

UNECE: Draft Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses

Steven Dunstan

Data Requirements and Quality Lead | UK Office for National Statistics, **Task Force Chair**

Sara Correia

Senior Research Officer | UK Office for National Statistics

Overview

- The **objectives** on which the guidelines were developed
- The **timetable** for approving the guidelines
- The **structure** and **content** of the guidelines
- Your **feedback / suggestions**
- Discussion: the **impact of COVID** on the quality of sources and the approaches used for assessment.

Aims

- To provide an overview of the draft guidelines for assessing the quality of administrative sources for use in censuses
- To invite feedback on any gaps, areas for development
- To establish whether any new developments / experiences in your countries might add value to the guidelines
 - Specifically, the impact COVID has had on the use of administrative data and the quality of these data.

TF members

Austria	Montenegro
Canada	Netherlands
Estonia	Poland
France	Portugal
Germany	Spain
Ireland	Turkey
Israel	United Kingdom (Chair)
Italy	United States of America

Eurostat

UNFPA

UNECE (Secretariat)

With extra contributions from:

New Zealand

Australia

Republic of Moldova

(Ian White – Expert)

The objectives

- Practical guidance relevant across the UNECE countries and across the different census methodologies.
- Building on the UNECE (2018) *Guidelines on the Use of Registers and Administrative Data for Population and Housing Statistics*.
- Drawing on:
 - Key literature / existing quality frameworks
 - Relevant projects e.g. CROS/ESSnet
 - Existing experiences across different countries

Timetable

Task	When
Presentation of draft to the UNECE Meeting of Census Experts	September 2020
Revise draft and submit to the CES Bureau for review	October to December 2020
Review of the report by the CES Bureau	February 2021
Consultation among the CES members	February to March 2021
Final Draft submitted to CES plenary session (June 2021)	April 2021

Structure of the guidelines (main chapters)

- Census methodologies and uses of administrative data
 - Quality framework (covering data and output quality dimensions)
 - Source Quality
 - Input Quality
 - Process Quality
 - Output Quality
- Stages of Assessment**
Covering:
- Key **Quality Dimensions**
 - Relevant **tools** and **indicators**
 - illustrative **case studies**

Key stages of assessment

Source quality

Quality assessment of data sources – based on metadata and communications with the supplier. Covering the data quality dimensions:

- Relevance (*representation and measurement*)
- Timeliness
- Coherence and Comparability (integration, change over time and domain), linkability
- Accessibility (Restrictions on data access and use; Public acceptability; Easy of data transfer and receipt; Interpretability – clear and comprehensive metadata)
- The Institutional Environment (confidence in supplier meeting the NSO's requirements)



Informs a decision on whether to proceed (or continue) with an acquisition and provides an understanding of risk and error (to be further assessed at the input stage)

Source quality: Case studies

1. **New Zealand:** assessment of administrative sources for use in the census for social and economic variables.
2. **New Zealand:** Privacy Impact Assessment (PIA) - against the accessibility data quality dimension.
3. **Estonia:** Improving administrative sources for the purpose of census (through legal and organisational measures).

Areas for feedback:

1. Have we covered the necessary data source quality dimensions?
2. Are there any areas of assessment you feel are missing?
3. We welcome your experiences and material you feel relevant.

Key stages of assessment

Input quality

Quality assessment of the raw data as supplied to the NSO. Through validation, data analysis and comparisons with other data sources. Covering the data quality dimensions:

- Readability and validation (on supply)
- Accuracy and reliability (representation and measurement error, re-supplied data comparisons)
- Timeliness and punctuality
- Linkability (quality of linkage variables)



Informing error corrections; transformations / statistical processing; and reporting.

Input Quality: Case studies

1. **Germany:** the quality of local population registers for the census, readability; accuracy; timeliness
2. **Poland:** The Polish Variable Quality System (statistical metrics for assessing accuracy)

Areas for feedback:

1. Have we covered the necessary data quality dimensions?
2. Are there any areas of assessment you feel are missing?
3. We welcome your experiences and material you feel relevant
 - We are interested in techniques, or thresholds countries use for deciding what are acceptable levels of errors (representation and measurement)

Key stages of assessment

Process quality

Processes to transform administrative sources for use in the census (accounting for errors and limitations identified at the source and input stages) and the quality of these processes.

- Record linkage
- Constructing statistical registers (“signs of life” methodology)
- Derived variables, through data integration (methods for choosing variables across administrative sources)
- Editing and imputation

Process Quality: Case studies

United Kingdom: measuring linkage quality when carrying out direct replacement of a census variable

Spain: Use of administrative data in construction of a census database: signs-of-life methodology

New Zealand: Process quality assessment when including administrative enumeration in the census

Areas for feedback:

1. Are we missing any processes you feel should be included?
2. We welcome your experiences and material you feel relevant
 - We are particularly interested in methods for assessing how good a process is, beyond what is covered in the guidelines to date.

Key stages of assessment

Output Quality

The quality of the census estimates produced through the integration of administrative data sources into the census design.

- Coherence and comparability (demographic analysis, comparisons with other sources)
- Accuracy (estimating coverage error)
- Timeliness
- Meeting user needs and balancing quality dimensions
 - Supplier feedback
 - Independent reviews
 - Sensitivity analysis
 - Quality reporting

Output Quality: Case study

Portugal: quality assessing the Statistical Population Dataset (SPD) and results

Areas for feedback:

1. Are we missing anything you feel should be included?
2. We have considered the need to further bring all stages together - e.g. with reference to total error frameworks

Your feedback / suggestions

1. Is there anything important you feel is missing from the guidelines?
2. Are there areas you think should be improved / expanded on in the guidelines?
3. Or areas (perhaps beyond the scope of the guidelines) that could warrant future development / the next set of guidelines?
4. The guidelines cover the indicators and metrics for assessing quality, but we're keen to include more on how NSOs decide what is "good enough".
 - Any experiences countries have on setting thresholds, rules of thumb, or the application of methods for deciding what is "good enough" would be valuable.
 - Perhaps countries have new experiences here in their response to the impact of COVID on their census / population statistics...

Administrative data quality: COVID impacts

1. How has the use of administrative data in the census changed through the COVID crisis?
2. What has been the impact of the crisis on the quality of the data used?
 - The impact on administrative data sources
 - The impact on the NSO collection (e.g. under a traditional census)
 - Decisions on use across these sources
3. Are there any examples of quality assessment done in this context, which we could incorporate into the guidelines?

Steven Dunstan – Task Force Chair; steven.dunstan@ons.gov.uk

Sara Correia – sara.correia@ons.gov.uk

Fiona Willis-Núñez – secretariat; fiona.willis-nunez@un.org