**Guidelines on the quality of multisource statistics**

Giovanna Brancato, Italian National Statistical Institute (Istat), brancato@istat.it

Gabriele Ascari, Italian National Statistical Institute (Istat), gabascari@istat.it

**Abstract**

*The Guidelines on the quality of multisource statistics are being produced in the framework of the ESS.VIP Admin project, as a work package of the Essnet Quality of multisource statistics.*

*They are intended to be practical and applicable in the National Statistical Institutes of the ESS, when several administrative sources are used exclusively or in combination with survey data to produce statistics. The guidelines are addressed to process managers who can use them in the design and the implementation of multisource statistical processes, to identify the factors having higher impact on output quality and to monitor and assess quality.*

*First of all a review of existing quality guidelines published at national level has been carried out, with the scope of identifying the elements relevant for the development of guidelines within the project. The structure and the content of the guidelines have been defined and reviewed taking into account the comments from Eurostat and the ADMIN network of contact points. A draft version of the guidelines is expected by the end of the current Specific Grant Agreement, i.e. in September 2018, however some areas, for which some methodological work is still required, are expected to be further finalised in the framework of the third specific grant agreement of the Essnet.*

*The paper will describe the quality framework adopted in the guidelines, will present the structure of the guidelines and a sample of the content. The guidelines are structured according to Eurostat statistics quality dimensions and take into account: the different basic data configurations when using multisource data, the phases and sub-processes where the errors generate, examples of measures and quality indicators.*

*This work is carried out in collaboration among the members of the Essnet.*

**Keywords:** output quality, statistical process error sources, multisource statistics

**1. Introduction**

These quality guidelines are developed in the framework of ESS.VIP Admin project of the European Statistical System (ESS) Vision2020 within the ESSnet titled Quality of multisource statistics (Komuso), and specifically in the Specific Grant Agreement 2 (SGA2), Work Package 1 (WP1). Previously, an SGA1 was completed with preparatory work. The general objective of the Working Package 1 is the development of quality guidelines that are practical and applicable in the ESS Member States National Statistical Institutes (NSIs) when administrative data are used exclusively or in combination with survey data to produce statistics.

The guidelines are oriented to the process managers of the NSIs of the ESS member states, who can use them in: supporting the design and the implementation of statistical processes using several sources, addressing the quality aspects that have to be assessed and documenting the quality of outputs based on at least one administrative source.

As the use of administrative sources is becoming more and more widespread for its advantages, various approaches on their integration with more traditional sources are being developed. Thus a need emerged for a comprehensive manual where the use of multiple sources could be outlined and systematised. Indeed, the main requirements for the quality guidelines can be summarised into: to cover the ESS quality dimensions and their meaning when administrative data are used together with survey data; to cover the different basic data configurations for combining sources with different characteristics; to consider the whole statistical production chain; to provide concrete examples and practical recommendations.

Quality guidelines, generally speaking, are agile and short volumes that can offer guidance and suggestions on actions to be taken in order to achieve a satisfactory level of quality (Lyberg *et al.*, 2001). Quality, in itself, not only means accurate and (as much as possible) error-free data but also accessible, punctual, relevant data and so on. Indeed, the multisource guidelines encompass all the ESS quality dimensions - Relevance, Accuracy and Reliability, Timeliness and Punctuality, Coherence and Comparability, Accessibility and Clarity (Eurostat, 2014) - dedicating a chapter to each dimension or pair of dimensions.

As a preparatory activity to the definition of the content and structure of the guidelines, a literature review of existing national guidelines was carried out. Based on its results, the structure of the guidelines was outlined and a sample chapter about the Relevance quality dimension was developed. A consultation on it occurred so that the main feedback could be taken over to the other chapters. As required by Eurostat, some of these chapters will be objective of the third SGA.

This paper will present the work done until now and what remains to be done. The main findings of the review carried out on existing quality guidelines are reported in section 2, the structure of the developed guidelines and their main characteristics are described in section 3. Section 4 will discuss the quality framework underlying the guidelines. Finally, section 5 will describe the remaining work expected to complete the manual.

**2. Review of existing quality guidelines**

The first activity that was carried out and that established the foundation of the multisource guidelines was the literature review on a selected group of national guidelines. The review was based on a questionnaire which focused on specific aspects to be evaluated, among the others, the quality dimensions considered in the reviewed manual, the use of the Generic Statistics Business Process Model – GSBPM (Unece, 2013), the inclusion of suggestions specifically oriented to multisource statistics or to the use of administrative sources, the quality indicators reported. This activity highlighted the common traits and the differences between national guidelines and has allowed us to come to a better understanding of what topics were discussed and how they were dealt with. Most national guidelines resulted more process-oriented than product-oriented; the use of GSBPM, or an adaptation of it, is quite widespread; while many manuals have specific sections dedicated to administrative data or addressed them within relevant chapters, a multisource approach was rarely described in detail, exposing a current gap regarding quality in a multisource environment.

**3. Characteristics of the guidelines**

Since quality is at the core of the production of official statistics by NSIs, each chapter, as explained in the introduction, discusses one or two quality dimensions and how they are affected - if they are affected at all - by the adoption of a mixed-source framework.

Regardless of the quality dimension taken into consideration each chapter follows the same structure and is designed not only to illustrate the corresponding quality dimension(s) in the new context, but also to describe the errors affecting the dimension and the GSBPM phases and basic data configurations where such errors could occur. Specifically, each section deals with the following points[[1]](#footnote-1):

* Definition of the dimension in the framework of multisource statistics
* Main sources of error affecting the dimension
* GSBPM and multisource basic data configurations
* Basic principles and quality guidelines
* Examples of measures/indicators

Not every dimension is impacted in the same way when moving from a traditional survey environment to a mixed-source environment; we acknowledge this in the first section of each chapter, describing the potential changes in the new context for the dimension analysed in the chapter. More detail about the errors that may affect the dimension is given in the second section. Of course, depending on the dimension considered, this section may be simple or longer as there is a different number of errors affecting each dimension. For example, while Accuracy may be threatened by multiple limitations of the administrative data (frame errors, representativity errors and so on), Relevance is mainly affected by validity errors.

A description of the relevant GSBPM phases and sub-processes in which these errors may occur follows, along with some considerations on the relation between the quality dimension and the basic data configurations defined in the previous SGA1. Configurations are ways by which data from different sources can combine. Different aspects have to be taken into account in these combinations: the overlapping of units and variables, potential coverage issues, the aggregation levels of the data and so on. Configurations are relevant because methods and applications may change significantly depending on the combination of input data used in a statistical process.[[2]](#footnote-2)

After this section the central part of the chapter starts, where principles and corresponding actions (guidelines) are presented. Principles are focused on three main lines of intervention: prevention, monitoring/adjustment, assessment. For each principle a series of actions are listed, from the simplest to implement to the less immediate ones. Actions can be focused on a single component of the process (survey or administrative component) or on their integration. They can be considered as short suggestions, without excessive technical or methodological details, that process managers can follow at different stages of the process. They also include some references to well established quality indicators that can be used to monitor or assess quality.

Indicators and quality measures, which are described next, on the other hand are meant to be technical applications of the most important aspects in quality monitoring or assessment for the dimension discussed in the chapter. They are examples drawn from the WP3 guidelines and provide a more practical perspective to the actions proposed. Since there are not established methods yet to assess overall output quality in a multisource context, these applications can be useful for process managers dealing with specific data configurations.

**4. The quality framework underlying the guidelines**

The quality framework adopted in the guidelines integrates together various quality perspectives. As already mentioned the guidelines are developed considering the ESS output quality dimensions and the statistical sources of errors that impact on them. Obviously some dimensions are affected to a very low extent by statistical sources of errors, if not at all (Agafiţei *et al.*,2015). Considering Accuracy, the statistical sources of errors are reflected in the Total Survey Error and its metric, the Mean Squared Error. However, error evaluations including all the sources of errors are barely carried out and the tendency is more towards the development of methods for the reduction of specific sources, especially required in the new production environment using multiple data sources, multiple modes and multiple frames (Lyberg and Stukel, 2017). This has led to systematise in the guidelines the classification of the sources of errors that can arise in the survey and in the administrative source components. As well experienced in the survey context, errors reduction and monitoring is realised through quality management of production process, thus the guidelines make suggestions on concrete activities and quality indicators that can be carried out in order to prevent and monitor the errors, respectively. The suggested quality indicators are those defined among the ESS quality and performance indicators (Eurostat, 2014), however inspiration from other references has also been taken (ESSnet – Komuso, 2016). Nevertheless, it is important to develop methods and increase the experience on the assessment and estimation of the expected impact on the final statistical results of the sources of errors. Therefore, in the guidelines also methods in that direction are suggested. In addition, reference to the “Guidelines on output quality measures” and applications developed in WP3 of the current SGA2 is made.

In the following, the statistical quality framework, represented by the output quality dimensions and the main sources of statistical error relevant for the dimension, is presented.

It is assumed that Relevance is mainly affected by the *validity error*, that is the error arising from discrepancies between the target concepts and the surveys questions, referred to as specification error in Biemer and Lyberg (2003) terminology, and discrepancies between the target statistical concept and the administrative concept, sometimes requiring and harmonisation process (Zhang, 2012) used to measure it.

Accuracy and Reliability are the quality dimensions more affected by statistical sources of errors. The taxonomy used in the guidelines for the type of errors arising in the survey and in the administrative data components is reported in Table 1.

Errors concerning the frame used for sampling selection and the administrative source used to complement data collection in a multisource process are grouped together in the category named *frame and source errors*. Undercoverage, overcoverage, duplications and errors in the auxiliary variables are errors possible both in the frame and in the administrative source, however their impact will be different. In the administrative source component, they can due to delays or failings in reporting administrative events (Zhang, 2012). It is assumed that missclassification errors in contact information are here not relevant since the units identified by the administrative data will not be contacted (although the relative variables will be relevant in the integration procedures).

The *selectivity error* category includes both the nonrespondent units and missing in the administrative data different from undercoverage, under the assumption that the existence of the underlying unit is known.

*Measurement error and item missingness* are grouped together since this is functional to the development of the principles and guidelines of this manual. The term item missingness is used instead of item non response to be more general and include also administrative source.

Considering *processing errors*, it is assumed that data capturing will not be present in the administrative data component. Mapping, Identification / Alignment and Unit Errors are categories pertaining to the use of administrative data (Zhang, 2012), since it is supposed that in planned observation, i.e. in the survey component, these operations are simple and less likely to generate errors. Integration is a relevant step in multisource statistics. It includes record linkage as well as matching procedures. The deriving errors are generally classified into false links and false nonlinks.

Lastly, model assumption errors may arise in every sub-process where a model is assumed. Here only some important steps are considered. It includes also analytic error (West et al., 2017).

**Table 1. Accuracy and Reliability and main sources of error in survey and administrative data (admindata) components**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sources of error**  | **Type of error included** | **Survey** | **Admindata**  |
| Frame and source error | Undercoverage  | x | x |
| Overcoverage | x | x |
| Duplications | x | x |
| Missclassification in the contact information | x |  |
| Missclassification in the auxiliary variables | x | x |
|  |  |  |  |
| Sampling error | Sampling error | x |  |
|  |  |  |  |
| Representativity error | Unit nonresponse | x |  |
| Missing in the accessed data set |  | x |
|  |  |  |  |
| Measurement error and Item missingness | Arising from: respondent, questionnaire, interviewer, data collection | x |  |
| Fallacious or missing information in administrative source |  | x |
|  |  |  |  |
| Processing errors | Data capturing error | x |  |
| Coding error | x | x |
| Mapping error |  | x |
| Editing and imputation error | x | x |
| Identification / alignment error |  | x |
| Unit error |  | x |
| Integration error | x | x |
|  |  |  |  |
| Model assumption error | Seasonal adjustment errors, editing and imputation, … | x | x |
| Model-assisted design based estimation error | x |  |
| Model based estimation error  |  | x |

It has to be mentioned that errors propagate. Some examples are: missclassification in the contact information may cause nonresponse in the survey component; integration error, i.e. false non link and false link, may introduce coverage and measurement error; item missingness in administrative source may cause under or over coverage if referred to variables reflecting an entrance or exit event.

Finally, it can be assumed that all the sources of errors considered for accuracy and reliability, impacting on the precision and unbiasedness of the estimates, have a consequent effect on coherence and comparability. However, the idea here is to focus on the most important sources for each quality component, not all.

With regard to the Coherence and Comparability quality components, it is assumed that validity and model assumption errors are the most important sources of statistical errors. In addition, some factors concerning the stability over time and space of the administrative data legislation may be relevant for these output quality dimensions.

Concerning Timeliness and Punctuality, a prominent role is played by the efficiency of the process performance and the possibility to have timely data sources (both administrative and collected by survey data). However, the progressive nature of the update of the administrative data can require a further evaluation (ESSnet - Komuso, 2017). Delays in the recording of administrative events in the administrative sources, can orient the process managers towards a later timing in the acquisition of the administrative data set, therefore affecting the final timeliness of the results.

Finally, Accessibility and Clarity are scarcely, if not at all, affected by the multisource nature of the statistics. Relevant issues in this area concern metadata management and reporting, and dissemination policies.

 **5. Future work**

As agreed with Eurostat, not all the principles and quality guidelines for each quality dimension have to be developed in the current SGA2, but some of them, for which further work and reflection is needed can be postponed to SGA3.

After taking into consideration different factors, among which the parallel work of other WPs, it was evaluated that some topics concerning Reliability, Coherence and Comparability could benefit from further analyses and experiences, and it was proposed to develop the relative chapters in the next SGA. The proposal was agreed with Eurostat.

The future developments will need to be carried out following the main structure and quality framework of the manual. However, once the remaining chapters are draft, the whole manual will need to be reviewed, and minor adjustments or integrations will be possible in order to make it fully coherent and complete.

**5. References**

Agafiţei M., Gras F., Kloek W., Reis F., Vâju S. (2015), Measuring output quality for multisource statistics in official statistics: Some directions, Statistical Journal of the IAOS, 31, pp. 203-2011.

Biemer P.P. and Lyberg L.E. (2003), Introduction to Survey Quality, J. Wiley, N.Y.

Eurostat (2014). ESS Handbook for quality reports. Available at: <http://ec.europa.eu/eurostat/documents/3859598/6651706/KS-GQ-15-003-EN-N.pdf> (Accessed 10 May 2018).

ESSnet on Quality of Multisource Statistics – Komuso (2016). SGA1, WP1. Checklist for Evaluating the Quality of Input data. Available at: <https://ec.europa.eu/eurostat/cros/system/files/essnet_wp1_report_final_version4.pdf> (Accessed 15 May 2018)

ESSnet on Quality of Multisource Statistics – Komuso (2017). SGA1, WP2. Quality measures and indicators of frames for social statistics. Available at: <https://ec.europa.eu/eurostat/cros/system/files/wp2-framequality-finalreport.pdf> (Accessed 15 May 2018)

Groves R. M., Fowler F.J.Jr, Couper M., Lepkowsky J.M., Singer E., Tourangeau R. (2004), Survey Methodology. Wiley, New York.

Lyberg L.E., et al. (2001), Summary Report from the Leadership Group (LEG) on Quality. Available at: <http://www.istat.it/it/files/2011/11/LEG-on-quality.pdf> (Accessed: 10 May 2018).

Lyberg L.E., Stukel D.M. (2017), The roots and evolution of the Total Survey Error concept. In Total Survey Error in Practice, editors P.P. Biemer, de Leeuw E., Eckman S., Edwards B., Kreuter F., Lyberg L.E., Tucker C.N., and Brady T.West, Wiley series in Survey Methodology, Hoboken, New Jersey.

UNECE (2013), Generic Statistical Business Process Model (GSBPM) (Ver 5.0., December 2013). Available at: [https://statswiki.unece.org/display/GSBPM/GSBPM+v5.0?preview=/97356247/97519760/GSBPM%205\_0.docx](https://statswiki.unece.org/display/GSBPM/GSBPM%2Bv5.0?preview=/97356247/97519760/GSBPM%205_0.docx) (Accessed: 16 May 2018).

West B.T., Sakshaug J. W., Kim Y., (2017), Analytic Error as an Important Component of Total Survey Error. In Total Survey Error in Practice, editors P.P. Biemer, de Leeuw E., Eckman S., Edwards B., Kreuter F., Lyberg L.E., Tucker C.N., and Brady T.West, Wiley series in Survey Methodology, Hoboken, New Jersey.

Zhang Li-Chung (2012), Topics of statistical theory for register-based statistics and data integration, Statistica Neerlandica, Vol 66, 1, pp. 41-63.

1. This outline was presented within the first deliverable of WP1 along with the results of the literature review. [↑](#footnote-ref-1)
2. The guidelines follow the configurations identified in the SGA1: complementary microdata sources, overlapping microdata sources, overlapping microdata sources with under-coverage, microdata and macrodata, only macrodata, longitudinal data. [↑](#footnote-ref-2)