegates use census population estimates as the denominator when estimating prevalence rates (the proportion of the population with an illness or condition at a given time) and the health ratios of health conditions. These estimates are based on both the census as the denominator and other data on the total number with a health condition to provide the numerator. This was important to a number of delegates:

> 'Census information is used to find out mortality rates and long term illnesses within ethnic groups in small data zone areas (5% of the population). This looks at death rates and health records looking back on trends dating back to the 1981 census.'

The census is also used longitudinally to determine how health varies across different groups and different geographies over time.

> 'The census is important for longitudinal study. Longitudinal study offers the opportunity to get new insights into Glasgow's mortality rate.'

The census is also used in specific areas to determine the context and potential reasons for social change, for example:

Footnote

4) Stratified random sampling is a technique which involves dividing a population or sampling frame into several, non-overlapping 'strata' (subgroups) according to a particular characteristic which reflects the variables of interest. Once the population or sampling frame is divided appropriately, simple random samples would then be selected from within each stratum.

> ‘I think understanding the de-population of islands is important. Simply knowing the number of people is inadequate. Contextual information is more important. Who are the people who are still living in low level areas and islands?’

> ‘I use it to analyse migration and the impact the expansion of the European Union has had on local authorities.’

Delegates expressed concern that household composition and family make-up would be harder to achieve using non-census methods and that the census provides a range of variables that are potentially harder to cover in administrative sources.

> ‘The census provides better data on household and family patterns than any other source. If administrative sources are used, the whole family doesn’t have to register with a GP or go to the same school, so it may not be as easy to gather family data as people think.’

2.14 History and Genealogy

It was clear that archived record level census data up to and including the 1911[5] Census was well used.

Census data had been used by delegates to ‘track down family histories and links between families’. Other delegates remarked that the census was often used by them as a ‘cross-over tool’ to ‘breach illegitimacy, name misspellings and street misspellings’.

Another example of family history use concerned the work of genealogists and the issue of probate and looking for possible family links to the deceased.

> ‘There is also revenue generated from genealogy and family tree research, which is particularly popular in Scotland. It is a hugely important issue for genealogists.’

2.14.1 Property History

It was evident from discussions that the rich information included in the census was used for more than simply tracing family histories. The census was also used for tracing the history of properties and how areas evolved through time.

> ‘I use the Census to trace the history of properties, sometimes even to identify their location. You can find street addresses that are, now, vanished. Tracing the enumerator’s route on the map.’

Footnote

5) Due to the Census Act, 1911 record level census data is the most recent publicly available.

2.14.2 Historical Tourism Tool

One area of work that involves genealogy as well as tourism was around the area of historical tourism.

> ‘The Census gives the pathway to historical tourism. A lot of North American tourists use the census to find their ancestors’

> ‘The second biggest usage of the internet is family history. Place of birth is one of the key things to make this work. Heritage tourism has become important.’

Census may be used abroad to discover the family heritage and then, through contact with agencies such as Visit Scotland, would be shown how they could visit the areas concerned (.e.g. the Islands), providing a possible tourism revenue to Scotland.

2.14.3 ‘Pointer’ Tool

Several members of the genealogy community indicated that the census when used with other records such as births, deaths and marriages could highlight possible avenues of research that would not otherwise have been apparent to genealogists. One example given was when someone had a rare job type and this allowed a positive identification of the individual in other record, allowing the specified piece of research to continue.

> ‘Census is a massive cross-over tool. Without it we cannot answer or confirm a lot of information regarding family history’

3. Perceived issues with the Census

Delegates at the events discussed some of the perceived limitations of the current census model.

3.1 Frequency of releases

For some delegates the current decennial frequency of the census was acceptable:

> ‘The ten year period seems a good interval’

However, for many delegates a key issue was the frequency of the census and the impact that this has on its reliability, quality, and utility. Delegates felt that the census becomes less useful as time passes, and the reduction in use of the 2001 Census is largely due to the decline in accuracy 12 years later:

> ‘The census was seen as relevant until about 2007 or 2008 then it wanes a bit.’

> ‘As years go by we trust less in the census. This is a real problem and it directly affects funding.’

> ‘[The census is valid for] four to five years, after that it's completely out of date. For instance, in our team we use ethnicity data for that period of time and afterwards we use a software to deduce ethnicity from the name and surname.’

Many delegates also expressed frustration at the time it takes from the day of collection to the publishing of population estimates and outputs.

> ‘Information on owner occupiers and private rental is good but due to the length of time a census turnover takes they are out of date by the time we get them.’

> ‘Its 2013 already so we are already 2 years on from when the census was done. It loses its relevance by 2017 and so we only actually get 4 years when it is relevant.’

Some delegates felt that the ten year frequency is not in line with changes in society. For example, some delegates felt that changes in the distance travelled and modes of transport are not captured in the data.

> ‘Travel habits have changed a lot in the last few years and they are not captured by the data.’

3.2 Housing and service provisions

Further implications of the ten year frequency of the census were also discussed by delegates in relation to housing and other service provisions.

> 'It has been a struggle in housing as the population grew too fast. We have a housing deficit and low numbers of affordable housing. Because it grew too fast we don't know where they are moving to, it's difficult to keep track. There are people who never come into contact with services. This not knowing is a huge problem. But the current census is not a solution, a more regular census would be better.'

> 'A new housing development is not caught in the 2001 census – in 2001 it was a field and 2011 we have however many houses. We have missed a whole set of data in the middle which makes planning for new schools and closing or merging them difficult.'

3.3 Mid Year Estimates and migration

Although the Mid-Year Estimates (MYEs)[6], which use the census as a baseline, are able to provide updated population figures by age and sex at a local level, the 10 year cycle of the census was still perceived to be a problem, by delegates, for reliability due to the possible problem of inaccuracies and cumulative effects in the Mid-year Estimates; these are regularly rebased following a new census.

It was discussed that the MYE need a regular census to keep them accurate. One delegate believed that 'the MYE become less accurate with every year that passes after the census' and felt that a more frequent census could limit inaccuracies.

Although the MYEs update the population estimates based on information on births, deaths, and migration it was felt that they do not fully capture migration or population change in the intervening period. Furthermore, it was felt that 'expansion of the EU and the changing economic climate' was also not captured. Ultimately, the current frequency was not felt to be in line, or representative, of the demographic and social changes that may occur over a 10 year period.

Other delegates raised the issue of needing more frequent reliable snapshots due to migrant populations. There was concern in all groups about the census missing peaks and fluctuations in the migrant population.

> 'Because of sudden influxes of migrant workers it is very difficult to keep track of the population. It is difficult to know our service needs. These are often urgent needs for families. We need help with local knowledge.'

Footnote

6) The mid-year population estimates provide an annual estimate for the population as of 30 June. They are based on census data and records of births, deaths and migration patterns.

‘We experienced one of the largest increases in migration in Scotland […] but how big is this phenomenon really? Currently we are just relying on anecdotal evidence. If this perceived influx of migrant workers is correct then support mechanisms need to be put in place. This requires more frequent and up-to-date data.’

3.4 Issues with geographical coverage

Aggregate statistics derived from the census are available for geographical levels from output areas to a national level. Due to data confidentiality reasons there are, however, minimum population and household threshold numbers for publishing output areas. In some cases an island on its own will not have sufficient numbers to meet the minimum threshold for population and household numbers and has to be included with another island or the mainland. The same rule applies for data zone areas but with different threshold numbers.

‘Islands are lumped into the mainland in some data zones. This is a problem. Census output areas for 2011 will be really helpful.’

A number of delegates felt that the current ten year frequency of the census made it difficult to understand social changes taking place within the area.

‘Aberdeen is a very interesting example. The oil sector has caused growth beyond what could be expected in other places. With a 10 year census it is hard to see exactly when the growth has been taking place.’

A similar problem was outlined for data on ethnicity, which is also unique to the census at lower geographical outputs.

‘In our small council the ethnicity variability is high and that means that data becomes outdated very soon.’

3.5 Statistical Disclosure Control Rules

Statistical disclosure control rules were also considered a barrier to analysis. In order to prevent information about an individual person being deduced from census outputs (as the census form gives respondents an assurance that their information will be treated as confidential) statistical disclosure techniques are employed to ensure that the risk of inadvertent disclosure in statistical outputs is minimised. This is applicable to all published statistics and would apply to all censuses.

‘Ethnic population counts often fall foul of disclosure rules.’

3.6 Historical data quality

Delegates at the genealogical event acknowledged that there were issues with data quality, especially in older censuses where respondents may have been illiterate and enumerators were not local.

> 'Historical Censuses contain wrong and inaccurate information'

4. Alternative data sources used

Delegates were asked about other data sources used in their work. A wide range of other data sources were used in their work, including surveys, administrative data and management information systems. A full list of sources used is given in Appendix B.

4.1 Surveys

Delegates used a range of government based surveys, including the Scottish Household Survey (SHS), the Scottish Health Survey (SHeS), and the Voluntary Population Survey. Delegates also mentioned the use of local or ad-hoc surveys where required but noted the expense in carrying out small scale specific surveys.

Delegates explained that these are used to supplement the census, rather than necessarily as an alternative.

> 'The census is supplemented by other smaller surveys, allowing us to create a picture by putting them together.'

The Scottish Government's harmonisation agenda[7], which has been running since 2005 to harmonise questions across Scottish Government surveys, was mentioned as a potential method to ensure consistency across surveys.

> 'The council have conducted annual household surveys which had 70% response rate. These could be developed. If 32 councils conduct household surveys and agree to ask the same 10 questions that might give you around 60% of the information required.'

4.1.1 Sampling

Delegates across the sessions discussed how surveys could be used as an alternative to the census in their work. Although most delegates agreed that surveys such as the SHS and SHeS are very useful sources, it was agreed that sample sizes across local authorities in Scotland was an issue. This was a recurring theme across all events.

In order to ensure that results can be said to be representative of the whole population a sufficiently large sample size is required. Usually a sample with between 1,000 to 2,000 interviews will allow for reasonably

Footnote

7) Further information can be found in the Survey Harmonisation section of the Scottish Government website.

robust analysis[8]. However, the sample size needs to increase if more detailed analysis of sub groups of the population, or aggregate statistics, are required. For example, if analysis is to include results for men and women separately then the sample will be divided in two (leaving 500-1000 per sub-group), but if there are five age categories to be analysed the sample will be cut five ways making the numbers in each sub-group smaller and the results less robust.

> 'The Scottish Household Survey provides a variety of useful information, but the problem is that it has a very small sample, only 100 houses in our council.'

> 'A sample survey is useless for actually identifying the locations for ethnic minorities. This is not currently part of administrative sources.'

Delegates stressed that not all local authorities have the same coverage, making it difficult for local authorities to accurately compare results nationally.

> 'Many years ago when I used the Scottish Household Survey, it had Moray in with the Highlands and not broken down. This made most of the data meaningless.'

Sampling is the process by which a feature of interest (or parameter) relating to a group of interest (or population) is estimated, by measuring its value in a smaller but representative sub group (or sample). The aim of sampling is to enable estimates or statistics that are as close as possible to the real value in the population. However, estimates obtained from samples can never perfectly match the true population parameters because of the information that is missing for the non-sampled population members. This can be an issue when the population is strongly grouped, e.g. by geography, and some groups may not be adequately represented in the sample[9].

> 'With samples people question how robust they are. There is always a question over how big the sample size is.'

> 'Although the sample size in the Scottish Household Survey is too low, I would not be in favour of removing it. It is a good and useful source.'

> 'The Scottish Health Survey is good but all we can do is compare Glasgow to the rest of Scotland and we can't rely on the sample being good enough in the rest of Scotland.'

Footnotes

8) Further information on survey design can be found in the Survey Methods section of the Scottish Government website.
9) The representation of clearly defined groups in a relatively small population can be improved by using sampling techniques such as stratified random sampling. For more information on the sampling strategies can be found with the Survey Methodology section of the Scottish Government website.

This was stressed in the Islands, where there was a perception about the problem of variability between islands which could limit the use of sampling techniques in surveys.

> ‘Samples are an issue as Shetland has a small population and the samples are just a sub-set of this and therefore can be less robust and have more issues and are really highly dependent on sample sizes.’

> ‘Both the Scottish Health Survey and Scottish Household Survey only use a small sample in the islands, and data is either meaningless or unreliable. Only 50 people are interviewed from the Western Isles in the SHS. In addition, sample surveys should take into account all islands of the Western Isles, as each island is different, and this is not always the case.’

> ‘We also use the Scottish House Condition Survey but as this only has 50 houses on the island that were sampled, this has issues with robustness.’

Delegates felt that to effectively plan services more detail is needed and samples should also be representative of groupings within specific islands, such as age-group, gender, ethnicity, and income-bracket.

The SHeS also has the option to boost[10] the survey but this is financed by health boards independently, therefore the option was used infrequently.

> ‘The expectation is that this cannot happen on a regular basis and each year the health boards analyse whether or not the data collected in the main SHeS is acceptable and most years this is the case.’

However, some delegates expressed concern about the potential response biases of survey methods. As it is important for the sample to be as representative of the population as possible, it is important that there are not differences in the characteristics of those who respond to a survey and those who do not. As the census is compulsory this was felt by the delegates to be less of an issue for the reliability of census data.

4.2 Administrative data sources

Delegates also used administrative data sources in their work. Administrative data is derived from information collected and maintained as part of an administration system, such as health records, vehicle licensing and tax systems.

The merits of a number of alternative data sources were discussed in more detail, particularly in providing information on hard to reach groups.

> ‘Some families never engage and if it wasn’t for schools and the school census we wouldn’t know about many of them.’

Footnote

10) Boosts can be used to increase the sample size for specific areas or characteristics.

Delegates discussed potential uses of the Community Health Index (CHI). CHI is a database in wide use throughout the NHS in Scotland containing data on patient demographics and some clinical information on aspects of healthcare screening and surveillance. The CHI number is, effectively, an NHS number and its use as a patient identifier makes it increasingly important to the implementation of 'ehealth' Electronic Health Records (EHR) and other Information and Communication Technologies (ICT) being introduced to healthcare in Scotland. Alternative uses of the CHI in addition to the NHS were discussed.

> 'CHI is better than the census at the moment, i.e. if planning for schools, we can see how many children have been born and are living in the area and who will be needing a school place in the coming years.'

However, issues of coverage and inaccuracies of the CHI were also discussed in other groups.

> 'The problem with CHI is that not everyone is registered or it is incorrect.'

Genealogy participants discussed government data sets as the electoral roll, property records, service records and maritime records which could be used to supplement birth, death and marriage registers.

4.3 Management Information Systems

The delegates also use a number of Management Information Systems (MIS) relating to their specific line of work including, Education management information software (SEEMiS), Social work information system (SWIFT and local authority Customer Relationship Management (CRM).

4.4 Monitoring tools

Some delegates were also concerned that many of the alternative and supplementary sources that they currently use have components that are actually derived from the census (for example, Scottish Neighbourhood Statistics and SIMD[11]) and were keen to emphasise how useful these sources were.

> 'SNS and SIMD are the main drivers for planning in our council area. They are useful for investment regeneration.'

4.4.1 Scottish Neighbourhood Statistics

Many delegates used Scottish Neighbourhood Statistics (SNS[12]). SNS is the Scottish Government's on-going programme to improve the

Footnote

11) Scottish Index of Multiple Deprivation (SIMD) available on the Scottish Government website.
12) Further information is available on the SNS website (www.sns.gov.uk)

availability, consistency and accessibility of small area statistics in Scotland. SNS has developed the systems to enable statistics across policy areas including information about benefits, education, health and the labour market to be brought together across a range of geography levels.

4.4.2 Scottish Index of Multiple Deprivation

The Scottish Index of Multiple Deprivation (SIMD) which is available on the Scottish Government website was used frequently by delegates. SIMD uses census data, in addition to a range of other sources of information, to calculate a measure of deprivation[13] within small data zone areas. Each data zone area has roughly the same population. The SIMD was considered to be a very useful tool for community planning by many delegates. Local authorities use SIMD to identify small area concentrations of multiple deprivation. There was recognition amongst the delegates that the SIMD has issues in areas with a low population density.

> 'There are problems with SIMD and its data zone areas. In the Highland area, because of the population density being low, often there are two or three areas with completely different demographics joined together. This can result in a data zone that can skew the apparent deprivation and 'hide' very deprived areas.'

4.4.3 Scottish Public Health Observatory

A number of delegates working in public health also used the Scottish Public Health Observatory (ScotPHO)[14]. The aim of the ScotPHO is to provide a clear picture of the health of the Scottish population and the factors that affect it. It contributes to improved collection and use of routine data on health, risk factors, behaviours and wider health determinants.

Footnotes

13) It incorporates several different aspects of deprivation, combining them into a single index. It divides Scotland into 6,505 small areas, called data zones, each containing around 350 households. Deprivation is defined more widely as the range of problems that arise due to lack of resources or opportunities, covering health, safety, education, employment, housing and access to services, as well as financial aspects.

14) The ScotPho collaboration is co-led by Information Services Division (ISD) Scotland and NHS Health Scotland, and includes the Glasgow Centre for Population Health, National Records of Scotland and Health Protection Scotland.

5. Implications of not having a Census

Delegates discussed what the main implications would be if the census outputs were not available. For many delegates the census is essential to service provision and to understanding their performance. Some delegates expressed concerned that if they did not have the census they would, ultimately, need to produce their own census-like event for adequate coverage at lower geographic levels, which would have large financial implications.

> 'We will have to use other estimates provided to us. We could also use local surveys but we will need to finance them.'

> 'You will need local solutions for more accurate statistics. Over time data will become more unreliable but it will not be a political priority to update your statistics, but you will need them to check if your policies work or don't work, or to present your case for attracting funding.'

A number of delegates stated that they have tried carrying out their own surveys but this has proven to be extremely expensive.

> 'LA also ran their own local survey but this is expensive to do and is not possible to repeat on a regular basis.'

> 'Health boards also have an option to run their own surveys, […] but this is expensive and not always an option.'

It was also noted that it would difficult to provide low level outputs without the census and that the importance of this is more likely to increase over time.

> 'I don't see pressure for information at data zone level going away.'

> 'The potential reduction in the number of councils across Scotland needs to be taken into account here. The impact of any reduction on databases, internal data sources would likely make the census and its flexibility down to Output Areas even more valuable.'

For some delegates the repercussions of not having the census would be very wide-ranging for resource allocations. For many delegates the census is essential to service provision and to understanding their performance. Resource allocation would not be as effective without the evidence from the census.

> 'If the census was not available there would be resource implications. It would have a domino effect as a lot of things are under-pinned by the census.'

> 'Without the census we would not know what service provisions are needed. We would not know how well we were performing. We would have to reinvent the census at lower levels.'

‘There would be huge cost implications as we need this information, and we are used to having this information at our fingertips.’

‘You could have housing investments built in the wrong places.’

‘General capital spending and planning across the board would be less reliable.’

‘We look to census for priority areas and fragile areas. Without census, money would go into the wrong areas.’

One delegated noted that the census was an essential part of what they do:

‘There would be a panic if we didn’t have the census.’

Delegates stressed the importance of high quality data to informing policy decisions:

‘Wrong decisions being made on the back of a lack of information or poor information is unquantifiable.’

In terms of legal implications, the delegates reiterated the legal requirement to monitor and assess equalities information. The importance of equalities indicators has been discussed in earlier outlining current uses of the census.

‘The only legal requirement of the Single Outcome Agreement is to keep reviewing data.’

Some delegates thought that they could rely on other data sources to meet Single Outcome Agreements (SOAs) details can be found on the SG website.

‘If the census was not available it would not be that bad as tend to use other data sets in the council for SOA. It would be more upheaval if the census was the only source of information.’

‘SOA is high on the agenda […] we could do without census as long as there was something on the population.’

A point was raised in one session which suggested that the census allowed local authorities to work in an open and non-biased way, ensuring that data quality was the same across the different departments and authorities.

‘The census is neutral and avoids infighting between departments in Local Authorities. No census could introduce bias at local level. If council X is richer than council Y they could have better data quality, for example.’

One delegate noted that current Gaelic policy assumed that there would be continuity in census data:

> **‘**There is a Scottish Gaelic Policy (available on the Scottish Government website) target based on there being a 2021 Census.’

There were also concerns about the impact of not having the census in terms of survey sampling design:

> ‘If we did not have the census it could increase sampling error and it would also interrupt time series data.’

A number of delegates said that ultimately, as Local Authorities need the data that the census produces, they would have to carry out their own surveys. Delegates believed that this would have financial implications and also ramifications for the quality of the data.

> ‘If you remove the census people might start commissioning their own surveys which would have low response rates and less reliable results. This would mean national savings are spent on sub-standard surveys throughout Scotland instead of one which can be relied on. I think local authorities would value and pay for having the census in their area.’

> ‘To see it as a cost saving nationally is a mistake. The census only costs about one pound per person per year. Is this a bad thing?’

6. Views on Future Models

Delegates were given an opportunity to discuss the different possible options for producing population and demographic statistics in future. These included a modernised census and using administrative data combined with data linkage.

6.1 Administrative Models and Data Linkage

One of the main alternatives to taking a census would be to use linked administrative data to produce population and demographic statistics.

Data Linkage is the joining of two or more administrative or survey datasets to increase the power of analysis possible with the data. The Data Linkage Framework (available on the Statistics section of the Scottish Government website) aims to address barriers to data linkage, and to widen the range of data linkages that can be carried out, without impinging inappropriately on personal privacy of data subjects.

At the events delegates discussed the merits and problems of using administrative data sources and data linkage. For some delegates this was based on their own experience of using administrative data and data linkage projects. Opinions on using alternative sources were generally positive, with delegates agreeing that there is huge potential for them to be used in future. There was also some discussion over international models.

> 'I've been told that in Finland they don't need a census as they have a register system in operation. Every time anything happens they report it. You could virtually run a census at any point in time.'

One delegate told the group of a linkage program which had been conducted in England, stating that it was very successful but did not receive enough political backing.

> 'There was a big program in England called 'contact point' (Encyclopaedia of Informal Education website) which joined children's data together. It worked for every child in England and the purpose of it was to highlight children in need. A massive cleaning operation went on. It even included records from children's charities. It seems to me that younger generations are happy to share information.'

> 'Health data is really difficult for us to get. […] It would be really helpful if the government could get health data and council data together and publish it with the same reliability factor as census.'

There were some interesting discussions on potential additional uses of the Community Health Index (CHI) relating to data linkage. CHI's use as a patient identifier makes it increasingly important to the implementation

of 'ehealth' Electronic Health Records (EHR) and other Information and Communication Technologies being introduced to healthcare in Scotland. Alternative uses of the CHI in addition to the administrative function for the NHS were discussed.

> ‘Since 1991 we have been taking snapshots of the CHI. We have postcode, age, sex and medical practice for everyone living in the area, so we are able to work out populations corresponding to SIMD data zones. We can use this to work out what populations were in the past, and can work out past history of geographical incidence of a disease.’

In addition, the CHI was suggested as a potential index data set for data linkage.

> ‘The CHI is a potential vehicle for a form of register. There is a unique identifier which should eliminate duplicates.’

Problems with the timing of updating or cleaning administrative sources were also discussed in relation to the CHI. The impact of this could potentially be that the population estimates drawn from the CHI could be inflated as patients who have died or moved from a GP’s practice could be left on the CHI system. One delegate commented:

> ‘Over-inflation is also a concern. We know that it takes GPs a long time to update records.’

One of the key benefits of data linkage projects discussed was validating local authorities’ own data. For example, a delegate stated:

> ‘Linking projects are good to validate our own data. When there are discrepancies you can take a note of the gaps. The potential of data linkage projects is exciting and we are happy that is very much in the agenda of the Scottish Government.’

Some delegates discussed work recently completed in North Lanarkshire council to create ‘golden records’, which are administrative records with very high accuracy.

> ‘Golden records can be created by validating who people are and where they live. There is a council which has achieved these already. It was done by working with the NHS Central Register (NHSCR). Councils are now looking in detail at data quality.’

One of the main concerns by delegates was the perceived coverage issues of administrative sources. Some delegates stated that some administrative sources can reveal additional information on certain groups that are hard to reach in the current census model.

> ‘We know from our own projects using the School Census data that census data was not completely correct for asylum seekers.’

‘The School Census is a good source of information, especially for ethnicity and changes in family structure.’

‘From the School Census we learnt that we have 148 languages in our council. Jobcentre Plus provide some statistics on ethnicity for jobseeker’s allowance.’

A number of delegates had also used data linkage to provide evidence for policy changes or implementations in their areas:

‘We used information from police data and linked it to Accident and Emergency regarding alcohol related incidents to estimate the cost of alcohol abuse.’

On the other hand, delegates noted that other administrative data sources, such as the electoral register, might have their own coverage problems and would therefore need to be used alongside other sources.

‘The electoral register could be used but there are coverage problems capturing vulnerable groups.’

‘The School Census is not so good if there are private schools in the area. They don't provide accurate information.’

6.1.1 Comparability

Delegates also had concerns about the practicality, cost and the time that it takes to maintain clean administrative data in order to carry out data linkage:

‘Could it be more work to clean up all the datasets than we think? The NHSCR is tidy but it is a full time occupation. Lots of people work on it every day. Could an alternative end up being more expensive than the census?’

Some delegates noted that administrative data sources were not conceived with a statistical use in mind, therefore it can be challenging and time consuming to ensure that these sources are consistent across different areas, especially if the remit of the source varies.

‘Using administrative sources is a great idea in principle, but it would be a great task to make data internally coherent, let alone on a national scale.’

‘My fear is that you have 32 different local authorities. There is likely to be no consistency in their data. One could have good data quality and one very poor data quality, which makes any comparisons pointless. For this idea to work you would need to make sure data from local authorities is as robust as it is in the census.’

‘Administrative sources vary tremendously in update cycles and categories. Trying to cross harmonise variables with administrative sources is very difficult.’

One delegate expressed concern that administrative sources would not be consistent and that this could add additional problems when trying to produce co-ordinated approaches across the UK.

‘It is important that census has a standardised UK wide output as it is difficult to have UK policy when you can’t compare statistics throughout it. An administrative source is not collected with this in mind.’

6.1.2 Legal Barriers

Delegates with experience of data linkage also warned of practical problems and potential legal issues when using administrative data.

‘I have various concerns about merging data. Our organisation merged with others and we consequently needed to merge data sets. It’s never as easy as it seems. We had a nightmare deciding which data is the stronger or more reliable. In reality it’s really challenging and there are a lot of technical and legal issues. There are so many regulations and agreements about how the data can be used. That is not to say that it can’t be done and it won’t be better – it’s just really difficult.’

‘We use administrative data with Jobcentre Plus. There are problems technically and legally as we need to look at actual names to get a proper match.’

‘We hit a brick wall with legal issues. I think Local Authorities are very keen to share information but the legalities get in the way. People are nervous of putting a foot wrong.’

‘The ICO (Information Commissioner’s Office)[15] is so strict on matching – privacy protection is so strict. Child protection issues are a nightmare.’

Communication problems between data holders were highlighted as a barrier to data sharing.

‘The problem is data holders won’t talk to each other. It would be great to link across, but it’s very difficult.’

Footnote

15) The Information Commissioner’s Office (ICO) is the UK’s independent public authority set up to uphold information rights. The ICO enforces and oversees the Freedom of Information Act 2000, Data Protection Act 1998, and the Privacy and Electronic Communications Regulations 2003.

The Data Protection Act (1998) (legislation.gov.uk website) and its implications[16] were seen as an issue for many delegates.

> ‘They [Local Authority] had to give up matching different data sources as individuals could be identified. This was a data protection issue and stopped the progress of the project.’

6.1.3 IT and Knowledge Transfer

The group emphasised problems with software and infrastructure that inhibited data sharing and linkage.

> ‘We have strong partnerships for data linkage but we don’t have the infrastructure.’

> ‘There are so many queries and access issues. We need to change it to make the software better – we are fighting through these barriers. Better software would solve the problems first.’

Additionally, delegates across the groups discussed inadequate data management skills and knowledge. Delegates claimed that information about how to access and what information is available is not passed on between employees.

> ‘Sometimes the people who ‘own’ databases have inherited them and don’t actually know how to use them. The people who created them have left and nobody knows how they work.’

> ‘One of the problems is that a few people run the system. When they go we need to build resilience so that we don’t lose information and how to access it.’

6.1.4 Quality and Accuracy

One concern amongst delegates was the flexibility of administrative sources to changing priorities and targets over time.

> ‘There should be core indicators that don’t change over time. Some questions change when targets change. It will be even harder to deal with when using administrative sources. They are more vulnerable unless ring-fenced by the legalities of the census. If this is all done you may be able to do an annual census. I can’t see how you would arrive at the same cost though.’

A number of delegates expressed concerns about the use of administrative sources as an alternative to the census. One delegate had experience of using administrative sources and found their reliability uncertain.

Footnote

16) More information on the Data Protection Act can be found on the Information Commissioner’s Office website.

‘Worryingly for me, when I looked into it, the match between census and other administrative sources was very poor in some areas.’

‘I don’t think an alternative to the census would be any better or faster. I have previously tried to use a database based around an address list and on top of the coverage being low, loading in information quickly was impossible.’

There were also some concerns by delegates about the validity of some administrative sources in measuring employment, for example. Delegates felt that benefits information alone is not adequate to understand the extent of unemployment as not all unemployed people claim benefits, and some people are not eligible to claim. A delegate explained:

‘You can just get those who are eligible in employment data. You can’t count unemployment in 16 and 17 year olds.’

6.1.5 Privacy and Security Concerns

Some delegates were also concerned about the impact that the use of administrative data could have on individual’s privacy. A delegate also said that they would have concerns about privacy if their personal information derived from administrative sources was used.

‘At a personal level I am concerned about a Big Brother society. The public will also have great concerns I’m sure. At least with the census you can choose what to answer.’

One delegate had concerns over the security of data from administrative sources.

‘We do a lot of linking projects which will help us in the future to do analysis. For instance, we are linking home carers with housing data, but these projects find a lot of problems with data security.’

‘Guarantee of confidentiality is important. People are suspicious of filling in personal data. If they can guarantee it is totally protected by law this would help alleviate the fears of those who think we are approaching a Big Brother society.’

6.1.6 Perceived Trust in Alternative Sources

As previously highlighted, a number of delegates argued that one of the most important aspects of the census is the perceived reliability and accuracy of the census.

‘There is a trust in what is provided, a trust in the census’ consistency and reliability. I think the problem with other data sources is that someone might say ‘I’m not sure about this data set’ and no one will use it. We need to build confidence in alternative data sources.’

One delegate raised the issue of this trust being related to branding and its official status and whether or not this is potentially transferable to other sources:

‘Is it possible that we all have more confidence in the census simply because it looks official? If you produced results from administrative sources and they had the NRS stamp on them would people trust them as much as they trust the census?’

A delegate also emphasised the need for transparency and the need for data to be available to the public:

‘I think the wider population need to understand the decisions we make – is it evidence-based or a hunch or prejudice? It needs to be based on information available to all.’

Delegates discussed what the next steps to using administrative data and data linkage might involve.

‘We need to run a pilot study. We need to see if it works and get the counts. It might be a change for the better.’

‘We need to show the benefits. It might take a very long time. It might be generational as we move on and use different technology.’

6.1.7 Impact on Genealogy

Genealogists discussed how administrative data models would impact on their work, especially if there would be no census type form involved. All members acknowledged that the impact of any changes would only become evident in 100 years time, but they understood that the direct consequences related to their field.

‘That [administrative only option] would have a negative impact on genealogy; but the Census did not start to satisfy the needs of genealogists, our research is a by-product of it.’

‘These options [administrative dependent options] will remove a really important source of information [for Genealogists].’

6.1.8 Beyond 2011 Timetable

One delegate expressed concerns over the timetable for Beyond 2011, indicating that the decision relating to whether to conduct a census in 2021 should not be made until the 2011 Census benefits realisation work had been completed.

‘A decision to replace census should not be made in time for 2021, it should be allowed to take longer. There will not be enough time to report on the value of new research based questions in the 2011 Census before a decision is made.’

‘Data linkage will take time. It should be pursued but I don’t think it will be ready to be an adequate alternative for a long, long time.’

6.2 Online Census

Delegates were, in general, extremely positive about the potential of the option to develop the online completion of the census.

> 'I think that developing the online option would be very good. It would allow more detailed questions to be covered for less money.'

A number of delegates were interested in the possibility of the traditional census being completed and returned electronically and felt that the online option should be explored further.

> 'The 2011 Census was the first census to have an online option. This is surely the way forward and would produce more accurate data.'

> 'In Canada they use the internet a lot more. There is a one-page sheet handed out with internet access code. You can encourage online participation.'

> 'Timeliness is an issue with the census. Even if it came out every 10 years but the results were produced quicker it would be better. It is now almost two years since the census happened and we don't have full results yet. I think greater use of online resource could speed this up in future.'

There was also concern over the impact of changing to a web-based format and how to ensure returns are accurate as well as the possible impact on response rates.

> 'There is a percentage of the population who will never complete it online, particularly older people.'

> '…people who are now in their 60s and 70s will not suddenly start using the internet in 10 years time.'

6.3 Short Form plus Survey

There was some uncertainty among the delegates about the short form plus survey option due to the restrictions on the questions applied to the short form and the impact that this could potentially have on the information needed for service provisions. Some delegates said that they would be very keen to see the short form and give feedback on its content.

> 'If you were to have the short form/long form method then I think we would need a mock-up of the survey and feedback to you on it.'

6.4 International Models

Some delegates were also interested in what work was being developed internationally on the census and Beyond 2011 programmes and how this could be applied to Scotland and the UK as a whole.

Delegates discussed the different census-type options that are used in other countries. Concern was expressed over the apparent disparity in the French rolling census model.

> 'We can't have a census like the French one that has new information for some areas and older information for others. This is an inequality.'

> 'In France, you do lose information on migration, and there are no questions on travel to work. There isn't a full picture at any one point in time. At low level demographics it can be difficult to produce this. It is more timely on population figure outputs though.'

> 'I'm keeping tabs on the census and looking at international developments. I'm interested in seeing if there are any alternatives.'

7. Future Requirements

Delegates were asked to discuss the following aspects of the information they use and their relative importance:

- Quality and accuracy
- Geographic coverage
- Frequency
- Aggregation

For some delegates all aspects were considered vital:

> ‘If you sacrifice any of those characteristics then the credibility of what we do is in question.’

At the events the common view on the relative importance of the key aspects was that this depended on what use of the census or variables were being discussed. For example, the view for housing may be very different for the view on education. It was also stressed that this depended on local outcomes and needs.

> ‘[Census releases] should be tied to government directives, government or parliament objectives, health boards and planning work.’

Some delegates felt that quality and accuracy were more important than the frequency of the survey. This was particularly important for measuring outcomes and understanding change over time. There was discussion on the expectations of local authorities and the pressure from users to have accurate data.

> ‘Quality is the priority – it is important that data is trustworthy and detailed.’

> ‘I would prioritise quality and accuracy, as reliability is the most important.’

> ‘If it’s not at the appropriate geography and accuracy it’s unusable. The frequency is desirable but not essential.’

Comparability was discussed and some delegates also pointed out the importance of having ‘fixed points’ in order to consider ‘long-term patterns’ or to carry out time-series analysis. Delegates were also keen to express the importance of being able to compare 10 year periods.

> ‘Comparability and consistency are important to look back through the years.’

7.1 Accuracy

Delegates discussed the importance of accuracy in the census and surveys in general.

> ‘Reliability is very important to historians. How do you capture a society that has changed a lot?’

Delegates noted that it would be better to have a more regular census to help improve the accuracy of the Mid-Year Estimates (MYE)[17].

> ‘It would be good to have something every 5 years to get the Mid Year Estimates back on track.’

Some delegates felt that accuracy was very important, with particular significance to the islands because of the small population numbers.

> ‘Small numbers can skew results easily when there is a minor change, unless we know the underlying population figures’.

The importance and role of coverage surveys was also discussed[18].

> ‘The census doesn’t have to be 100% accurate; it just has to be accurate enough. Coverage surveys play a very large part.’

7.2 Frequency of Updates

The optimum frequency of data was also discussed. There was a wide variation in the responses from delegates and was highly dependent on their area of work.

Some delegates thought that more frequent data would become burdensome as they could potentially end up in a continual cycle of updating their information, which would also have resource implications. Some delegates agreed that more than once a year would be unnecessary as seasonal fluctuations could make the data difficult to interpret.

> ‘I would be happy with every two years but every year would be the best for strategic assessments. Economic questions would be better as often as possible.’

Delegates debated how often they would like to receive population and socio-demographic information at many events the 10-year frequency was criticised, however this view was not held by all delegates.

Footnotes

17) The mid-year population estimates provide an annual estimate for the population as of 30 June. They are based on census data and records of births, deaths and migration patterns.

18) The coverage survey is conducted independently of the census and is designed to calculate how many individuals may have been missed by the census enabling accurate population estimates to be calculated.

> ‘I think it’s unlikely you would be able to get census accuracy annually. Having the census every 10 years is acceptable because you can depend on it. The General Health question is worth waiting 10 years for as it is a reliable source to base information from.’

> ‘It would be useful to have robust population figures more frequently than at present but perhaps supplemented or benchmarked against a ten-yearly (as present) full census.’

> ‘Information seems to be out of date. It’s already two years out of date and we can only use it for five years after that. I would prioritise frequency. That would be fantastic.’

Changing populations was cited as a reason for wanting data on a more frequent basis.

> ‘People are more geographically mobile than they used to be, especially in cities, so more frequent than 10 years would be good.’

> ‘We need to keep it up to date – every two years would allow capture of transitory population.’

> ‘This should be on an annual basis, any longer and it is not possible to measure change but practicalities have to be taken into account.’

Some delegates felt that the frequency of releases should be focused around planning cycles and to ensure that there was data available to match these requirements.

Delegates related their ideal frequencies for accessing the data to their different roles and uses of census data. Delegates discussed that this would allow Local Authorities and Health Boards to revise their case every three to five years in line with the planning cycle.

> ‘The data is important to feeding into SOAs and producing data for the 3 year periods and also the yearly targets. SOAs have also now changed so it is a rolling plan to produce statistics on a regular basis.’

However, others felt that there was a need for flexibility in the frequency and approach:

> ‘Although most interventions are over extended periods, for example, disease prevention, others work on a shorter timescale and therefore a flexible approach might be required.’

Other solutions were proposed including a five year census combined with other datasets for the intervening years and a topic focused census on an annual basis.

7.3 Geographical Coverage

Geographical coverage was also considered to be important, and delegates recognised this would be especially important for people in rural and island areas.

> ‘It’s the geographical spread and sample size across the country that is important.’

> ‘Geographical coverage is also important as we need to pinpoint specific towns with problems, especially with depopulation.’

The delegates thought that quality and accuracy go hand in hand with these aspects and are also vital.

Geographical level of data was deemed to be very important for those involved in community planning and travel to work maps.

> ‘It is best to collect data at a low level. We are confident in census data because it has come from non-aggregated data.’

However, this was no consistent view across events and some delegates felt that the geographical coverage was viewed as being ‘useful but not vital’. One of the groups discussed the importance of data zones and the changing expectations for information at this level.

> ‘We have worked hard to build awareness of the availability and it is now embedded. There are expectations and we are asked for information at data zone level more and more.’

Some delegates prioritised accuracy and geography combined:

> ‘If it can’t be geographically located then why are we actually collecting it at all? Obviously quality is essential but if you have highly reliable data and you don’t know where it’s from it’s pointless.’

7.4 Aggregation

Delegates also discussed the level of aggregation of variables needed. It was agreed that sensible aggregations of variables could be used but allowing an option for more detail if required.

> ‘At the moment we need to have accurate age groups, as there is an ageing population and are using the census for that.’

Many delegates stated that they would need data that was aggregated at low levels, therefore providing the most detailed information. Related to this delegates emphasised the importance of being able to produce cross tabulations. Cross tabulations summarise data based on two or more of the characteristics (or aggregate groups) and can show the interrelationship between two or more characteristics. For example, one of the delegates said that they would like to produce cross tabulations on language and income:

‘What Gaelic speakers are earning, i.e. cross tabulation, is important. If data was collected by other means but could still be cross tabulated that would be fine.’

‘I really value the fact that we can do cross tabulations. We don’t have any other sources that can do this.’

‘All of these things are important. Comprehensive cross tabulation is the most important aspect. Coverage in terms of attributes should be up there.’

For many delegates, however, aggregation or disaggregation was not more important than the other factors.

‘In our case we’re not interested in individual level records, so we think accuracy, frequency and geography are most important.’

‘I would give up some disaggregate information if the frequency was better.’

‘I don’t mind the 10 year wait if I can have data which has not been subject to aggregation.’

Aggregation by protected characteristics under the Equalities Act 2010 (on the legislation.gov.uk website) were also important to some delegates:

‘Equality laws deal with protective characteristics and how results are presented. We have to provide an equal service to the whole community but we don’t know the community. We need something to advise us.’

‘The core demographic statistics by age and gender is easy to obtain, it's the other characteristics that's the problem and there's no other way to collect that information other than a big survey.’

It was felt that any aggregation of variables, e.g. only providing a small range of job classifications would negatively impact the work that genealogists do.

‘Family historians are interested in the family picture not necessarily a “Victorian” family as many families now have different relationships/names. This information is used as building blocks by family historians.’

‘Occupation is important as it can lead you to other records.’

8. Equalities

8.1 Introduction

National Records of Scotland (NRS) ran a series of equalities specific discussions. Comments relating to equalities from the non equalities specific Beyond 2011 stakeholder engagement sessions have also been included to help build a comprehensive picture.

Many of the delegates were from public bodies and their current use of census information was related to the public sector Equality Duty, which states that no one should be denied opportunities because of their race or ethnicity, their disability, their gender or sexual orientation, their age or religion. Therefore it is important that any requirements by our users are recorded.

8.1.1 Public Sector Equality Duty

In order to comply with the specific duties within the public sector Equality Duty (section 149 of the Equality Act 2010 both PDF documents available on the Press for Change website) some public bodies currently rely on information derived from the census. The public sector Equality Duty applies to public bodies and others carrying out public functions. It supports good decision-making by ensuring public bodies consider how different people will be affected by their activities, helping them to deliver policies and services which are efficient and effective; accessible to all; and which meet different people's needs.

The Equality Duty is supported by specific duties, set out in regulations. The specific duties require public bodies to publish information to show their compliance with the Equality Duty, at least annually; and set and publish equality objectives, at least every four years.

The information they publish must show that they had due regard to the need to:

- eliminate unlawful discrimination, harassment and victimisation and any other conduct prohibited by the Act;
- advance equality of opportunity between people who share a protected characteristic and people who do not share it; and
- foster good relations between people who share a protected characteristic and people who do not share it.

The following characteristics are protected characteristics, outlined in the Act:

- age
- disability
- gender reassignment
- marriage and civil partnership
- pregnancy and maternity
- race
- religion or belief
- sex
- sexual orientation.

8.1.2 Single Outcome Agreements

Considerations of equalities characteristics are specifically important for Local Authorities (LAs) when producing and maintaining Single Outcome Agreements (SOAs). SOAs set out how Community Planning departments in councils will contribute to the Scottish Government's National Outcomes (further information available on the Scottish Government website). The purpose of a SOA is to identify areas for improvement and deliver better outcomes for the people of Scotland as well as individual Local Authorities. The concordat set out the terms of a new relationship between the Scottish Government and local governments. In coming to this arrangement, national and local government share a commitment to equality and legal duties and are subject to the same legal requirements in relation to equalities.

In order to fulfil these specific duties and SOAs some public bodies currently use data from the census as relevant evidence. The equalities information derived from the census is outlined below.

8.2 Equalities Information

The 2001 Census covered the following 'protected characteristics'[19]:

- Age
- Disability (a question on general health and whether or not have long term illness, health problem or disability which limits daily activities or work)
- Marriage
- Race (question refers to ethnic group)
- Religion or belief (one question on religion belong to and a second on the religion an individual was brought up in)
- Sex

A list of the relevant equality questions are included in Appendix A. The 2011 Census extended the disability question to include information on the nature of the condition and a question about the ability to understand and speak British Sign Language (BSL). The marriage question was also extended to include civil partnerships.

Information provided in returned census forms[20] are analysed and aggregate statistics are derived. Aggregate statistics are a collection of summary counts of the numbers of people, families or households resident in specific geographical areas possessing particular characteristics, or combinations of characteristics drawn from the themes of population, including the outlined protected characteristics, work, and housing, for example. The aggregate statistics (relating to the protected characteristics) are available at a range of different geographies from census outputs areas to a national level. Information at record level in the form of Sample of Anonymised Records (SARS) is also available[21].

Footnotes

19) Copies of the 2001 Census questionnaire (PDF NRS website) and the 2011 Census (PDF Scotland's Census website)

20) Adjustments are made in the datasets and outputs for people and households who may have been missed out and not been recorded on a form, or may have been counted more than once by being recorded on more than one form. Some records are also swapped for disclosure reasons.

21) More information about the methodologies around Census 2011 can be found on the Scotland's Census website.

8.3 Uses of the Census

8.3.1 Service Planning

In order to comply with the public sector Equality Duty and Single Outcome Agreements (SOAs), aggregate statistics derived from the census are used by public bodies to help them to plan policies and services which are efficient and effective.

> 'We use the census as the basis for designing our policies for equality and diversity at LA level and understanding the population we deliver services to. The census allows comparing the population profile between LA and from there define how local policies should differ.'

> 'Census data allows understanding of the needs of the different equality groups and is used to collect evidence to propose new services.'

Information from the census is also used to help identify those equality groups that are most in need of assistance, ensuring resources are effectively deployed in the right area. Some delegates argued that this is becoming increasingly important as demographic and social changes will entail an increase in the demand for public services, which will also be intensified and prolonged by the economic climate.

> 'Due to cuts in budgets the college has to make sure that they reach deprived areas and pin point down to post code and output areas.'

8.3.2 Monitoring Outcomes

The census currently provides LAs and NHS Boards with baseline information on protected characteristics. Statistics from the census are used to monitor the equality outcomes across a range of services.

> 'Census data is used as a benchmark to assess the outcome with regards equality group's policies.'

Other public bodies also monitor equality outcomes for their services in order to assess the equality of opportunity to participate in services and events.

> 'We collect the participation in cultural events from different organizations and compare it with the population proportion for those groups, to understand what groups are not being involved in the same proportion. The council has specific objectives in terms of the ratios of participation from the different groups.'

8.3.3 Unique Estimates

Many delegates stressed the importance of the census as being the only source of equality information, particularly relating to ethnicity.

> ‘The census is the only source of information on ethnicity. We are required to provide information and compare ethnic minority groups’

> ‘For us ethnicity information from the census is very important and you can’t get it from other sources.’

> ‘Only the census captures equalities data.’

> ‘We rely on the census heavily for information on ethnic groups to understand service needs and initiatives.’

> ‘Census has data on ethnicity which would otherwise be hard to obtain.’

Equality statistics derived from the census are available at a range of different geographies from data zones to a national level. This aspect of the statistics was considered very important by a number of delegates to allow effective planning and targeting of resources.

> ‘No other survey or data source gives the data on a lower geographic level other than the census.’

Census data was considered the most accurate data source by many delegates. The accuracy of the data was related to the legal requirement to take part in the census by a delegate:

> ‘Because the census is a legal requirement it gives more reliable data.’

A number of the delegates also stated that they were looking forward to the full release of the 2011 Census data, particularly to review the data from the new question on disability.

> ‘We need census data to understand how many people will have needs in terms of learning disabilities and carers. We are very interested in seeing the results from the 2011 Census on this.’

8.3.4 Policy Development

A number of delegates across the different groups also raised the issue that it is the perceived authority and trust in the accuracy of the census that is important to their work, particularly in relation to justifying planning and policy decisions.

> ‘When plans and policies are developed people object. If we don’t have the data we cannot defend and justify what we are doing.’

Some delegates from equality groups also said that the census is a vital source of evidence, which can be used to influence organisations and to lobby.

‘The census is critical for identifying need and for ensuring people can speak with confidence in arguments when trying to influence organisations.’

‘For equality groups covered by the census the outputs provide evidence to lobby the UK government for disability issues.‘

8.4 Limitations to the Census

8.4.1 Frequency

A number of delegates were critical of the 10 year frequency of the census. They felt that the census does not adequately capture changes in society and population figures at the current frequency.

‘10 year frequency is far from ideal as the population, particularly in Edinburgh, has changed dramatically since 2001, even in the period from 2011 – 2013 there has been significant changes, particularly in the Eastern European community. There is likely to be more Romanian migrants to the city this year so even the 2011 results will not accurately reflect the city’s population. ‘

‘The 10 year cycle of the census makes it very difficult to know what the ethnic profile of Scotland is. Scotland may not have been particularly diverse in the 1990s but it is now due to the dispersion programme of asylum seekers and other government initiatives.’

As the 2001 Census data is now out-of-date many delegates were waiting for the more up-to-date 2011 Census data to be released before using census data in their analyses.

‘Government based equality outcome agreements can no longer be based upon Census 2001 statistics as they are irrelevant and out of date.‘

‘The delay in 2011 census is causing problems to teams working with minority groups as the numbers from 2001 are clearly out of date.’

‘The crux of the issue is how to collect equality information that is timely and accurate […] the 2011 delay is very frustrating as you can’t predict changes […] it is hard to make a case for a particular cause if you don’t have accurate figures […] its difficult to argue your corner.’

8.4.2 Question Design and Definitions

As referred to in the introduction, the census requires individuals to complete a selection of equality questions using pre-determined classifications, which some delegates have argued not to be appropriate or understood by all individuals. This is considered to be a problem, in general, with survey methodology when asking survey respondents to classify their identity, which many feel to be very complex and subjective.

> ‘We all have multiple characteristics and information can be diluted by the census rather than capturing its richness. You need to try and balance rigor against fluidity.’

> ‘You have to be careful if you adopt a targeted approach to data collection because if you go looking for people then you’ll find them so who are you missing along the way?’

Although extensive research, consultation and question testing is carried out on the ethnicity question in intra-census years to ensure it reflects changes in society, it was not possible to reach agreement with all equaltiy groups about the categories included and the structure of the question[22]. Some delegates disagreed strongly with the implied definition of ethnicity relating to skin colour and the structuring of the ethnicity question[23].

> ‘The problem is that it mixes up ethnicity with pigmentation. If you are going to ask about pigmentation then just do it.’

The sensitivity of an individual attributing themselves to a certain category can also be influenced by specific histories and politics, such as the historical sensitivities surrounding Jewish people being identified by the state. This can lead to Jewish people not identifying themselves as Jewish in the religion question in the census, although there was a general call by Jewish Board of Deputies to ‘tick the Jewish box’.

> ‘Because of the holocaust and other incidents older generations do not like disclosing their religion to official bodies. ‘

8.4.3 Lack of data on some protected characteristics and subgroups

There are several factors that NRS considered when developing the questionnaire for the 2011 Census. These factors apply both when deciding how many pages to include in the questionnaire, which questions to ask, and how many tick-box response options to include for each question. They apply to all questions, not just those relating to equality groups. The various factors include the impact on respondents (including the acceptability of questions and respondent burden); operational considerations (cost of printing, delivery and collection and processing); and time (for delivery, collection and for processing). All of these factors are dependent on the length of the questionnaire and the questions asked. A few stakeholders expressed their frustration that the reason given for not including more equalities categories was the lack of space on the census form to include this extra information.

Footnotes

22) The 2011 Census Recommendation Paper (PDF available on the Scotland's Census website) outlines the reasons for the design and inclusion of certain questions.

23) In a paper outlining the prioritisation tool that was created to decide which ethnic groups to include as tick-boxes and which to use write-in spaces (also used by NRS), ONS stated the ethnic group question was never intended to establish the ethnic’ composition of the population as it might be understood by sociologists, anthropologists and historians, rather, to capture in a common sense or pragmatic way the categories of person that were likely to be victims of ‘racial’ discrimination. (Moore and Hickman 2007).

> ‘The argument for not including the question is that there is not enough room in the census.’

> ‘We are repeatedly told that the box is not big enough.’

The coverage of the protected characteristics was also discussed by delegates and that in some cases there was little to no information collected for the characteristics.

> ‘We would like to have accurate statistics [on deaf sub-groups]. We have lobbied and campaigned for this. When we campaign for other issues we are asked for the numbers. We can only ever give rough figures.’

> ‘To us it doesn’t matter which way you gather information if we can access trusted evidence we will be on our way to getting true equality. The biggest barrier that we have is that we cannot get the numbers.’

A question on sexual orientation was not included in the 2001 or 2011 censuses due to concerns about individual privacy and public acceptability of including a question in a compulsory household survey and the quality of the resulting data. Prior to the 2011 Census NRS conducted a small scale postal survey as a first step to understand public attitudes to a sexual orientation question and the feasibility of mounting the question in a census. Census style forms were sent to a geographically representative random sample of Scottish households. In this test the percentage of respondents who did not provide useful data was around 31%. As this outweighs the percentage of respondents who declared a non heterosexual sexual orientation, the survey results called into question the accuracy of data gathered by such a question and hence the utility of any such data. Although sexual orientation questions have been included in some of the large social surveys, including the Scottish Health Survey, and Scottish Crime and Justice Survey, some delegates believed that the number of Lesbian, Gay, Bisexual and Transgender (LGBT) people are believed to be under-reported:

> ‘Whilst [our] estimate is 6-8% of the population in Scotland is LGBT, this reduces to 2% from employee records, and down to 0.8% in the large social surveys which ask a sexual orientation question.’

One delegate explained that not having detailed enough statistics limits the evaluation of their services.

> ‘It would be interesting to know how many people would benefit from using [the service] I don’t know how many people use it, who doesn’t have access to it, which people are waiting for equipment or which people have unused equipment. We need data to help improve outcomes. We need evidence to be able to say who it is benefiting from it and allocate properly our budget.‘

A number of delegates discussed how the absence of data collection can reinforce feelings of disadvantage and being invisible to the rest of society.

‘Although it can be difficult to capture information on non-settlement gypsy/travellers, they can take offence at being overlooked and excluded.’ [24]

‘Right now the biggest issue is that we are invisible. Companies and others do not think about deaf blind people. Our place in society is not reflected in statistics.’[25]

The misunderstanding that the number of deaf blind people cannot be calculated in the census highlights the need for education on the potential uses of the census data. User education will be an ongoing body of work throughout 2013 and beyond.

8.4.4 Statistical Disclosure

A delegate outlined the issue of statistical disclosure control to equality subgroups. Statistical Disclosure Risk (SDR) is assessed when statistics are published, based on information provided in confidence, that might allow an intruder to uncover the identity or attributes of a statistical unit (e.g. an individual, household, or business unit)[26]. The UK 2011 Census Statistical Disclosure Control (SDC) policy position is based on the principle of protecting confidentiality set out in the National Statistics Code of Practice which includes the guarantee that ‘no statistics will be produced that are likely to identify an individual unless specifically agreed with them’.

In a census context, where thousands of tables are generated from one database, the risk of attribute disclosure occurring can be addressed by introducing uncertainty about the true value of small cells. For the 2011 Census the following suite of methods will be used, including:

Footnotes

24) There has been a programme of liaison to target groups traditionally poorly enumerated in the census, with the aim of improving response rates from these groups. In the 2006 test a well-developed community liaison network was put in place which engaged the appropriate Scottish Government agencies, the local authorities, ethnic communities and organisations representing groups such as asylum seekers who were known to be housed in the test area at the time. Frequent contact with the gypsy/traveller community was a strong feature of this network.

25) The 2011 Census included a question on long term conditions with the options of ‘Deafness or partial hearing loss’ and ‘Blindness or partial sight loss’. Cross tabulating the responses to the question will allow for the number of deaf blind individuals to be calculated. There was also a question on the ability to ‘understand’ ‘speak’ ‘read’ and ‘write’ English, Scottish Gaelic and Scots. A further question on language other than English at home was also asked with the options for ‘English only’ ‘Yes, British Sign Language’ and ‘Yes, other’ with the option to write in other language. Questions about ability to read and write English should provide useful information on deaf sub-groups when combined with the question on language and other questions in the census.

26) More information on the reasons for potential disclosure and methods to prevent it can be found on the Scottish Government website.

- Modifying some of the data before the statistics are released through ‘record swapping’, where records with similar characteristics are swapped with a record from another geographic area.
- Where the number of people or households in a geographic area falls below a minimum threshold, the statistical output - except for basic headcounts - will be amalgamated with that for a sufficiently large enough neighbouring area;
- Restricting the number of output categories into which a variable may be classified, such as aggregated age groups;

For some more detailed tables, where the impact of disclosure control on the usefulness of the data is too great, special access arrangements will be put in place for approved researchers, as defined in the Statistics and Registration Service Act. A consequence of disclosure control methods is that the detail about very small groups can be lost from publications[27]. A delegate stated that minority groups are more likely to ‘slip through’ and not be included in outputs at the lower geographcial levels.

8.5 Alternative Sources Used

Delegates were asked about other data sources used in their work. A wide range of large scale surveys and administrative sources were used.

8.5.1 Surveys

Across the events and meetings, delegates were asked about any other surveys they possibly used to gather equalities information. The delegates listed a number of large social surveys such as the Scottish Household Survey (SHS) and the Scottish Health Survey (SHeS), the Annual Population Survey (APS) and the Labour Force Survey (LFS). However, many delegates felt that the sample size of these surveys was not considered large enough to derive any accurate equality findings and the lack of data at lower geographical levels limited the use of the surveys.

> ‘The Scottish Household Survey gives you a Scottish wide picture, but it is not useful at low geographical levels. The matrix of the SHS is very limited.’

Many of the delegates conducted their own research which was seen as especially important for the collection of information on sexual orientation and groups that were known to be hard to reach in the census and other surveys, such as Gypsy/Travellers.

> ‘Edinburgh College conduct 3 surveys per year for students and separate surveys for staff. The equalities section of these surveys asks for age, sex, ethnicity, sexuality and gender. ‘

Footnotes

27) Statistical disclosure control is applied to all official statistics under the UK Statistics Authority Code of Practice for Official Statistics (further information available on Scotland’s Census) website.

> ‘The census is good to refer to and good to know that there are [Jewish] people in every LA but we use our own anecdotal evidence. We know that if we put on an event or where we go Jewish people will come along.’

A delegate believed that the public may be more likely to disclose sensitive information in non-official surveys:

> ‘The accuracy of protected characteristic statistics can be affected by the way the data is collected. Some people will not answer controversial questions in official forms but they may do it in a survey.’

A number of delegates outlined a number of problems with Local Authorities (LAs) and other organisations collecting their own equalities data:

> ‘Capturing that information by LA surveys will be very expensive. Each question and response category costs money, we are limited in resources and we’ll make a compromise between the information we need and the resources. For instance, in our LA we will [restrict] ethnicity categories to just a few groups, or even sometimes just white and non-white (including maybe a specific group), and that does not give you the information you need for providing the right services to communities. There are other ethnicity groups that we know are important.’

8.5.2 Accessibility Issues relating to survey methods

A number of delegates also mentioned the various barriers that can prevent certain equality groups from fully accessing surveys. They emphasised the importance of looking at how surveys are carried out to understand if and how individuals might be excluded from taking part in them. They raised the issue of the use of telephone interview techniques potentially excluding deaf groups.

> ‘Anything that requires the person to communicate will be excluding for the groups we are interested in. For us the methodology of capturing data is very important.’

8.6 Administrative Data Sources

Delegates outlined a number of administrative sources that provide information on equality groups. Delegates working in education mentioned a number of administrative sources including Higher Education Statistic Agency (HESA) data on School Leaver Destinations, School census, School Pupil Survey, Scottish Qualification Awards (SQA) results, and University and College Admissions Service (UCAS)

statistics which include information on age, ethnicity and disability. Jobcentre Plus statistics were also used for information on ethnicity and jobseekers allowance.

8.6.1 Coverage Issues

Issues concerning the coverage of administrative data sources was mentioned by a number of delegates:

> 'Using admin data sources which are supplemented by surveys brings up issues of coverage and representativeness…hard-to reach groups tend to get missed as they're on the margins of society and you're not going to get them by relying on admin data and surveys.'

> 'Ethnicity information is complete for around 90% of in-patient records. It can be more difficult to get GPs to capture ethnicity information. Hospital admissions capture some marginalised individuals but the data is limited to age, sex, ethnicity and linguistics.'

> 'We lose a high percentage of the schools population [in the School Census], as private schools do not provide the data.'

8.6.2 Quality Issues

There were a number of concerns about the quality of data relating to equalities characteristics derived from administrative sources:

> 'Although social workers are supposed to compile equality group information they are often under too much work pressure to be concerned with the quality of the data.'

8.6.3 Accessibility and data sharing

The delegates had concerns about the accessibility of administrative sources:

> 'Other departments within our LA may have the information we need, and although there is a feeling that sharing it should be the way forward, there is also a concern with lack of compliance when sharing or to release information we shouldn't.'

> 'I get the feeling the information we need is being collected but we don't know where it is, how it is collected or how to access it. Probably schools and social work departments collect related information but we don't have access to it. Even if we had it we are not connected very well, so an overall picture of the problem is not available.'

> 'There's a lot of information to be harnessed through record linkage…the opportunity to discover a wealth of additional information that is not captured elsewhere, however there is a tradition of not sharing data in the NHS.'

A delegate also expressed concern about the classifications in different administrative sources and the problems that this could cause for data linkage:

> ‘There can be differences between primary and secondary care information as they are operated by independent contractors – there is no transferability between the two.’

8.6.4 Qualitative Research

A number of delegates also mentioned that they use a mixture of quantitative and qualitative evidence.

> ‘At this moment, from an equality perspective, we rely on ad-hoc reports and in qualitative research. I belong to a multi equality forum which is useful to understand what is important for the different equality groups and contrast that with official information but again this is qualitative information.

> ‘The community interest bank[28] is another source of qualitative data.’

Qualitative information can reveal some unique information about sensitive information such as the experience of discrimination and hate crime:

> ‘Qualitative studies have revealed that being LGBT in a rural area is different and they are more likely to face discrimination and hate crime.’

Footnote
28) PDF document available on the Scottish Government website.

8.7 Future Options

A number of groups also gave opinions on the possible future methods for producing population and socio-demographic information. A number of delegates were positive about an increased use of electronic census forms:

> 'Digital inclusion is the way forward…the completion of electronic census forms is a good move but you need to be careful not to exclude anyone and to make sure that technology is accessible.'

Some group members were interested in the short form/long form census method. They were particularly interested in the potential of developing surveys which contained a 'spot light' feature whereby questions were concentrated on a particular theme (i.e. having a transport feature which perhaps included questions on how the disabled travelled) or adapted to target a particular equalities group in a specific area.

> 'We know that there are concentrations of deaf people within geographical areas relating to specific industries such as in the North East of England and Glasgow. Is there a possibility of asking different questions in these areas?'

8.8 Future Information Priorities

The Beyond 2011 programme is looking into how much detailed information on equality groups can be collected from the census or whether it could be within the remit of other general surveys and administrative sources. The need for specificity and greater detail versus the cost of collecting this information and the chances of being able to release it given disclosure control constraints would need to be balanced.

The delegates discussed their future information priorities and the majority of delegates agreed that the ability to aggregate or disaggregate information on the protected characteristics was important.

> 'LA information on all protected characteristics will be an improvement on what we have now. If it's done below the LA it helps focusing the services in the areas that need them.'

> 'We'd like to see questions referring to all protected characteristics.'

A number of delegates across the groups said that they needed more detailed information on the disability question.

> 'We should be able to drill down the disabled group at least by type of disability, speech, sight, sensory.'

> 'We need a higher breakdown in disability.'

A number of delegates said that they wanted more detailed information on the different deaf groups and different communication disabilities:

‘There has to be work on the disability question and more specifically on the deaf question to capture BSL (British Sign Language) and deafblind.’

‘In the Northern Ireland census there was a list of disabilities which contained an option to select ‘communication disability’ or ‘communication support needs’. Some people have receptive communication disabilities. I think it might be of benefit to have a question like “do you have difficulty making yourself understood or understanding others?”.’

‘There is a difference between short term and lifelong communication difficulty. I want to be able to deduce that ‘X per cent of people with cerebral palsy use our services, for example. Breaking down the type of communication needs is important.’

As the census does not capture information on LGBT, a number of delegates have requested that any future method should provide information on sexual orientation.

‘We need an LGBT question. There is no reliable source for this information and we have to estimate it based on the number of customers of organisations such as Stonewall.’

Delegates also emphasised the necessity for detailed information on ethnic groups.

‘For what we use the census for, data categories are not precise enough for the groups we require, and that’s important to provide the right services. Black/White category is too wide a separation, we need sub-categories by country of birth, as that sometimes determines how affluent that group is and how efficiently we can target services. There are also health related issues with each ethnic group by country of origin.’

8.8.1 Geographical Level

For many of those delegates involved in service planning, the ability to provide information at different geographical breakdowns was seen as an equally important aspect of aggregation and disaggregation.

‘Obtaining information at LA level is good, but lower geographical levels, the ward, is also of interest. We need to know some information at local level, for instance the language spoken in the schools, particularly in deprived areas.

‘In terms of geographical level, equality tends to be council wide but community councils also need to understand their population profile. People need to be reminded that there are areas within their city that are not as affluent as the average.’

‘We tend to use national level data but Local Authority level when making funding applications.’

> ‘The numbers at LA level are so small that you end up arguing in terms of number rather than principle. It is difficult to justify a service for less than 5 people. For such small numbers these issues are best solved at a national level. ‘

Although national level statistics can help inform policy decisions, the experience of individuals with protected characteristics can be considerably different between cities and rural areas and there is also a need for information at a local level to monitor outcomes across Scotland.

> ‘Local Authorities are basing policy decisions on national level statistics but the experiences of LGBT are different from cities to rural areas. Stonewall have qualitative evidence that in rural areas LGBT have a different experience of being gay. There are less clubs, etc. and it’s more of an online experience […] In rural areas there is a much greater difficulty coming out, and they are more likely to face discrimination.’

8.8.2 Frequency

The majority of delegates agreed that the frequency with which they receive population and socio-demographic information should be increased.

> ‘2001 Census data feels completely out of date now. Having newer data right now, even sacrificing certain accuracy, will be helpful right now.’

The majority of delegates across the discussions agreed that the ideal frequency of population and socio-demographic information should be around every five years. Across the groups, many delegates stated that having the population and socio-demographic information in line with their planning processes was the priority.

> ‘To conduct a census every five years would be ideal as the College produces its Strategic Plan every three years and relies on the census for this. ‘

> ‘It would be better to gain information more in line with the outcome process.’

> ‘The changes our work leads can take three years or more to implement, so it is not necessary for us to have more frequent statistics than that.’

The frequency of data collection was considered to be important to the accuracy of the data. As ‘people can be sceptical about the accuracy of Mid Year Estimates, high quality information every five years, with a tiered approach of something in-between to act as a marker’, was seen to be the ideal level of frequency for some delegates.

> ‘Five years would probably be a reasonable window. In two years there could be a spike but five years allows some smoothing.’

8.9 Implications of not having the Census

The groups discussed what the main implications would be if the census was not available. Delegates felt that the main impact would be that they might not be able to meet their Equality Duties.

> 'The NHS uses information from the census to help meet its legal requirements under equalities legislation.'

A number of delegates who rely on census data for service provision were concerned that they would not be able to effectively plan services and allocate resources.

> 'Without the census, it would be difficult to know what resources were required and to protect funding.'

> 'If there was no census, it would be difficult to demonstrate and justify why the funding for projects and for small voluntary organisations was needed.'

In terms of other legal implications, the delegates reiterated the legal requirement to monitor and assess equalities information. Some delegates claimed that they could rely on other data sources to meet SOAs.

> 'If the census was not available it would not be that bad as [we] tend to use other data sets in the council for SOAs. It would be more upheaval if the census as the only source of information.'

> 'SOA is high on the agenda […] we could do without the census as long as there was something on the population.'

Some delegates reiterated concerns about the accuracy of alternative sources to the census:

> 'The consequences of not having a census in 2021 would mean that equality statistics would need to be gathered from local sources. Over the years, this data could become wildly inaccurate.'

9. Conclusion

The discussions that were held with delegates were very useful and allowed the Beyond 2011 team to see how the census is used across Scotland. Delegates emphasised the importance of equalities information derived from the census to their work.

The contents of this report will be added to the bank of information currently being collected by the Beyond 2011 team to aid the Equality Impact Assessment for the programme. It will also be used in the evaluation of any possible methods for the future provision of population and socio-demographic information in Scotland.

Delegates will also be kept informed about any future events relating to Beyond 2011 and were sent a link to the Beyond 2011 formal consultation in Spring 2013.

Appendix A – 2001 and 2011 Census questions relating to the protected characteristics

2001 Census	2011 Census
Gender	
What is your sex? Male Female	What is your sex? Male Female
Age	
What is your date of birth? day/month/year	What is your date of birth? day/month/year
Marital status	
What is your marital status (on 29 April 2001)? Single (Never Married) Married (first marriage) Re-married Separated (but still legally married) Divorced Widowed	On the 27 March 2011, what is your legal marital or same-sex civil partnership status? Never married and never registered a same-sex civil partnership Married Separated, but still legally married Divorced Widowed In a registered same-sex civil partnership Formerly in a same-sex partnership which is now legally dissolved Surviving partner from a same-sex civil partnership
Disability	
Over the last twelve months would you say your health on the whole has been: Good? Fairly good? Not good?	How is your health in general? Very good Good Fair Bad Very bad
Do you have any long-term illness, health problem or disability which limits your daily activities or the work that you can do? * Include problems which are due to old age. Yes No	Are your day-to day activities limited because of a health problem or disability which has lasted, or is expected to last, at least 12 months? *Include problems related to old age. Yes, limited a lot Yes, limited a little No
	Do you have any of the following conditions which have lasted, or are expected to last at least 12 months? * Tick all that apply. Deafness of partial hearing loss Blindness or partial sight loss Learning disability (for example, Down's Syndrome) Learning difficulty (for example, dyslexia) Developmental Disorder (for example, Autistic Spectrum Disorder or Asperger's Syndrome) Physical disability Mental health condition Long-term illness, disease or condition Other condition, please write in

Religion	
What religion, religious denomination or body do you belong to? None Church of Scotland Roman Catholic Other Christian, Please write in Buddhist Hindu Jewish Muslim Sikh Another Religion, please write in	What religion, religious denomination or body do you belong to? None Church of Scotland Roman Catholic Other Christian, Please write in Buddhist Hindu Jewish Muslim Sikh Another Religion, please write in
What religion, religious denomination or body were you brought up in? None Church of Scotland Roman Catholic Other Christian, Please write in Buddhist Hindu Jewish Muslim Sikh Another Religion, please write in	
Race (Ethnicity)	
What is your ethnic group? A White Scottish Other British Irish Any other White Background, please write in B Mixed Any Mixed background, please write in C Asian, Asian Scottish or Asian British Indian Pakistani Bangladeshi Chinese Any other Asian background, please write in D Black, Black Scottish or Black British Caribbean African Any other Black background, please write in E Other ethnic background Any other background, please write in	What is your ethnic group? A White Scottish Other British Irish Gypsy/ Traveller Polish Other white ethnic group, please write in B Mixed Any Mixed or multiple ethnic groups, please write in C Asian, Asian Scottish or Asian British Indian, Indian Scottish or Indian British Pakistani, Pakistani Scottish or Pakistani British Bangladeshi, Bangladeshi Scottish or Bangladeshi British Chinese, Chinese Scottish or Chinese British Other, please write in D Black, Black Scottish or Black British Caribbean, Caribbean Scottish or Caribbean British African, African Scottish or African British Other, please write in E Other ethnic group Arab, Arab Scottish or Arab British Other, please write in

Appendix B – Alternative Data Sources Used

Administrative Sources

Scottish Qualifications Authority (SQA)
Register of SASINES (ROS)
Community Health Index (CHI)
University and College Admissions Service (UCAS)
Jobcentre Plus
School Census
GP Services Data
Benefits Data
National Health Service Central Register (NHSCR)
Council Tax records
Her Majesty's Revenues and Custom (HMRC)
Driver and Vehicle Licensing Agency (DVLA)
Electoral Register
Citizen's Account
Employment statistics from the Department for Work and Pensions (DWP)
Higher Education Statistics Authority (HESA)
National Address Gazetteer
Corporate Address Gazetteer
Property Records
Service Records
Maritime records

Surveys

Business Register and Employment Survey (BRES)
Family Resource Survey
Scottish Household Survey (SHS),
Scottish Health Survey (SHeS)
Scottish House Condition Survey (SHCS).
Voluntary Population Survey
Health and Language Needs Survey
Labour Force Survey (LFS)
Annual Population Survey (APS)
School Pupil Survey
Local and ad hoc Surveys

Management Information

CACI (Marketing and Information solutions)
Education Management Information Software (SEEMIS)
Social Work Information System (SWIFT)
Local Authority Customer Relationship Management (CRM)

Monitoring Tools

Scottish Longitudinal Study (SLS).
Scottish Neighbourhood Statistics (SNS)
Scottish Index of Multiple Deprivation (SIMD)
Scottish Public Health Observatory (ScotPHO)

Other Government Sources
NOMIS (official labour market statistics)
Community Innovation Statistics (CIS)
Sport Scotland
Other NHS statistics

Other
Citizen panels
Fuel Poverty Forum
Experian (Commerical)