

## International Steering Group – Note of Key Points

22 September 2022 : 9am – 11 am (BST)

### Attending

Professor James Brown – UTS

Abby Morgan – Stats New Zealand

Owen Abbott – ONS

Professor David Martin – University of Southampton

Andrew Morris - Health Data Research UK

Esta Clark – NRS

Caroline Ellis – NRS

Bethanie Fenney – NRS

Tom Macintyre – NRS

Ralph McDevitt – NRS

Uche Okoro – NRS

### Key Actions

1. The group discussed whether a representative from OSR could attend a future meeting or be invited to join as an observer. Consensus was that this would be wholly appropriate and would provide good transparency of process.
2. Minute from 1 September agreed and ready to be published.
3. NRS provided an update on progress with Admin Data:
  - Health-PBPP submitted on 09/09 with tier 2 scrutiny on 28 Sept (Interview panel). The panel will circulate questions from tiers 1 & 2 in advance, providing an opportunity for Chair of ISG (and others) to provide input as necessary. Best outcome would be conditional acceptance to proceed with separate DSAs; final outcome would follow a week later. Chair has offered availability to attend the session scheduled for 28 September.
  - NRS to secure the majority of data by December and have it ready for linking in the clear. Processing will entail minor editing of existing code that was prepared for de-identified data. ONS experiences will be useful as this project is progressed.
  - This administrative project is under-pinned by statute and is not a research project. NRS and NHS have already exchanged communications to such effect.

- Stats-PBPP for School Pupil Census is on track for submission.
  - Acquisition of datasets is on track. NRS are still pursuing DWP for individual and aggregate data. NRS submitted an application to HMRC on 21/09 to get permission for their aggregate data to be shared. Recommendation from ISG is to get clarity on which geographies can be secured whilst adhering to non-disclosure rules.
  - Local Authorities council tax exemption data to be pursued by NRS, but this will probably only be from some of the 32 LAs. NRS will push to secure this data from lower-responding areas. This data will also be useful in identifying unoccupied properties (second homes).
4. NRS provided an update on CE estimations.
- Small CEs – to increase sample size, NRS only collected minimal information in CCS with just name, DOB and sex along with communal establishment type.
  - HESA – where NRS only have postcode with missing address, then those records can be linked through other administrative data. If NRS plans to go down a modelling route (rather than simple DSE) then household information will be required. For large communal establishments, like student halls, this will be lower priority. NRS was commended for getting high response rates from these CEs.
  - Prisons data can be complicated due to prisoner status (e.g. remand), and ‘denial’ from any CCS home response where a prisoner may be included as usual resident at home address. Stats NZ handled prisoners in the following ways:
    - i. In census: census returns in combination with admin data was used to identify the portion of the prison population who had not responded to the census and to create census admin records for those individuals.
    - ii. In coverage adjustment: prisoners were excluded from coverage estimation as the census process effectively ensured they had near 100% coverage. Stats NZ’s coverage estimation assumed coverage patterns for people in non-private dwellings were the same as those in private so they would have over-adjusted the prison population if they had been left in during estimation.
  - NRS have secured MOD aggregate data for CE estimation – age band, sex by site (posting). It is ‘behind the wire’ on base population that is needed, excluding personnel living in the community. MOD data reports ‘Residents on the site’ so that should be accurate for CE estimates.
  - A critical issue for modelling will be the number of CEs included in CCS. If that is small, then modelling opportunities for NRS may be hampered due to lack of information to factor in effects.
  - For large CEs the admin data is likely to be higher quality than CCS linking and DSE approaches. Recommendation from ISG is to trust the admin data and to

use that. The error profile on Admin data will generally be smaller than the error profile created using statistical processes. NRS should check quality of admin data for different types of CEs, and use the admin data where quality is deemed good, noting the varying impact dependent on size of CEs. Be transparent over strategy for each type and size of CE.

5. AOB

- International Census Forum (ICF) scheduled for 26-30 September would be a good time to set out a timeframe for next steps on estimation and for ONS to know what input they would be able to contribute.

22 September 2022