

# National Statistician's Data Ethics Advisory Committee Minute

Tuesday, 1 February 2022  
Meeting held via video conference

## **Present**

### **Members**

Dame Moira Gibb (Chair)  
Stephen Balchin  
Rob Bumpstead  
Vanessa Cuthill  
Monica Magadi  
Isabel Nisbet  
Marion Oswald

### **UK Statistics Authority**

Lily O'Flynn  
Daniel Towler  
Simon Whitworth

### **Office for National Statistics**

The National Statistician, Sir Ian Diamond (for item 1)  
Michael Cole (for item 3)  
Alison Reynolds (for item 3)  
Zoe Sargent (for item 3)  
Tansy Arthur (for item 4)  
Helen Colvin (for item 4)  
Josephine Foubert (for item 4)

### **Other**

Marc Verlot, Cabinet Office (for item 4)

### **Apologies**

Colin Godbold  
Emma Uprichard

## **1. Talk and Q&A with the National Statistician**

- 1.1 The Chair welcomed members and the National Statistician to the 27<sup>th</sup> meeting of the National Statistician's Data Ethics Advisory Committee (NSDEC).
- 1.2 The National Statistician firstly thanked the Committee for the continued advice and assurance that it provides to the statistical system in a timely and efficient manner. Sir Ian noted the importance of ethics in ensuring there is public trust in the way in which the Office for National Statistics collects, processes and analyses data.
- 1.3 The National Statistician also explained to the Committee his vision for the ONS over the next year. This placed importance on producing faster economic indicators, improving understanding of the changing labour force, improving the inclusivity of data holdings, as well as focus on the dissemination of Census 2021 results and continued work to inform the future of the Census.

- 1.4 During discussion, the members of the Committee raised the following questions:
- i. The Committee questioned how the sensitivity of different datasets can be considered when processing and linking them to other information, and how it should be ensured that appropriate ethical safeguards are in place for this. The National Statistician stated that no data should be out of scope due to the sheer potential and public benefit of data linkage. However, the NSDEC was assured that while this may be the case, all proposals to use the data would have to be in the public good and ethically sound.
  - ii. The Committee also raised the importance of maintaining the confidentiality of data, as well as privacy within analysis, which extends to how the products of analysis are deployed. The National Statistician assured the Committee that the most appropriate technical security controls are in place to enable this, and that the involvement of ethics and research accreditation are essential safeguards.
  - iii. The National Statistician also stressed the importance of transparency and the publication of outputs to ensure that the public good from doing the research is maximised.
  - iv. The NSDEC also emphasised to the National Statistician how useful the data ethics compliance reviews are in assuring the Committee, and also the public, that the advice the Committee give, is actioned by the research teams.

## **2. Minutes and matters arising from the previous meeting**

- 2.1 Members approved the minutes from the previous meeting and correspondence.
- 2.2 Daniel Towler updated the Committee with progress on actions from the October 2021 meeting, and correspondence in the interim period. All actions were completed or in progress.

## **3. Resolving conflicting ethnicities within administrative data when producing admin-based ethnicity statistics. NSDEC(22)01.**

- 3.1 Alison Reynolds, Zoe Sargent and Michael Cole from the Social Statistics Transformation Analysis and Research (SSTAR) division, ONS, presented this item. This item considered a proposed approach to resolving conflicting ethnicities within administrative data, to feed into work which aims to provide insights on population, migration and society using a range of data sources. This derived from an identified issue within linked administrative data where several ethnicities may be recorded against a single individual, which prevents each data record from having a single accurate recorded ethnicity. The admin-based ethnicity dataset produced through this research will be used to produce aggregate-level statistics only, and not to identify the ethnicities of specific individuals. This work will form part of the evidence base for the 2023 National Statistician's Recommendation to the government on the future of population and social statistics, which will outline the future of Census arrangements. This project is seeking advice from NSDEC on the feasibility research phase and if it were to be implemented as a standard method for producing official statistics in future, further ethical consideration would be sought.
- 3.2 The proposed approach, labelled 're-allocation of 'other'', seeks to build upon a previously explored approach which assigns data records with the most recently recorded ethnicity from administrative data records. This most recent approach, however, leads to an inflated proportion of data records being assigned an ethnicity of 'other'. Therefore, under the newly proposed approach, if the most recent ethnicity in an administrative data record was recorded as 'other', this data

record would be allocated the next available ethnicity within administrative records subject to certain conditions, such as to the next available ethnicity not being refused. The 're-allocation of 'other'' approach is applied to fewer than 1% of records but is important for maintaining statistical accuracy at the aggregate-level.

- 3.3 In discussion of the paper, the Committee acknowledged the need for the research, and the public good, but raised the following points:
- i. It was noted that approximately 10% of the population would not be present in the administrative datasets used in the method at present, the Committee stressed the importance of ensuring that further work is conducted to address missingness within the data.
  - ii. The NSDEC noted the apparent risk of comparing these estimates to the 2011 Census, as a baseline for accurate ethnicity aggregates, due to this dataset being over 11 years old, which increases the possibility that aggregate ethnicity statistics have changed since then. The Committee was assured to see plans of using the 2021 Census, when available, to continue to confirm the accuracy of the approach.
  - iii. The Committee urged caution on what this may be used for. While it may be useful for population estimates at higher geographic levels, it could be less accurate at a more granular level, meaning certain populations may be misrepresented.
  - iv. It was stated that the administrative data sources used for this work should be recorded by self-identification, meaning that individuals would provide their own ethnicity. The Committee asked for evidence regarding the extent to which this happens in practice in the administrative data used.
  - v. The Committee emphasised the importance of communication to improve public understanding of how ONS is using administrative data to produce ethnicity estimates so that the public could have confidence in the way in which this is done.
  - vi. Given the uncertainty as to whether all the data is actually self-identified in practice, the Committee raised the ethical risk of re-allocating someone's purposeful, self-identified ethnicity of other, with an ethnicity that may have been selected by a third party. It was acknowledged that there is also a risk that by retaining the ethnicity of other, an individual's ethnicity entry may be retained as a value reported by a third party rather than changing the record to their self-identified ethnicity.
  - vii. Furthermore, while it was noted that ethnicity may be re-allocated to a previously recorded ethnicity, even if this was self-identified, this may have changed to other over time due to the fluidity of one's ethnicity. The Committee questioned the use of a cross-sectional baseline as a comparator, as this would not account for the fluidity of ethnicity.
  - viii. The Committee stressed the importance of considering individuals' ethical consent of re-allocating an ethnicity of other, as it may be the case that an individual wants to be recorded as other due to a lack of suitable alternatives.
- 3.4 The NSDEC appreciated the public good and need for this work to support recommendations around the future of Census but concluded that further justification and assurance on points vi, vii, and viii are required before final advice could be provided on this work.
- 3.5 **Action – ONS to provide assurance on all points raised in section 3.3, with particular focus on the final three points, and present this to NSDEC so that final advice can be given on this approach.**

**4. Disability Experiences Survey- capturing data on the lived experiences of disabled people. NSDEC(22)02.**

- 4.1 Tansy Arthur from the Survey Division (ONS), Josephine Foubert and Helen Colvin as Disability Statistics Leads (ONS) and Marc Verlot (Cabinet Office) were present for this item.
- 4.2 The Disability Experiences Survey has been sponsored by the Cabinet Office to feed into the national disability strategy and also provide HM Government, and in particular Cabinet Office, with new vital insight in the life experiences of people with disabilities to strengthen the policymaking process and help in the evaluation of existing initiatives aimed at improving provision for people with disabilities. This project is a new household survey to monitor the experiences of people with disabilities. The survey will seek to help ensure that the lived experiences of people with disabilities are robustly captured across key areas of life and allow ONS to deliver more granular data on the topic.
- 4.3 The development of this survey is split into four phases, these include: the question design which will include cognitive testing, a pilot study, the main study, and data analysis.
- 4.4 The following points were raised in discussion:
- i. The Committee queried what the measure of success would be to determine whether this study progresses from the pilot phase, into an annual and potentially longitudinal study.
  - ii. The Committee stressed the importance that the data collected from this study has an outcome beyond better statistics. It was noted that the Disabled community are sometimes asked to provide information of lived experiences which they consider is not then used to effect improvement.
  - iii. It was queried as to whether the suggested sample size would be large enough to pick up very specific disabilities, which may not fit directly into the nine impairments categories being used to inform the representativeness of this study from the Government Statistical Service (GSS) harmonised impairment question.
  - iv. It was suggested that the researchers look at the Inclusive Data Taskforce's work (IDTF), which included disability, to ensure their work benefits from the IDTF's recommendations.
  - v. The Committee appreciated that respondents would be able to access the survey in Welsh but noted the need for further confidentiality procedures for the translation of participants answers from Welsh to English for analysis.
  - vi. Due to this being a household survey, the Committee questioned whether the burden would be disproportionate on large households, which could subsequently lead to misrepresentation.
  - vii. The Committee stressed the importance of proportionate adjustments to survey materials for participants with different cognitive abilities, as the use of standard materials may restrict the accessibility of the survey participant information, which could impact on individuals' decisions to take part.
  - viii. It was noted that proxy responses may be required should participants not be able to answer themselves. The NSDEC suggested that the researchers should learn from existing best practice and policies across ONS where proxies have previously been used in survey collection.
  - ix. The Committee requested clarity on the Cabinet Office's involvement in the study as well as an understanding of and justification for any data Cabinet Office will receive following the study. It was also stressed that identifiable data should not be shared with the Cabinet Office.

- x. The NSDEC confirmed that any ethical approval from this submission relates only to the survey collection and initial analysis by the ONS, and therefore further uses of the data by the Cabinet Office or other parties would not be in scope of any ethical approval from NSDEC without further details.

**4.5 Action – the ONS to address points raised by the NSDEC in a revised application and provide this to the Secretariat**

**5. Updates to the Ethics Self-Assessment tool. NSDEC(22)03.**

- 5.1 Daniel Towler, from the UK Statistics Authority, updated the Committee on work being done to update the ethics self-assessment tool. The self-assessment tool was used 258 times in 2021, and has a broad audience across ONS, the GSS and the wider academic community, to enable researchers to review the ethics of their projects throughout the research cycle, using an easy-to-use framework.
- 5.2 The presentation outlined the changes made to the self-assessment tool, and how the iterations had been grounded in the views of users. The Committee was informed that further engagement activities are planned to gain agreement on the changes.
- 5.3 The NSDEC appreciated the update and was supportive of the approach the UK Statistics Authority Data Ethics team had taken to ensure that the user community had the ability to feed into the re-development of the ethics self-assessment tool.

**6. Any other business**

- 6.1 The quarterly Data Ethics Compliance Review, NSDEC22(04) was presented to the Committee. This presented a data ethics compliance review of the Mental Health of Children and Young People in England 2020 (MHCYP 2020). The data ethics compliance review has confirmed that the research team have taken the advice of the committee and implemented all actions agreed via correspondence in June 2020.
- 6.2 The UK Statistics Authority's Centre for Applied Data Ethics (CADE) also presented a recent draft guidance piece on inclusivity, NSDEC(22)05. This guidance has been developed in response to the UKSA's Inclusive Data Taskforce report and the National Statistician's response and recommendations.
- 6.3 The Committee highlighted that it is important to be clear that inclusivity is not just about ensuring all voices are heard. It is also about the way in which those voices are brought into the conversation and that research and statistics has to be inclusive in allowing participation. The NSDEC noted how comprehensive the guidance was and suggested that additional consideration is given to ensuring inclusivity of participation.
- 6.4 The next meeting will be held on 27 April 2022.