

**UNITED NATIONS STATISTICAL COMMISSION and  
ECONOMIC COMMISSION FOR EUROPE**

**CONFERENCE OF EUROPEAN STATISTICIANS**

**Work Session on Statistical Data Editing**

(Ottawa, Canada, 16-18 May 2005)

Topic (i): Editing administrative data and combined data sources

**EVALUATION OF EDITING AND IMPUTATION SUPPORTED BY ADMINISTRATIVE  
RECORDS**

**Supporting Paper**

Submitted by the Central Bureau of Statistics, Israel <sup>1</sup>

**I. INTRODUCTION**

1. Administrative files support editing and imputation (E&I) processes directly and indirectly: On the one hand, they enable error detection and cold-deck imputation that enhance the end product, and on the other hand, administrative files improve model based imputation and lead to a better prediction of missing values (Ardal & Ennis, 2001; Ruotsalainen, 2002). In their direct use, they are used as an independent reference against which the edited data are compared. Once more than one administrative data-source is used, the comparison criteria are less direct and rely on integrated estimates. Furthermore, administrative data are used to simulate an alternative editing process after which the edited files are compared (Barcaroli & D'Aurizio, 1997; Di Zio et al, 2002).

2. In many cases, administrative files are considered to contain true values; on a micro level, individual records are linked and values are borrowed, and on a macro-level, shapes of frequencies are adopted and variance and variability are kept. However, administrative files carry their own errors and since they are initially intended to serve other agencies and reasons, the definition of their population, variables and categories do not always correspond to those of the edited file. Hence, the administrative data should support the formation of a reference-file, accepted as a good proxy for the true values, rather than serving as such. The reference-file can be used as a ruler by which other data-files are measured and edited, but it can also be the target data-file itself. When it is the target file, the selection of data sources usually use registers as pivot-files, since they cover whole populations and carry their basic identifying attributes (Blum, 1999; Roos et al, 2001; Shlomo, 2003). Other characteristics are selected from different administrative records, while controlling coherence within the created file.

3. The logic behind the building of a reference file serves the process of multi-source data collection. The selection of the best data source for populations and variables, while controlling quality within the created file, includes embedded editing and uses editing logic and tools. This reduction of future editing by applying its rules in the collection stage rather than postponing it to the post-collection stage, expands the meaning of editing to include processes of editing-prevention actions.

4. This extensive use of administrative data necessitates quality evaluation and control, throughout the statistical production process. The evaluation encompasses products, like the data sources, the raw file and the end-file, and elected processes among which are E&I. The evaluation apparatus depends on the available resources; it relies on internal indices, calculated within the treated file, on external attributes like frequent use or interests involved that lead to accurate and updated registration, and on external sources,

---

<sup>1</sup> Prepared by Olivia Blum [blum@cbs.gov.il](mailto:blum@cbs.gov.il)

usually administrative ones. The available resources stipulate different combinations of quality dimensions, which are relevant to the imputation-type involved, and answer the needs of the end users.

5. The structure of this paper is as follows: Section II refers to the merits of using administrative data for editing and imputation and other supporting and supported processes. Section III refers to the quality dimensions relevant to the different imputation procedures, stipulated by available data sources. Section IV expands editing to a multi-source data collection and section V closes the arguments with contemplation about future developments.

## II USING EXTERNAL DATA-FILES FOR E&I: CONCEPT AND MERITS

6. External information supports both parts of data editing: editing, in its narrow meaning, and imputation. *Detecting* possible or certain errors in a data-file requires the implementation of logical rules within or between data sets, or the existence of a ruler or a reference-file against which the values are compared. *Correcting* errors in a data-file entails a well-defined model to predict the value of a variable, the availability of records, variables and values to enable the implementation of the model, or an external true value, or its proxy, to be imputed. While detecting errors is better off with the support of additional data sets, the correction will more often rely on them, if they are available and surpass a quality threshold.

7. An illustrating analogy would be the editing of a proverb. One can locate the suspected or certain wrong word in the proverb from the sentence itself, once it doesn't make sense. However, the introduction of external information, that includes proverbs, enables the comparison, the indication that there is a difference between the sentence to be edited and the external proverbs, or the identification of the wrong word itself. The value added to the error location process by the external information is two folded: the tagging of the sentence as erroneous even if it makes sense (avoiding false-positive errors), and the specification of the word that causes the sentence to be illogical (the cause of the failed edit-check). Correcting a word in a sentence, when no additional information is brought about, results in a logical sentence whether it forms the right proverb or not. However, it is just rational to turn to accessible proverbs-book (true file) or quality texts with proverbs, rather than to choose a word from a collection of possible ones, since they form an already agreed upon, accepted proverb.

### A. Mechanisms and Merits

8. Administrative records support E&I via three main mechanisms: the enrichment of the relevant information needed (Roos and Roos, 2001), the expansion of the ability to create a ruler or a relatively accurate reference-file, and the continuous quality assurance performed throughout the statistical production process.

9. The additional data files allow for a better specification of an imputation model, either by adding variables not included in the treated data file, or by adding records that improve the representation of small groups in the edited file. In a hot-deck imputation, which can be considered as a special case of a model-based imputation, the administrative records improve and refine the selection of sort-variables and enrich the pool of possible donors. Moreover, the augmented information is used directly, on the individual level, in a cold-deck imputation. In the previously phase of editing, administrative records enable error locating by identifying unexplained differences in values of variables involved in failed edit-checks.

10. In their role as an external ruler, the administrative records support the editing process in locating errors and corroborating values in variables involved in failed edit check; they improve the ability to identify the erroneous variable. Moreover, they help to avoid false-positive errors, through a comparison between the edited file and the reference-file. As for the imputation process, if specific variables in a single administrative data source are found to be qualitative enough to serve as true values, cold-deck imputation is called for. This is usually the case in registers, like the population register, where variables like place and date of birth are accurate for people who were born in the state that the register refers to, after the establishment of the register. Moreover, registers supply the population or the sampling frame, and as such, they facilitate the identification of missing records in the edited data-file and enable cold-deck imputation of missing units.

11. However, if no single register or administrative source is found to carry accurate values, every additional file with variables common to both, improves the ability to create a reference-file with relatively

accurate values. This process involves the selection of the best value using different processes: *the majority rule*, which implies the selection of the value that most files carry independently; *the qualitative-file rule*, that tests the file rather than the variable and implies the selection of values from the data source with the highest quality indices; *the corroborated-variable rule*, implies the selection of values of variables that were corroborated by empirical tests.

12. The third mechanism, in which E&I is supported by administrative records, is the continuous quality assurance. It is more of a preventive measure. Edit-checks and values comparisons between the processed data and the administrative data are carried out in order to identify missing values, errors in the collected data and errors added during data processing. Since it is done online, the error source itself can correct the error or indicate that the edit-check is incorrect, almost immediately. Future E&I can be significantly reduced, especially the micro-level errors.

13. Recapitulation-Table 1: Merits of using administrative records in E&I, by mechanism.

	<b>Mechanism</b>	<b>Supporting Editing</b>	<b>Supporting Imputation</b>
1	<i>Enrichment of information</i>	<ul style="list-style-type: none"> <li>• Corroborating values.</li> <li>• Locating possible erroneous variables in failed edit-checks.</li> </ul>	<ul style="list-style-type: none"> <li>• Enabling cold-deck imputation.</li> <li>• Improving the specification of the imputation model, including small area estimates and sort-variables in hot-deck.</li> </ul>
2	<i>Creating a Ruler or an accepted Reference-File</i>	<ul style="list-style-type: none"> <li>• Corroborating values.</li> <li>• Locating the erroneous variables in failed edit-checks.</li> <li>• Avoiding false-positive errors</li> </ul>	<ul style="list-style-type: none"> <li>• Enabling qualitative cold-deck imputation.</li> <li>• Improving model-based imputation, including frame related errors.</li> </ul>
3	<i>Continuous Quality Assurance</i>	<ul style="list-style-type: none"> <li>• Reducing micro-level editing:</li> <li>• Online editing, which enables error detection during data collection.</li> <li>• Avoiding errors accumulated throughout the production process.</li> </ul>	<ul style="list-style-type: none"> <li>• Reducing the need for imputation.</li> </ul>

## **B. Other Aspects of E&I Supported by Administrative Records**

14. Extending farther the meaning of E&I, incorporates other aspects of data processing to those who benefit from the support of administrative records, through the enriched information mechanism. Three of the most routine aspects are record linkage, coding and imputation of new variables.

15. Record linkage is a tool to facilitate and enable heterogeneous statistical processes. One of which is the definition and identification of the entity behind the record. Once the records of the same individual are linked, critical variables can be corrected using cold-deck imputation. For example, when the individual's unique ID number in a survey is not identical to the one in the register, but a linkage, based on other variables, is possible, the ID number in the survey data-file can be corrected. It enables further linkages with less variable-rich data-sources. Statistical record linkage on an aggregate level, using a group profile, supports macro editing and hot-deck imputation and it also improves model-based imputation through the enrichment of relevant data.

16. E&I are embedded in the process of coding; especially in the more complicated variables, like industry and occupation. Actually, each coding process, which is not based on exact match, engages editing. It is implemented directly in the text, or indirectly, by the interpretation of the text within a context, drawn by other variables. The administrative data support the interpretation by creating or enhancing the context.

17. As for imputing new variables, there are two main aspects to be considered; the direct cold-deck imputation of variables not included in the basic data-file or of new indices calculated with the support of external data, and the replacement of the data source by an administrative one. New variables and indices are possible due to the unique information in designated data files, built for administrative purposes. For example, an index of well-being, on the locality level, can be calculated based on census data or it can be better specified by including data from other sources, like the subsidies per capita from the social security

institute and the air pollution level from the environmental ministry. The inclusion of data from other sources in the editing process is going a long way, from the substitution of erroneous values to the substitution of data source. In this context, it is a preventive measure, carried out when an extensive E&I is foreseen, either for correction or for supplementing the data. Further implications are elaborated in section IV.

18. Recapitulation-Table 2: Extension of E&I supported by admin records, by process

	<b>Process</b>	<b>Process Feature</b>	<b>Enabled E&amp;I Processes</b>
1	Record linkage	Individual level	Cold-deck imputation
		Aggregate level	Macro editing
2	Coding	Text	Text editing
		Context	Context interpretation
3	Imputation of new variables	Calculated indices and variables	Direct and indirect cold-deck
		Substitution of data source	Ex-ante E&I

### III USING EXTERNAL DATA-FILES FOR E&I: QUALITY DIMENSIONS

19. Evaluation of quality of an edited data set depends on whether there are other data sets that have been used in the E&I process or can be used for evaluation. Moreover, the supporting data influence the scope of the possible evaluation of the E&I results. The number of administrative files available, the interdependency between them and their coverage and content have an effect on what quality dimensions are applicable.

20. Quality dimensions that are directly related to E&I processes are coherence within, consistency between, comparability, completeness and accuracy. Other dimensions are relevant to the output product and are not directly related to E&I. Among them are core dimensions like relevance, accessibility, timeliness and punctuality, and non-core dimensions like response burden, confidentiality and data integrity. The following theoretical analysis refers to those directly related to E&I.

21. The concepts of coherence and consistency are used at times as synonyms with regard to quality dimensions, however, it is functional to separate them, as implied in Holt and Jones (1998); Coherence is measured within a file and refers to the relative number of failed edit-checks (Elvers and Nordberg, 2001). Consistency is measured between data sources and refers to differences of definitions, data collection processes, etc. These two dimensions are independent in so much as the data set can be coherent within but not consistent with other data-sets, or it can be consistent with others but not coherent within. Another closely related quality dimension is comparability. Comparability is a necessary but not sufficient condition for consistency. Data sets can be comparable in spite of the differences between them, since they share data elements like population and variables. The results of the comparison may point at the attributes that cause the inconsistency and require harmonization. E&I improve quality measured in terms of all three: coherence, consistency and comparability.

22. Another multifaceted quality-dimension is completeness. It refers to missing values and missing records before E&I, and those remained missing after editing. Incompleteness can be a result of a choice (not to impute values in specific variable, for example), impossibility (when no model can be specified or too many values are missing, etc.) or as a result of unawareness (like missing records due to partial frame). Improvement of quality as a product of E&I measures relative completeness rather than the absolute one.

23. The most common and basic quality dimension in E&I processes is accuracy (Holt and Jones, 1998). Accuracy is a relative term, measured against accepted or agreed upon values, hence the importance of the availability of external information. As most quality dimensions, accuracy measured by bias and variance, support decision-making during editing and serve a process of continuous quality assurance, which feeds back the E&I.

24. The above quality aspects are not always applicable. They depend on the data sources available and on the editing processes that have taken place, which in turn, depend on the data sources used. The increasing abilities to implement E&I processes and to evaluate the derived statistical quality is unfolded as data sources from different types are added, either by record linkage or with longitudinal data (Roos et al, 2001);

25. If the edited file is the only data source to be used, one may detect errors by activating edit-checks that are implied by logical rules, concerning the relations between variables within the dataset. One of the main problems is how to determine what value involved in a failed edit-check is the wrong one. The idea of erasing minimum values that satisfy maximum rules is one of the solutions suggested and is applicable in many datasets. The limited abilities to manipulate data in order to correct it are also manifested in the imputation process. Cold-deck imputation is not possible, while hot-deck and model-based imputation depend on the richness of the attributes within the dataset and the correlations between them. More variables enable a better specification of an imputation model if they are correlated and characterize homogeneous groups. In this realm of a single data-file, editing is done on a micro level. Macro-level problems are not easy to detect and once they are identified, correction is quite impossible since no external information is available. For example, a macro-level failure can be a too high proportion between children and women in the fertility age. How and where to correct is a question to be answered with more information than the edited file alone. As for the quality dimensions, quality can be evaluated by internal cohesiveness and by completeness related to the attributes within the file, whether they are item nonresponse or erased errors, and completeness related to unit nonresponse within the sample. Problems of a frame are not presented.

26. If the edited file is a part of a survey conducted in several waves, or part of an ongoing survey, the E&I possibilities are expanded. Potential errors are detected by comparing waves, cold-deck imputation can be used by borrowing background variables or by borrowing slow-pace changing attributes, and model-based imputation is enhanced, by including changes over time. A limited macro level editing is also introduced, if an accepted dataset has been formed. Evaluation of quality is also broadened by consistency and comparability between datasets of the same survey and by a limited accuracy dimension, measured against the best 'true file' the same survey in different time points could supply.

27. The existence of other surveys or administrative files widens the information scope that comes from independent sources. Macro editing has the supporting information needed for detecting errors and correcting them, and cold deck imputation is expanded to more variables and has more candidate values to be imputed. Furthermore, the introduction of additional data, coming from independent sources enables a better-specified imputation model. The quality evaluation is more reliable, especially the accuracy dimension, since this wealthy information facilitates the development of an accepted reference-file.

28. The uniqueness of registers, with regard to other administrative sources, is the completeness of the frame and the continuous updated information. Registers are reference-files and as such, they influence directly and indirectly processes and results. Hot-deck imputation is an example of an indirectly influenced process. After exhausting the cold-deck imputation with the support of the register, hot-deck imputation is engaged for missing values of variables not included in the registers. The register information improves the stratification of the population in the edited file, needed for hot-deck imputation to be performed within homogeneous strata. Other sources of information may do it as well, but registers have the advantage of having the whole population, as censuses do in infrequent points of time. This attribute contributes to the completeness dimension when evaluating coverage of a relevant population.

29. Recapitulation Table 3: Quality dimensions by imputation type and by data source

	<b>Sources of data</b>	<b>Imputation type</b>	<b>Level</b>	<b>Quality dimension</b> (Relevant to E&I)
1	File to be edited	Model-based Hot-deck	Micro	<ul style="list-style-type: none"> <li>• Coherence within</li> <li>• Completeness (sample related)</li> </ul>
2	File to be edited Neighboring waves (same survey) Previous rounds (same survey)	Extended model-based Hot-deck Cold-deck	Micro & Macro (Partial)	<ul style="list-style-type: none"> <li>• Coherence within</li> <li>• Consistency between sources</li> <li>• Comparability over-time- same survey</li> <li>• Completeness (sample related)</li> </ul>

				<ul style="list-style-type: none"> <li>• Accuracy (limited)</li> </ul>
3	File to be edited Other surveys Administrative files	Extended model-based Hot-deck Extended cold-deck	Micro & Macro	<ul style="list-style-type: none"> <li>• Coherence within</li> <li>• Consistency between sources</li> <li>• Comparability between sources</li> <li>• Completeness (sample related)</li> <li>• Accuracy</li> </ul>
4	File to be edited Registers (Census)	Extended model-based Extended hot-deck Extended cold-deck	Micro & Macro	<ul style="list-style-type: none"> <li>• Coherence within</li> <li>• Consistency between sources</li> <li>• Comparability between sources</li> <li>• Completeness (sample and frame related)</li> <li>• Accuracy</li> </ul>

#### IV. EVALUATION OF E&I IN A MULTI-SOURCE DATA COLLECTION

30. The use of several data sources in the data collection phase presents questions with regard to the meaning of editing and the statistical production-process that it encompasses. There are many reasons for the use of administrative sources to substitute direct data collection. However, in the editing context, a rational decision making, which leads to indirect data collection and integration of several data sources, have to see the merits in terms of a lesser need for E&I, easier implementation and a high quality end-product.

31. Editing has always been a part of data collection. Field supervisors have checked the paper questionnaires and returned them to the interviewers in order to correct errors or to complete missing values. The introduction of computer assisted data collection, has allowed for avoiding errors by implementing online edit-checks, where the interviewers can correct either the answers given or the data capture, while still interviewing (Bethlehem and Hofman, 1995; Wein, 2004; Kozjek, 2004). The use of administrative records as an alternative data source deepens editing avoidance by bringing its implementation to an earlier stage of the statistical data production, the selection of data source.

32. Since the issue of data quality is not the only reason for multi-source data collection, editing is not shunned completely and new types of errors are introduced. Consequently, evaluation of the data to be used is a pre-requisite. Accuracy and completeness are the critical quality dimensions that support decision-making concerning sources of information to be used, in order to form the population frame and fill-in its relevant attributes. For example, a good population register has to be used as a pivot, since it covers the designated population (Roos et al, 2001; Shlomo, 2003) and its geo-demographic attributes. So is the case with a business register. However, if registers are not available, the population frame is a result of an extensive record linkage between data sources, which have been evaluated and identified as relevant, regarding completeness of units and their unique identification. The incorporation of the population attributes is also stipulated by previous evaluation that refers to the data source and the relevant variables in it. For example, files of authorities that provide subsidies have a relative accurate geographic address since the people have an interest to be located. They may also provide information about wages and income although downward biased.

33. Quality of a data source at this stage is a result of frequent use and interests. Frequent use acts as an ongoing quality check. For example, files of authorities that sell services, like telephone and electric companies, are used and updated on an ongoing basis. Therefore, they are expected to include all service buyers with the attributes relevant to the transaction. As for interests, they involve subsidies and other benefits, given by the file holder. They serve as an incentive to be registered in the administrative file with true attributes, as is the case with children allowances given by the social security institute. Yet, the very same interests may instigate over coverage and biased information, in order to be eligible for the benefits. Therefore, corroborating information is needed; either from another administrative source or from a field survey. Editing, in its error avoidance meaning, is inherent to this evaluation and selection process. Moreover, editing is also embedded in the integration of data sources, implemented in harmonization and ongoing quality assurance processes, in order to ensure coherence, consistency and comparability.

34. Despite the careful building of the raw data file, additional editing is required since attributes like data availability and costs interfere with the quality considerations. Imputation is needed since sources of data are either carrying missing values or having values that lead to failed edit-checks, and therefore are erased. Macro editing is required since it is more difficult to perform online, while building a file. The problem is that the external accepted reference-files have already been incorporated into the data collection stage. It means that all these files can be used partially, if at all, by using directly and indirectly variables and values that have not been used. The availability of other administrative files, not used in the data collection stage, may improve the E&I processes. Editing, at this stage, plays its traditional role of correcting the errors accumulated in previous stages; errors made in the selection process and errors originated in the data sources themselves.

35. Evaluation of the end result is the third time the data are evaluated; the first is the evaluation of the separate data sources, the second is the evaluation of the raw data file, and the third is an out-going quality evaluation that refers to the end file. It is a result of the expansion of the editing life cycle, in Luzi and Manzari (2000) terms. Evaluation of the end result includes quality dimensions measured in relative terms and refer to the quality change contributed by E&I. The main challenge at this stage is finding or building an independent reference-file to serve as the truth, since the process of building the file and editing it has already used evaluation and selection of the best data source.

36. There are two main approaches for evaluation in this situation: The first is to point at the distances of key attributes from the truth, which is represented by a new accepted reference file, and the second is to evaluate key processes that lead to the end file (Nordbotten, 2000).

37. Qualitative administrative files that have not been used or a combination with files that have been used can serve the first approach. It is possible to use an already used file since the result of the different combination may vary significantly from the one created and used before (although it questions the stability of the selection results). An alternative data source can also be a designated field survey, used as another source of information and not THE true file since it carries its own unique errors (Poulsen, 1997). Comparing the differences between the raw and the reference file, and between the end and the reference file brings about the effect of E&I (Granquist, 1997; Luzi and Manzari, 2000). The implementation of this approach is costly and concentrates mainly on the accuracy and completeness dimensions of quality.

38. The second approach is the evaluation of the processes rather than the products. Its implementation means checking decision rules, repeating processes and checking the effect of different processes within the realm of the statistical operation. For example, there is a leading principle to define edit rules; it can be a number of checks, like the maximum checks possible in a given computer resources, or the minimum checks needed to trace errors between key variables only. An alternative principle can be a context related one, like edit-checks that keep maximum coherence within the file. The activation of different principles, possibly on samples, enables the evaluation of the principle chosen and the edit-rules implied.

39. The process of evaluating processes may end up with many segmented results or it can add up to a single evaluation picture. In order to have the later, the evaluated processes have to be complementary. For example, the evaluation of imputation should include the evaluation of the main processes and methods used, in order to assess the quality of the imputation product, rather than the evaluation of the effect of cold deck imputation solely. Evaluation of processes may also call for a field operation and may end up being quite costly. Its advantages reside in cases where administrative files are not available and where critical selected processes have to be evaluated. The quality dimensions attached to the evaluation would be those that were aimed for, in the different processes.

40. Recapitulation Table 4: Quality considerations in a multi-source data collection

	<b>Stage</b>	<b>Editing Type</b>	<b>Quality Considerations</b>
1	Building the raw data-file: Selection and integration of data-sources	Preventive, ongoing	<ul style="list-style-type: none"> <li>• Completeness and accuracy refer to single sources in addition to the integration of the selected sources.</li> <li>• Coherence, consistency and comparability should stipulate the integration</li> </ul>
2	E&I of the raw data-file	Traditional: correction of	<ul style="list-style-type: none"> <li>• The result of the integration carries</li> </ul>

		accumulated errors	source and selection errors. • Evaluation is against alternative sources or alternative selection.
3	Evaluation of the end-file	--	Two approaches: • Creating a new true-reference-file with a combination of old sources or with new ones, administrative or field products. • Evaluating processes rather than products.

## V. FURTHER CONTEMPLATION AND CONCLUDING REMARKS

41. In a realm where multiple data sources are available and accessible, the boundaries of editing and imputation become fairly blurry. Administrative files support errors detecting and correction, yet they are evaluated as a potential data source with the very same logic and rules used for editing. Both uses aim to the same goal; a data-file of the quality traits that the end users ask for or need.

42. As a result of the moving boundaries and because of the extensive use of administrative data, the line between editing and evaluation is constantly moving. It happens in spite of the different goals of the two, since evaluation tools and dimensions are also used for the definition of edit-rules. Moreover, the process and results of the selection of data that have passed a quality threshold, when developing a reference true-file, is actually narrowing down the options opened for imputation. It dictates the imputation processes and essentially lays down part of the imputed values.

43. In this sphere, there is a real danger of creating a tautology by using administrative data that have supported editing, for evaluation purposes. Since the unlimited wealth of administrative information is still utopia, and since more and more statistical offices strive to exhaust the use of administrative data as an alternative data-source for a fieldwork operation, it seems that independent evaluation operation will be THE process to be carried out in a field survey.

44. The challenges of evaluating a complex product, that relies on administrative information, in a fieldwork operation, is in getting the big picture from which future work will benefit. The samples are required to be efficient since they are meant to enable the quality evaluation of the administrative files, the different processes that led to the final statistical product and the evaluation of the end product itself. If the result is a large sample and elevated costs, direct data collection will be again part of the possible data sources used for a survey or the only source for it. One way or another, statisticians will have to equip themselves with editing knowledge and tools that will enable a qualitative editing of any source of information or the product of their integration.

## References

- Ardal, Sten and Sherri Ennis (2001). "Data Detectives: Uncovering Systematic Errors in Administrative Databases". Proceeding of Statistics Canada Symposium 2001: Achieving Data Quality in a Statistical Agency: A Methodological Perspective.
- Barcaroli, Giulio and L. D'Aurizio (1997). "Evaluating Editing Procedures: The Simulation Approach". UNECE Work Session on Statistical Data Editing. Prague, Czech Republic.
- Bethlehem, Jelke and Lon Hofman (1995). "Macro-Editing with Blaise III". IBUC International Blaise Users Conference. Helsinki, Finland.
- Blum, Olivia (1999). "Combining Register-based and Traditional Census Processes as a pre-defined Strategy in Census Planning". FCSM Research Conference. Washington DC, USA.
- Di Zio, Marco, Orietta Luzi and Antonia Manzari (2002). "Evaluating Editing and Imputation Processes: The Italian Experience". UNECE Work Session on Statistical Data Editing. Helsinki, Finland.
- Elvers, Eva and Lennart Nordberg (2001). "A Systematic Approach to Quality Measurement and Presentation". Q2001, European Conference on Quality and Methodology in Official Statistics. Stockholm, Sweden.
- Holt, Tim and Tim Jones (1998). "Quality Work and Conflicting Quality Dimensions". 84<sup>th</sup> DGINS Conference. Stockholm, Sweden.

- Granquist, Leopold (1997). "An Overview of Methods of Evaluating Data Editing Procedures". Statistical Data Editing: Methods and Techniques, Vol.2. UNECE Statistical Standards and Studies No. 48.
- Kozjek, Pavle (2004). "Error Reporting and User Interface for Post-Collection Data Editing". IBUC International Blaise Users Conference. Ottawa, Canada.
- Luzi, Orietta and Antonia Manzari (2000). "Data Editing Methods and Techniques: Knowledge To and From Users". UNECE Work Session on Statistical Data Editing. Cardiff, UK.
- Nordbotten, Svein (2000). "Evaluating Efficiency of Statistical Data Editing: General Framework". UNECE. Geneva.
- Poulsen, Marius Ejby (1997). "Evaluating Data Editing Process Using Survey Data and Register Data". Statistical Data Editing: Methods and Techniques, Vol.2. UNECE Statistical Standards and Studies No. 48.
- Roos, Leslie L and Noralou P. Roos (2001). "Of Space and Time, of Health-Care and Health". Journal of Health Services Research and Policy, Vol.6 (2).
- Roos, Leslie L., Ruth-Ann Soodeen and Laurel Jebamani (2001). "An Information Rich Environment: Linked-Record Systems and Data Quality in Canada". Proceeding of Statistics Canada Symposium 2001: Achieving Data Quality in a Statistical Agency: A Methodological Perspective.
- Ruotsalainen, Kaija (2002). "Use of Administrative Data in Population Censuses – Definition of Main Type of Activity as an Example". UNECE Work Session on Statistical Data Editing. Helsinki, Finland.
- Shlomo, Natalie (2003). "The Use of Administrative Data in the Edit and Imputation process". UNECE Work Session on Statistical Data Editing. Madrid, Spain.
- Wein, Elmar (2004). "Improvement of Data Editing Processes". Q2004, European Conference on Quality and Methodology in Official Statistics. Mainz, Germany.

-----