

National Statistician's Data Ethics Advisory Committee Minute

Wednesday, 27th April 2022
Boardroom, Drummond Gate, London and via video conference

Present

Members

Dame Moira Gibb (Chair)
Colin Godbold
Monica Magadi
Marion Oswald

UK Statistics Authority

Lily O'Flynn
Tia-Mae Surtees
Daniel Towler
Rhys Nadin (for item 5)

Office for National Statistics

Nicky Rogers (for item 2)
Jonny Tinsley (for item 3)
Dani Evans (for item 4)
Sarah Beasley (for item 4)

Other

Romi Gupta, Department for Health and Social Care (DHSC) (for item 3)

Apologies

Stephen Balchin
Rob Bumpstead
Vanessa Cuthill
Isabel Nisbet
Emma Uprichard
Simon Whitworth

1. Minutes and matters arising from the previous meeting

- 1.1 Members of the Committee approved the minutes from the last meeting in the interim period via correspondence. The minute for the project that was received via correspondence in March, were approved by the Committee.
- 1.2 Daniel Towler updated the Committee with progress on actions from the February 2022 meeting, and correspondence in the interim period. All actions were completed or in progress.
- 1.3 Members were notified that those members that had sent apologies for the meeting, had sent preliminary comments on the projects. The Secretariat confirmed it would raise these comments in discussions at relevant points throughout the meeting.

2. 2021 Census Cohort Study: Proof of Concept. NSDEC(22)06.

2.1 Outcome – The Committee advised that the research team collaborate with the Centre for Applied Data Ethics to consider the ethical risks identified before returning to the Committee with a proof-of-concept application.

2.2 Detail of the project and discussion:

2.3 Nicky Rogers from the Health Population and Methods Transformation division, ONS, presented this item. This was a presentation to overview the transformed population and social statistics system. At the heart of this is a system of demographic accounts which seeks to estimate the key components of population and population change. The project seeks to build a 2021 Census-based Cohort Study to provide a prospective longitudinal resource.

2.4 Given the significance and long-term nature of the project, this presentation introduced the Committee to this work. The research team assured the Committee that the NSDEC would be consulted as the project develops. This includes presenting a project application on the first phase of the work, the proof-of-concept study, in July.

2.5 The NSDEC appreciated the early sight of the project, but raised the following points:

- i. The Committee sought clarity on whether this project seeks to use the whole 2021 Census population as its cohort and would be a longitudinal study. It was confirmed that this is the aim for the project.
- ii. The NSDEC stressed that it should be absolutely clear in any communications on this project that it is a whole population study, and that the use of the word cohort may suggest that this work only considers a proportion of the population which would therefore be misleading.
- iii. It was also stressed that significant public acceptability work would be vital in the development of this study. Linked to this, the NSDEC advised that the ONS consider issues relating to whether data subjects would be able to opt-out of their data being included in this study.
- iv. The ONS is advised to carefully consider and document the legal basis for this project, to ensure that there is a strong legal grounding for all data acquisition, use and processing related to this project, and ensure compliance with Data Protection Legislation.
- v. The Committee wanted clarity as to whether all of a data subject's administrative data would be linked to their records as part of this study. Justification of why the data used in this project is required to support this work is necessary.
- vi. The Committee requested evidence that this work will be successful and produce a high-quality data asset that would realise significant public benefits. The Committee further stressed that any limitations should be clearly communicated so that the onwards use of this data for accredited research projects and official statistics isn't assumed to be of the same or similar quality to the Census.
- vii. The NSDEC actioned the research team to ensure the Centre for Applied Data Ethics (CADE) is engaged in all phases of this work as it develops so that the project considers ethics by design and approaches NSDEC for independent advice at relevant points throughout the project design.

2.6 Action – The research team to ensure that CADE is engaged in all phases of the work and approach NSDEC at relevant points throughout the study.

- 2.7 Action – ONS to present a high-level overview of the work that is feeding into the 2023 recommendation to understand how all of the work that ONS is undertaking in this space links together.**
- 3. Shared Outcomes Fund Project: Understanding the Long-Term Health and Employment Outcomes Due to COVID-19 on Disproportionately Impacted Groups. NSDEC(22)07.**
- 3.1 Outcome – The Committee supported this work subject to assurances to the points outlined in 3.5 being provided to the Secretariat.**
- 3.2 Detail of the project and discussion:**
- 3.3 Romi Gupta from the Department for Health and Social Care (DHSC) and Jonny Tinsley from the Health Analysis and Pandemic Insight directorate, ONS, presented this item. This project aims to link in data from DWP and HMRC on benefits and earnings to the already linked data that includes Census, death registrations and health data. This will enable analysis of labour market outcomes (employment, earnings) against health outcomes (COVID-19, chronic health conditions and reduced access to care), by personal characteristics such as disability, ethnicity, sex and low-level geography. The aim is to understand population characteristics and geographic areas that have been disproportionately affected by COVID-19, either directly through infection, or indirectly through economic effects or reduced access to healthcare.
- 3.4 The Committee had previously been made aware of this project in an addendum of the project “to determine the population-level relative risk of hospitalisation or death that COVID-19 presents to people with different socio-demographic characteristics and co-morbidities.” This set out that feasibility testing to see if DWP data could be successfully linked to Census, to subsequently add into the COVID-19 health data asset. The NSDEC requested that this project is submitted to the Committee for review, should the feasibility work prove successful.
- 3.5 The Committee acknowledged the public good that could be achieved from this work. The following points were raised in discussion:
- i. The Committee requested that the research team provide additional assurance that the public engagement work undertaken to support this project is sufficient. The NSDEC recommends that the research team further considers public attitudes towards the linkage to benefits and tax data to Census more generally. Specifically, where this data may be used to understand impacts of Covid on labour market outcomes for minority populations or groups at greater risk of disadvantage, NSDEC recommends that further public acceptability work is undertaken to ensure that such groups are comfortable with the use of such data and understand how analysis may feed into policy interventions that impact them.
 - ii. The research team clarified that the population coverage of the study is limited to working age because of its focus on employment, and that ONS cover non-working age populations, such as the elderly, in other health work.
 - iii. The Committee stressed that the results of the work actually do lead to the public benefit being realised, and therefore emphasised the importance of transparency to ensure that research project outcomes that utilise this dataset are easily accessible and can have impact.
 - iv. The NSDEC recommended that the research team review the impact that uses of this data asset are having, especially in relation to changing policy priorities as the UK recovers from the Covid-19 pandemic.

- v. Due to ONS being responsible for record level analysis on this data, the Committee sought confirmation that other government departments would only have access to statistically disclosure controlled aggregate statistics produced as a result of this linkage work.
 - vi. Furthermore, the NSDEC stressed that uses of this linked dataset, both now and in the future, are for statistical and research purposes only.
- 3.6 It was agreed that the NSDEC should be consulted as major amendments to this work develop in the future, while less significant updates to the project can be supported internally from an ethics perspective by the UK Statistics Authority's Centre for Applied Data Ethics.
- 3.7 The Committee recognised a recent increase in the number of significant data linkage projects submitted for review, similar to this project, and therefore agreed that the NSDEC should engage in a wider piece of work, with the support of the Centre for Applied Data Ethics, to identify key ethical considerations relating to such data linkage projects, particularly those that involve larger, more sensitive data.
- 3.8 Action – The research team to provide assurances to the secretariat.**
- 3.9 Action – The NSDEC to have a wider conversation on the key ethical considerations relating to the linkage of data, as outlined in point 3.7.**
- 4. An investigation into the most appropriate weighting approach for “Prefer not to say” and “Don’t Know” answers to the Sex question on the National Survey for Wales (NSW). NSDEC(22)08.**
- 4.1 Outcome –The research team provided the NSDEC with further detail and clarity on the suitability of the weighting methodology, via correspondence, as per the Committee’s request. Based on this, the NSDEC supported the recommendation subject to the following advice:**
- i. **This should be a short-term solution, and should not be used for longer than it needs to be, and;**
 - ii. **The outputs are fully, and clearly, transparent about how the weighting has been applied.**
- 4.2 Detail of the project and discussion:**
- 4.3 Dani Evans and Sarah Beasley from the surveys team in ONS presented this item which sought ethical consideration for a weighting methodology being used to deal with an issue in the data following data collection for the National Survey for Wales (NSW).
- 4.4 The NSW has a question asking the respondent’s sex with the response options of “Male”, “Female” and spontaneous options of “Prefer not to select male or female” (“PNTS”), and “Don’t know”. The collection of the data in this manner has caused an issue with analysis as traditionally, data is weighted using Age, Sex (male and female) and local authority data. However, this method does not account for cases where sex is not “male” or “female”. The proposed solution to weighting “PNTS” or “Don’t Know” Sex answers on NSW, finds an average weight of a male and female in each age/region group to then apply those to the non-male/non-female responses to ensure that these responses are fairly included in all analysis without a sex breakdown.
- 4.5 During the meeting, the Committee acknowledged the situation, and raised the following points within the discussion:

- i. The NSDEC wanted clarity on how short-term the use of the survey question in this format would be, as the current formulation of the survey question requires the use of this non-standard method. The Committee advised that it should not be used for longer than it needs to be due to the recognised limitations of the survey question.
- ii. Furthermore, the NSDEC urged the team to ensure that the methodology is fully transparent in all published outputs, while ensuring appropriate caveats and detail on the methods limitations enable the public to be fully aware of how the data has been processed, why this method is preferable to others, and how this composition of the survey question should be avoided.
- iii. The NSDEC requested further detail on potential consequences of the suggested weighting methodology; including further detail on the questions in the survey to elucidate what this survey may inform.

4.6 In response to the comments made by the Committee, the research team provided the NSDEC with a paper via correspondence shortly after the meeting. The paper outlined the following:

- i. The Committee was assured that this suggested weighting methodology was better than the other proposed solutions. In general, this was due to improved inclusivity for the respondents who answered “PNTS” or “Don’t Know” as by weighting them in the suggested way would mean they could be upscaled to the population.
- ii. The research team acknowledged that the potential consequence of the weighting would be that those who answer as “PNTS” or “Don’t Know” would be excluded from any sex breakdowns, meaning any outputs informed by these breakdowns would exclude these respondents. However, it was stated that this is a consequence of the nature of the questions asked, rather than the methodology, and if this methodology was not to be used, the “Don’t Know” and “PNTS” respondents would be underrepresented in all breakdowns.
- iii. The research team confirmed that technical report will be published alongside each published annual dataset. This report provides detail on the weighting methodology as standard, and the adapted methodology will be clearly presented including why this method is preferable to others. This includes the appropriate caveat that “PNTS” or “Don’t Know” responses are not included in sex breakdowns.

4.7 The Committee appreciated the further assurances and supported the recommendation on the basis that this solution is not utilised for any longer than it needs to be and that outputs are fully transparent about how the weighting has been applied.

5. Data ethics compliance review plan 2022/23. NSDEC(22)09.

5.1 Outcome – The Committee supported the forward plan for the next year and suggested that further work to assess compliance after a project has concluded would be useful.

5.2 Detail of the project and discussion:

5.3 Rhys Nadin, from the data protection team, UKSA, presented this item to the NSDEC. This paper outlined the plan for data ethics compliance reviews for 2022/23, which includes projects external to ONS, which was previously

requested by the Committee. The paper includes the selection criteria for projects to be chosen for review, which focuses on those projects with higher sensitivity. Furthermore, future plans to increase the level of assurance provided to the Committee (subject to resourcing). This includes more reviews of projects, plus compliance checks of ethics self-assessments.

- 5.4 During discussion, the following points were raised:
- i. The Committee queried what would happen should the compliance review conclude that a project, external to ONS, was non-compliant with the advice of the Committee. Assurance was provided that the NSDEC could recommend to the UK Statistics Authority and the Research Accreditation Panel that the project's accreditation be removed where non-compliance is identified as undermining the project's ethical standards.
 - ii. The NSDEC suggested that reviews of projects that have concluded would be useful to ascertain whether the NSDEC's advice on transparency, outputs and presentation, interpretation, and use of results were implemented. The NSDEC added that it would also be beneficial to see further consideration of the extent to which projects have achieved their originally stated public good, or whether they are on track to achieve this.
- 5.5 The Committee reaffirmed the value that the data ethics compliance reviews provide in assuring the NSDEC of the impact of their consideration and appreciated the update on those projects that will be reviewed in the coming year.

6. Any Other Business

- 6.1 The NSDEC suggested that it would be useful to hold a workshop for members to discuss strategic questions that the Committee has identified as particularly pertinent to the current research, statistics and data ethics landscape. This should include:
- i. a discussion of the UK Statistics Authority's approach to the consideration of data ethics, and whether additional ethical safeguards should be embedded within data linkage and acquisition processes,
 - ii. the extent to which the public good of a research project is achieved in the way set out at research design phase, and how NSDEC might gain more assurance that the public good of research projects that it considers are realised, and
 - iii. detail on the ongoing work and future plans for the 2023 recommendation, which seeks to inform the future of the Census and population statistics. This is so the Committee can better understand how strands of work, relating to the recommendation, fit together, broader public engagement around the recommendation, and how ethics will be considered throughout.
- 6.2 The Committee agreed that the hybrid meeting worked well, and supported its use in future to allow members to join in person, where possible.
- 6.3 The next meeting will be held on 5 July 2022.