

2018 Census External Data Quality Panel: Minutes of Meeting on 6 March 2019

Date and time	6 March 2019, 9am to 3:30pm
Location	Auckland University of Technology 16 th floor, 56 Wakefield Street Auckland 1010
Present - panel members	Richard Bedford - Chair Tahu Kukutai Donna Cormack Alison Reid Thomas Lumley Len Cook Barry Milne Ian Cope
Present – Stats NZ	Kathy Connolly, General Manager Census Vince Galvin, Chief Methodologist Gareth Meech, Senior Manager Census and Secretariat Steph Prosser, Senior Analyst Census Adele Quinn, Manager Census Analytics Christine Bycroft, Principal Statistician Chris Hodgins, Manager Census Geography (afternoon session via phone only)
<p><u>Meeting minutes</u></p> <p><u>Review previous minutes and action points</u></p> <p>The panel confirmed that the minutes contained a fair summary of discussions at the February meeting.</p> <p>The Chair and secretariat worked through the action points from the February meeting, removing those that were now closed.</p> <p>Alison Reid has been confirmed as co-chair who will work with Richard to ensure the work of the panel continues in his absence from May 2019.</p> <p>The panel noted that hosting some meetings in Wellington would be a positive step.</p> <p>The Action Points have been updated in the table at the end of this set of minutes, but there were some specific conversations regarding the Privacy Impact Assessments (PIAs) to be noted:</p> <p>AP3-10: Carol outlined the work she has been doing to update the PIAs, making them consistent and more accessible. A statement about the Stats and Privacy Acts, making it clear that we are operating under the exemption that applies to statistical and research use, will be added to an appendix before the April announcement.</p> <p>The panel suggested the statements about PIAs should be linked on the website and made more readily available. It was also noted that there are some broken links to previous documents and it should be clear which documents have been superseded.</p>	

Report back from 27 Feb in-committee meeting and suggested changes to Terms of Reference

At the 27 Feb meeting, the panel spent time discussing how to provide the best value. Ideas included creating a set of materials that reflect on issues discussed as a more formal way to document views. These would not be formal reports and may be just framed by a set of questions. The panel is currently preparing a series of short papers for Stats NZ on iwi data, privacy and confidentiality and one on what the census is used for and the challenges faced.

The discussion moved to the terms of reference. A proposed updated version was sent to the Chair just prior to the meeting and requires the Chair to review the TOR before being presented to the panel. At the in-committee meeting, the panel agreed the panel should continue to meet until the around the time of the first release of census data, but wanted it acknowledged that the extension will not be easy for panel members as they have busy timetables. There was also concern expressed that the longer that the panel continues, the more difficult it becomes for individual members to keep silent about what is being discussed.

Carol commented that it is vital for the panel to continue to meet until first release of census data and that the advice being given by the panel has been extremely important and influenced many decisions. Stats risks making decisions in isolation and becoming too insular without the regular panel meetings.

There was also a discussion at the 27 Feb meeting about what constitutes official statistics.

Census programme timeline update

Kathy presented a slide on the timeline we are targeting noting that not much has changed since the February meeting. There were questions from the panel about whether the timeline remains feasible. The Stats NZ response indicated increased confidence in the methods being implemented but it was noted that the timeline still has risk attached.

Methods session

Christine reviewed in some detail the methods paper that had been uploaded to the secure site shortly before the meeting. Christine emphasised that we are people missing from households were only being replaced when there was a high level of confidence that the replacement matched the missing individual.

There was quite a wide-ranging discussion about the communication challenges Stats faces, including comment on:

- What are the key questions the users are going to have in April?
- What are the key challenges?
- If the external reviewers publish their report in July, how will this affect the messaging at the time and at the time of first release?
- Depending on the timing of the public discussion about what went wrong and why, this might weaken the impact of the published census data and some positive results arising from the accelerated use of admin data sources.
- Stats should produce more information in April about the quality of what is being produced by the 2018 Census.
- With Stats being silent in the media, the media themselves only have the attackers to talk to, as there is no stated defence.
- When discussing the changed definitions of responses, how will people know that they can trust Stats?

- The NZ public did not have a discussion about the extensive use of admin data in the 2018 Census. Stats noted that greater use of admin data was flagged, but not admin enumeration. How confident is Stats that the public will understand and be comfortable with the use of admin enumeration? There are risks that the public will not be happy with admin enumeration of family structures, parents and children.
- Stats needs to up front with the public about what linking has been used inside and outside the IDI 'spine'
- What is happening with the use of admin enumeration for people in NPDs?
- More elaboration is needed from Stats about:
 - o The distribution of admin records around the collection date
 - o Summary of who is on the margin
 - o With item imputation – how much is the method stereotyping people eg. smoking
 - o That there will be no unit donor imputation (which is good) but will still use it for item imputation
- The people in admin enumeration are different eg. more Māori, Pacific and migrants
- Stats NZ need to clarify some of the detail around administrative data such as:
 - what we do from both a usual residence and census night perspective
 - how up to date is administrative address information
 - when do different administrative data sources get updated
 - why we put some people in a meshblock rather than a household
 - how does the administrative prioritised usual residence address get created
 - how we will handle administrative addresses in 2023 as a planned activity as there are lags in the address information coming through
 - how many of the people who were included in the the census file from administrative sources were from Auckland

During a discussion about how the new admin enumeration methods may improve the key demographic variables, the panel requested an update on the quality of other variables where there are no other data sources – smoking being an example.

AP6-1: Ask the chair about whether the panel would benefit from a data quality overview of a variable with a limited admin enumeration component..

One of the panel members ran through a description of the regularly updated counts slide so that all panel members had a shared understanding. It was noted that the admin data sources will not neatly fit the “census day” point of collection. For example, some admin data will be from sources containing data generated at a range of dates around census day. Some agencies don't update their IDI data frequently enough. Stats staff will need to explain which admin sources were used and any caveats about their quality and timeliness that data users will need to know.

A question was also raised about how confident Stats staff are that people have been placed in the right households and meshblocks (including for electoral purposes). Christine responded that where we do not have confidence in placing people in households we have placed them in a meshblock. Stats staff noted that they have been working with the electoral commission on Central Auckland's population and electoral enrolments. There are a number of possible explanations, including high numbers of hard to reach students and Auckland being cited as the local destination on immigration cards when this not necessarily the case.

The panel asked about how the methods have treated all population groups – and that Stats should produce a paper that covers all segments. Stats responded saying there are a number of

internal technical advisory group papers that cover several data areas. One panel member noted there is limited understanding on the quality of household and family data. The panel suggested Stats needs to summarise and explain all the changes in a technical report and then work on the key messages.

The panel discussed whether users want to see information relating to change between census or is there more interest in actual levels and frequencies in the data. In other words, how have the characteristics of a set of people changed compared to level relating to those characteristics?

The panel discussed whether it is clear about what has and has not been captured in the admin data as a result of the methodology decisions made. Of interest here is understanding the characteristics of meshblock admin enumerations compared with the characteristics of enumerated meshblock populations without the addition of admin data.

Prison and defence force population data

Stats described their approach to using Department of Correction and Ministry of Defence admin records to supplement census forms. Stats is planning to use prison and defence admin data for key demographic information but not about iwi due to the quality of the iwi data. Some members of the panel expressed serious concerns about the use of prison authority data in place of other methods as there are additional issues around control and access to people's data. Coercion can be an issue and some members were very uncomfortable with the ethical implications and possible challenges. Stats responded by informing the panel that a PIA will be published that explains our use and that we have been very keen to use as many good quality admin data sources to improve data in this census.

The panel requested a methods document that covers a range of topics including the method, PIAs, iwi data decision and 'what is a response?' – although some may need to be split.

AP6-2 Present list of papers and approx. completion (or draft) timing so panel can base own reports off Stats documents (rather than meeting minutes).

Programme benefits and performance indicators

The panel were keen to understand performance measures, especially in terms of benefits. Carol presented brief slides on the topic noting that response rates will look different as this is a proxy-transition to an admin based census model. Coverage calculations will change to reflect the new model. Carol was in the process of reworking the benefits, given the failures, to determine where the new focus should be to ensure benefits are still delivered.

The panel noted that it would be a mistake for Stats to say the response rate is higher, and that it should 'take the hit' early to keep it away from the census data releases.

Stats discussed how it has referred to response rates since census day.

Stats noted that there were a lot of KPIs and we should have focused more on the key count and undercount KPIs.

Dwelling frame quality

Chris Hodgins dialled in from Christchurch to answer questions about the paper provided to the panel on the dwelling frame. There was a question about the occupation of holiday homes – were there registers available to help? Stats explained that a cautious approach was taken in the field,

documenting validated numbers and field changes to give confidence to users. It was noted that a few paragraphs need to be added to the overview paper about the dwelling frame quality.

AP6-3 Add numbers on 'address not found' to the table at the end of the dwelling frame paper.

Quality plan update

Gareth gave a short presentation on the progress of quality assurance including some more information about the variable quality rating scale. Gareth would send a TAG reviewed paper on the subject to the panel prior to the 12 April meeting. The sensitivity analysis slide attracted questions about what else is possible to be completed. Local government cases were discussed and whether a smaller council such as Gisborne could be approached alongside larger councils such as Auckland and Christchurch.

AP6-4 Upload variable quality rating scale paper to workspace before 12 April.

The conversations then started to refer to the variables and admin enumeration sources spreadsheet that was provided to the panel prior to the meeting. The panel found the document useful, but Stats noted it is draft and there are some errors including a reference to MOJ sourced data which needs to be checked and corrected if necessary.

The panel asked questions about whether Stats is doing enough to look at how close the census response data is to the admin data used to increase coverage. Christine noted work on this, including statistical assessment of relationships between the two data sets, is being planned.

Kathy gave an update about the meeting with a group from the iwi chairs' forum. The panel noted that equity of access is important when talking with some but not all stakeholders. Not all iwi are represented by the iwi chairs' forum. All iwi are highly likely to be planning to use 2018 Census information for decision-making on and they need to know as soon as possible what data may be available.

The additional slides with updated Māori ethnicity and descent data generated several questions which were addressed by Adele. The panel asked for an update on electoral Māori descent data which is not yet available. A draft paper on ethnic mobility written by Robert Didham will be uploaded to the workspace before the next meeting.

AP6-5 Upload ethnic mobility paper to workspace.

The next meetings dates are:

- in-committee only 1 April, Auckland
- 12 April, in Christchurch
- 22 or 23 May in Auckland
- June date to be confirmed, but Wellington venue to fit in with PANZ conference.
- July TBC

Action log

Ref	Date raised	Description	Owner	Date required	Progress	Status
Meeting 2 actions						

AP 2-4	23/10/18	More information requested about what investigations we are doing into 2013 counts and response rates – are we assuming that 2013 is correct?	Panel	By Dec meeting	<p>9/11 Information presented that include 2013 response rates and counts. Gareth to check with Richard whether this should be closed by 23 Nov.</p> <p>6/12 Gareth to discuss with Richard</p> <p>18/2 Stats may need a specific 2013 methods paper to document any inconsistencies in approach and will be captured in methodology papers and metadata documentation. Gareth to upload proposed method papers (completed 18 Feb). Panel to review paper and offer comments by 6 March.</p> <p>11/4 Summary spreadsheet of proposed papers to be published uploaded on 10 April. Feedback needed from panel by 30 April.</p>	Open
AP 2-6	23/10/18	Need more detail about the totality and components of the imputation methods initially planned and now being implemented.	Gareth	By 9 Nov	<p>Sessions at 30 Oct meeting and priority at full 9 Nov meeting.</p> <p>9/11 Information presented at meeting. Gareth to check with Richard whether this should be closed by 23 Nov.</p> <p>6/12 Gareth to discuss with Richard</p> <p>18/2 Spreadsheet detailing variables and imputation processes to be sent to panel prior to 6 March meeting.</p> <p>6/3 Methods paper and variable source spreadsheet sent to panel.</p> <p>Confirm with Richard whether this item can be closed before 12 April in committee meeting.</p> <p>11/4 Gareth to check with chairs.</p>	Open
AP 2-10	23/10/18	Request for data on birthplace.	Christine	By Dec meeting	<p>Topic for December meeting</p> <p>9/11 To be presented at Dec meeting.</p> <p>7/12 Data added to workspace by 7 Dec</p> <p>18/2 Richard to discuss further with Christine</p> <p>6/3 Richard pop insights set of data showing net gains and losses by birthplace over inter-censal period Mar 2013 to Dec 2017. Not</p>	Close

					on workplan, but planning further work on using admin data more effectively. Country of birth on list. This request will be dealt with outside the panel.	
AP 2-16	23/10/18	Re electoral implications, more information is required about the representativeness of the data and impacts of imputations. A range of scenarios, including constitutional impacts for Māori, would be useful	Kathy and Gareth	By 9 Nov	Discussion planned for 9 Nov meeting. 9/11 No additional electoral data was presented (as not complete yet). Updated draft electoral calculations to be presented to panel in Dec meeting. 6/12 Electoral data has not yet been run – targeting Feb meeting update. 18/2 Updated electoral data to be presented to 6 March meeting 6/3 Data is not ready for presentation to the panel. To be provided to the panel before or at the 12 April meeting. Note that Dot loves Data has been contracted to look at the sensitivity of the detailed electoral calculations. 11/4 Slides to be included for April meeting presentation. Suggest Close.	Open
Meeting 3 actions						
AP 3-5	9/11/18	More information about what admin data is used in the method and for what purpose. Privacy Impact Assessment update. Add to agenda for December.	Gareth	By 7 Dec	6/12 Did not have time for this item in the Dec agenda. Will present information in Feb. 14/2 Requested list for all variables about what commitments we have made for use of admin data in the past and now. Rolled into AP2-6. 6/3 As AP2-6 and 3-10 cover this action and are still open, close.	Close
AP 3-10	9/11/18	Clarification is needed about what data is used in the census administration enumeration (referred to as IDI data in the presentation), and a list of the current variables being imputed, where it comes from, whether the data suppliers are aware of its use and whether it uses identifiable or de-identified person data.	Panel members	By 7 Dec	6/12 Leave open, not addressed yet 14/2 PIA includes detail, but hard to follow. Suggest adding a flowchart to aid communication. Panel to review the PIA and give feedback. 6/3 One panel member had provided feedback. Carol stated that the PIAs on the website require further revision work has started	Open

					on this. Has worked with legal counsel on creating a generic statement about the legal basis of the use of admin data. A statement will go into an appendix before the (likely end) April announcement. Leave open. 11/4 Legal basis document added to shared space. Suggest close.	
AP 3-11	9/11/18	Present criteria about the thresholds being used to make decisions about methods.	Gareth and Christine	By 7 Dec	6/12 Some information in Dec meeting present, more info to come after meeting. Leave open. 14/2 Panel will be using Canadian quality framework. Request more information about quality plans at March meeting. 6/3 An update about the quality measures was presented at 6 March meeting. Variable quality rating scale paper to be sent to panel before 12 April meeting. 11/4 Variable rating scale paper loaded to workspace 10 April. Will need time to discuss suggest on agenda for May meeting. Leave open.	Open
AP 3-14	30/10/18	The Dwelling Frame evaluation report to be shared with the q/a panel	Adele	14 Feb 2019	14/2 To be uploaded to shared space by 6 March. 6/3 Uploaded to workspace	Close
AP 3-15	30/10/18	Provide information on births and removals from the Dwelling frame during the live operation	Adele/ Chris H	14 Feb 2019	14/2 Some information was presented to the panel in November. More to be loaded to workspace before March meeting 6/3 Uploaded to workspace	Close
Meeting 4 actions						
AP 4-5	7/12/18	Document how multiple de-identified data points in admin sources have been protected from identification once they are assembled.	Christine	By 14 Feb meeting	14/2 Assign to methods team. Confirm priority with Richard. 6/3 Panel sent paper prior to meeting and meeting session discussed at length. Close	Close
AP 4-6	7/12/18	Document 'What is the legal basis that it is OK to do what we are doing with admin data and what are the ethical implications we have considered?' 'What commitment has Stats NZ given to respondents and data suppliers about what we are doing with admin	Carol	By 14 Feb meeting	14/2 Although Liz discussed the issues with the panel at meeting on 5 Feb, panel still wants a statement from Stats about the legal basis of admin sources. Without, it is a risk for the panel to endorse. High ethical standard set by the	Open

		data? What have they said? What are the key data sources?’			Stats NZ at the 2017 data summit. 6/3 This matter continues to be a source of concern to panel and Stats is addressing some of these concerns in the variable list (see AP 3-10) 11/4 On agenda for April meeting. Suggest close after meeting.	
AP 4-8	7/12/18	Consult over when Stats NZ will return to the panel with a decision about public release of iwi data	Richard and Gareth	By 14 Feb meeting	14/2 Stats to cover in meeting on 14 Feb. Check with Richard if enough information to close. 6/3 iwi data leaders meeting key messages sent to panel. Decision made that Stats will not talk publicly about data quality before the April announcement. However, need to have conversation with iwi before then. Leave open until iwi data messaging is confirmed. Panel is writing up document to be sent to Stats about panel view about quality of iwi data. The panel noted that the 2018 Census webpage has a link to information about iwi that implies it will be available. Stats to review and update. 11/4 On agenda for April meeting. Suggest close after meeting.	Open
AP 4-9	7/12/18	Update panel once more detail has been completed on the quality framework including an ordering of decision making criteria.	Gareth	By 14 Feb meeting	14/2 Planned item on 6 March agenda. 6/3 Variable quality rating scale paper to be sent to panel before 12 April. 11/4 Paper uploaded, set time in May agenda. Leave open.	Open
AP 4-13	7/12/18	Create a table to summarise variable quality measures.	Gareth	By 14 Feb meeting	14/2 Did not have time to present, planned for 6 March meeting. 2/6 Added to workspace	Close
AP 4-17	7/12/18	Check panel members availability for April meeting.	Gareth	By 21 Dec	14/2 Meeting is booked to 12 April in Auckland, but will check with panel members via email on 22/2. 6/3 Proposed venue for April meeting is Christchurch. Logistics to follow. Close.	Close

Meeting 5 actions						
AP 5-1	14/2/19	Need to clarify admin enumeration, admin imputation and other terminology. Thomas created a short paper of definitions to be reviewed by Stats NZ.	Adele, Christine	By 6 March	6/3 Thomas has written a paper for the panel about definitions of admin vs survey data. Kathy to send to SNZ staff for review. 11/4 Written response to be developed by Stats for each paper written by panel. Date to be confirmed. Leave open	Open
AP 5-3	14/2/19	Load technical papers presented to internal Technical Advisory Group to the shared space including approach to prison and defence admin enumeration.	Gareth	By 6 March	6/3 several papers added, more to come. Any specific requests welcome	Close
AP 5-4	14/2/19	Additional ways for information to get to the panel during the rapid changes	Gareth	By 6 March	Regular emails each week	Close
AP 5-5	14/2/19	Find and load 2018 Census business case document and valuing the census to the shared space	Gareth	By 6 March	6/3 Added to workspace	Close
AP 5-6	14/2/19	Revise terms of reference mainly around the mechanics of the panel purpose, then for the panel to review and sign-off from Liz and Richard	Richard	By 6 March	6/3 Sent to Richard for initial review. Richard to review by 1 April. 11/4 Richard and Alison reviewed and spoke with Carol and Liz. To be confirmed at April meeting, effective from 29 April. Suggest close.	Open
AP 5-7	14/2/19	Panel members to add links to the previous research and papers on Māori descent and ethnicity they already have and provide any further feedback to Stats NZ.	Tahu and Donna	By 6 March	6/3 Link added to shared documents	Close
AP5-8	14/2/19	Include 'what is a response' paper in the list of method papers that would assist the panel to write the reports	Christine	By 10 April	6/3 No update 11/4 Paper uploaded to shared space. Suggest close.	Open
Meeting 6 actions						
AP6-1	6/3/19	Provide the panel information on smoking question data as an example of a variable without other admin enumeration sources.	Adele	By 12 April	11/4 Unable to provide counts of lower priority variables for April meeting. Suggest smoking data and associated quality rating topic for May meeting.	Open
AP6-2	6/3/19	Present list of papers and approx. completion (or draft) timing so panel can base own reports off Stats documents.	Gareth	By 12 April	11/4 Updated list loaded to shared space. Suggest close.	Open

AP6-3	6/3/19	Add numbers on 'address not found' to the table at the end of the dwelling frame paper	Chris H	By 12 April	11/4 To be completed by end of April. Suggest keep open.	Open
AP6-4	6/3/19	Upload variable quality rating scale paper to workspace before 12 April.	Gareth	By 12 April	11/4 Uploaded to space. Suggest close.	Open
AP6-5	6/3/19	Upload ethnic mobility paper to workspace	Gareth	By 12 April	11/4 Original uploaded on 29 March. Updated version 2 uploaded on 10 April. Suggest close.	Open

Released under the Official Information Act