



Office for
National Statistics

**3rd Meeting of the Methodological Assurance Review
Panel**

Agenda & Minutes

**24 & 25 June 2019, Office for National Statistics
London**

Office for National Statistics Methodological Assurance Review Panel Meeting

Agenda

24 & 25 June 2019

Drummond Gate London

Chair: Sir Bernard Silverman

Day 1

Time	Item	Presenter
1 10.30-11.15 45mins	Introduction & Actions from previous meeting	Rachel Skentelbery
2 11.15-12.45 90mins	eQuestionnaires	Steve Woodland
12.45-13.15 30mins	Lunch	
3 13.15-13.45 30mins	Maximising Response Strategy (EAP113)	Orlaith Fraser
4 13.45-15.15 90mins	Response Chasing Algorithm, Field Prioritisation Algorithm & Response Profiles (EAP114, EAP115 & EAP116)	Orlaith Fraser & Victor Meirinhos
15.15-15.30 15mins	Break	
5 15.30-17.00 90mins	Field Operations Simulation (EAP117)	Orlaith Fraser & Kim Ward
17.00-17.20 20mins	Summaries & Actions	Rachel Skentelbery

Day 2

Time	Item	Presenter
6 09.05-10.35 90mins	Quality Assurance Strategy (EAP118)	Pete Large
10.35-10.45 10mins	Break	
7 10.45-11.45 60mins	Statistical Design for the Population Statistics Transformation	Andy Teague
8 11.45-12.00 15mins	Update on Data Linkage Advisory panel	Sarah Henry & Bernard Silverman
12.00-12.30 30mins	Lunch	

9 12.30-14.00 90mins	Integrated Population and Characteristics Survey (IPACS) (EAP119)	Pete Jones
10 14.00-15.15 75mins	Enhanced Outputs – VOA data (EAP120)	Fern Leather & Andy Mealor
15.15-15.30 15mins	Summaries & Actions	Rachel Skentelbery

Attendee List

External Panel Members

Sir Bernard Silverman (Chair)
Prof Natalie Shlomo (External Panel Member)
Prof David Martin (External Panel Member)
Dr Nik Lomax (External Panel Member)
Dr Oliver Duke-Williams (External Panel Member)

Office for National Statistics

Rachel Skentelbery (Vice-Chair, Chief Methodologist)
Cal Ghee (ONS Panel Member)
Owen Abbott (ONS Panel Member)
Sarah Henry (ONS Panel Member)
Jon Wroth-Smith (ONS Panel Member)
Steve Woodland (Presenter)
Charlotte Hirst (Presenter)
Orlaith Fraser (Presenter)
Victor Meirinhos (Presenter)
Kim Ward (Presenter)
Pete Large (Presenter)
Andy Teague (Presenter)
Pete Jones (Presenter)
Andy Mealor (Presenter)
James Redmore (Secretariat)
Gareth Powell (Secretariat)

Actions

Agenda item¹	Action
[3,3]	A43 – Panel requested that a case study be completed on an area that has changed a lot between the old and new geographies e.g. Stratford.
[3,4]	A44 – Panel members to speak to SDR about the use of K means
[3,5]	A45 – Panel requested a paper on contingency planning for disasters.
[3,5]	A46 – What effect the 2019 rehearsal results have on the field operations simulation model are to be brought back to the panel at a later date.

¹ [M,N] – M denotes the panel number, N denotes its position on the agenda.

[3,6]	A47 – To return to panel with updated paper including software QA, how to involve Local Authorities in the QA process, and scoping where the start and end of the QA process is & what progress has been made.
[3,10]	A48 – How does the VOA and address list relationship inform decisions e.g if address on one and not on the other how can we use this knowledge?
[3,10]	A49 – Update address list – to see how the VOA list compares with the census list.
[3,10]	A50 – Change methodology from Cramer's V

Minutes

1. Actions from the previous meeting

1.1 – Update on action A19 on why both types of perturbation are necessary – to be sent to panel members after meeting. [Update: this has been sent to the panel members]

1.2 – Update on action A20 on homelessness from the previous panel “How homeless people will be counted and if there should be a separate survey to count them.” Panel request this action continues with Jon Wroth-Smith as action owner.

1.3 – Action A21, the Panel requested to see eQuestionnaires at spring panel. This will be presented at today's panel. Within this discussion, the panel no longer required a dedicated behavioural insights session, closing actions A21 and A39.

1.4 – Actions A22 and A23 on the Hard-to-Count index and Census Coverage Survey will be presented at a future panel.

1.5 – Continue with action for optimization method A24 – with a dedicated future session.

1.6 – The panel recommends to the ONS that they organise a scoping exercise with the ATI for use of machine learning for matching. Matching is to be the focus area, but other areas are to be considered for use of this method in the future.

1.7 – Multiple Matching actions were closed. Meetings with Data Science Campus and the Alan Turing institute completed A25. A26 was closed by the development of the matching strategy, with it noted the possible requirement for online versus paper responses require assessment in future. A27 was closed as work on the household and individual matching has been progressed with a session on this now scheduled for a future meeting. Both probabilistic and deterministic matching has been progressed, with RMR work being considered in this, closing A28. Updates to the clerical matching systems have been scoped, with improved software programs included closing A29.

1.8 – Close remove false person's actions. Action A30 has been closed since the distributions of removed date of births are now outputted as a key diagnostic for analysis. The proposal in A31 to use probability has been considered and is being developed as an alternative method for the 2021 Census, to be used if the current method proves ineffective. A32 has also been closed since as with the 2011 Census, it was confirmed a detailed report will be issued detailing response rate and defining what a response consists of.

1.9 – A33 has also been closed as the response chasing algorithm, covered at item 4 in this meeting, describes Matching quality versus DSE quality.

1.10 – Figure 1 from E&I was amended as requested, closing A34. Update on action A35 to provide data on the number of reused donors is in progress.

1.11 – Item 7 contained information of the intended ADC design and use of IPS, closing actions A36 and A37.

1.12 – Panel members reviewed and fed back on the paper in A38, closing this action. Actions A40, A41 and A42 will be the subject of presentations for future panels.

1.13 – Request from the panel to formalise the action log for presentation – presenting only the necessary actions that need to come back to the panel.

2. eQuestionnaires – Steve Woodland & Charlotte Hirst

Background

Action from the March 2019 panel. It was requested that the eQuestionnaire was brought to the June 2019 panel, the panel wished to see the layout and style of questions. The 2021 Census questionnaire will be primarily online, with the development team designing a questionnaire that would work across both modes. The presenters took the panel through the questionnaire design process and demonstrated answering various questions within the questionnaire to demonstrate validity.

Discussion & Suggestions

2.1 – There appeared to be no option for a half-sibling.

2.2 – Do access codes create a barrier to completing the census? Is there a method to make this easier to access? Such as a 16-digit number split in 4.

2.3 – The panel raised concerns about filling in the section on having another address – this question can be interpreted in a number of ways. Could be unclear for people who travel for work and stay in hotels that they are not required to fill in this section.

2.4 – There is a need to make it clear that the input for dates is numeric, rather than characters.

2.5 – Census date – do we need this on the 'who lives here' question?

2.6 – EQ and routing – we need to ensure that this can be looked back on in the future when the questionnaire is no longer in use.

2.7 – Panel queried about autofill and if it will autocomplete census forms, and whether or not this should be turned off.

No actions.

3. Maximising Response – Orlaith Fraser & Victor Meirinhos

Background

This paper outlines the maximising response strategy for the 2021 Census, the goal of which is to not only maximise response but to minimise variability in subpopulations and areas. This methodology includes ensuring those who are unable or unwilling to respond online can fill in the census, targeting specific hard-to-count groups, predicting patterns of non-response and efficient allocation of resources for follow-up visits.

The panel commented that this work is excellent and thanked ONS for doing this.

Discussion & Suggestions

3.1 – Panel asked how changes to geographies will affect outputs and the model.

3.2 – Look at international practices with reminder letters – ONS have already looked at some different practices.

Actions

A43 – Panel requested that a case study be completed on an area that has changed a lot between the old and new geographies e.g. Stratford.

4. Response Chasing Algorithm, Field Prioritisation Algorithm & Response Profiles – Orlaith Fraser & Victor Meirinhos

Background

These papers summarise the methodology for the Response Chasing Algorithm (RCA), Field Prioritisation Algorithm (FPA) and Response Profiles. The RCA enables live return rates to be compared with expected response patterns, identifying shortfalls in return rates and recommends interventions to address problems. The FPA is designed to reduce variability within local authorities by targeting low response output areas. This is completed by optimising the route for field visits. Response profiles model volume of responses expected during census operations, modelling where, when and how people are going to respond.

Discussion & Suggestions

4.1 – Implementation of k-means in MATLAB rather than SPSS.

4.2 – Important to have a measure of variability on the dashboard.

4.3 – Necessary to look at demographics to determine what the best predictors of response are.

4.4 – Panel questioned whether the regression models were necessary and what they add to the process.

Actions

A44 – Panel members will follow up with SDR on the use of K-means.

5. Field Operations Simulation – Kim Ward

Background

The Field Operations Simulation (FOS) model is designed to simulate the 2021 census collection operation, including all aspects of the collection operation. The FOS produces outputs that support resourcing decisions and its primary function is to estimate the volume of field resources required for the 2021 Census.

Discussion & Suggestions

5.1 – Think about what experiments could be done as part of Census.

5.2 – Don't rely on the model in the live environment there must be appropriate checks in place.

5.3 – Need a view on what to do if response is only slightly below target.

5.4 – Panel suggested using an ARIMA model for the forecast, or alternative appropriate time series models. Test if the model is following a daily cut of the 2019 Census rehearsal, this will determine its accuracy.

5.5 – Is there a plan B in the event the model is an inaccurate predictor of reality?

Actions

A45 – Panel requested a paper on disasters and what the effect on the census would be, what should happen as a result and what are the consequences to the Field Operations Simulation.

A46 – What effect the 2019 rehearsal results have on the model are to be brought back to the panel at a later date.

6. Quality Assurance Strategy – Pete Large

Background

This paper outlines the Quality Assurance (QA) Strategy for the 2021 census. The quality of census outputs is determined by user's confidence in the results and that they are fit for purpose; ensuring there is no need for adjustments following the first release; and leaving a legacy of methods, tools and skills for Quality Assurance for post-2021.

Discussion & Suggestions

6.1 – Ensure QA of the software, specifically the way in which the code is written.

6.2 – Write any documents to ensure they are understood – try to avoid use of statistical words e.g. reliability and validity

6.3 – Needs clearer scope of where QA stops and starts – this is to be updated in the paper.

6.4 – Need to be careful about QA methods doing the same thing in different ways e.g. mortality rates & fertility rates.

6.5 – In order to keep local authorities involved, send them a template to fill in.

Actions

A47 – To return to panel with updated paper. Including software QA, LA options and how to involve them, scoping of start and end of QA & progress made.

7. Statistical Design for the Population Statistics Transformation (Integrated Statistical Design) – Andy Teague

Background

Integrated statistical design is supporting the transformation of population, migration and wider statistics within the ONS. The purpose of the transformation is to modernise, improve

efficiency, reduce cost, reduce response burden, and ensure the ONS remains current and relevant at the forefront of statistics.

Discussion & Suggestions

7.1 – Discussion around how accurate outputs from admin data are.

7.2 – The challenge of obtaining household information from admin data.

7.3 – Some issues around the ONS I.D. number and privacy issues were raised.

7.4 – What are the international practices?

No actions

8. Update on data linkage Advisory Panel

8.1 - Bernard and Sarah provided a brief update on the Data Linkage Advisory Panel, what had been discussed at that meeting and how it is linked to the census.

9. Integrated Population and Characteristics Survey (IPACS) – Pete Jones

Background

As a part of the transformation of population, migration and social statistics the ONS seeks to integrate more administrative and non-survey data sources into official statistics. This will reduce the reliance on large population and business surveys, though not eliminate them completely. This paper outlines the ONS's methodology for this process.

Discussion & Suggestions

9.1 – Suggestion to compare with NISRA's methodology on their use of GP visit data.

9.2 – Discussion about whether this survey could be mandatory? How do we persuade people to fill in the survey?

9.3 – What level of response rates are required to make this beneficial?

9.4 – Can the different methodologies be integrated? Using V3 and V2 where each is more appropriate.

9.5 – What's the impact of the survey design for PCS in further waves of LMS?

No actions

10. Enhanced Outputs – VOA Data – Andy Mealor

Background

This paper details the findings of feasibility research into whether administrative data variables from the Valuations Office Agency (VOA) are suitable to undergo edit and imputation (E&I) within the standard census framework when linked to 2011 Census questionnaire data. The

research in this paper focused on the imputation of the number of rooms variable from the VOA data.

Discussion & Suggestions

10.1 – The panel suggested looking at why VOA data has high missingness within central London. Working with the VOA to understand this.

10.2 – Find an alternative to using Cramer's V, perhaps utilising a correlation instead for ordered variables.

10.3 – Look at using different methods e.g. predictive means methods. Panel concerned about selectivity and the issue of the unlinked data.

10.4 – The way that this is communicated is very important – why are we collecting, what is it used for, what's the continuity plan? Be clear that this has been produced in a very different way before.

Action

A48 – How does the VOA and address list relationship inform decisions e.g if address on one and not on the other how can we use this knowledge?

A49 – Update address list – to see how the VOA list compares with the census list.

A50 – Change methodology from Cramer's V.