Quality improvement of administrative registers statistically exploited to generate the indicator-based decision-making system in the State of Yucatan, Mexico

Amparo Ballivian[1], Indu John-Abraham[2], Alejandro Medina[3], Julio Ortegon[4], Ignacio Romero[5], Federico Segui[6,7]

[1,2] World Bank Group, Washington DC, USA
[3] International Consultant contracted by the World Bank, Aguascalientes, Mexico
[4] Secretary of Planning and Budget of the State of Yucatan, Merida, Mexico
[5] National Institute of Statistics, Geography and Informatics (INEGI), Merida, Mexico
[6] Expert on Quality in Official Statistics contracted by the World Bank, Montevideo, Uruguay
[7] Contact author: Federico Segui, e-mail: fsegui@ine.gub.uy

Abstract

Currently, politicians and administrators are increasingly entrusting on measured values and statistics to lead developments within economy and society. The purpose of an indicator-based decision-making system is to provide objective evidence of facts and actions taken. Hence, the system should assure transparency and help to improve the effectiveness and efficiency of governmental programmes. A government project in the State of Yucatan – Mexico aimed to strengthen the state statistical, geographic and evaluation information system (SIEGEY in Spanish) to support proof-based decision-making in the State of Yucatan. The diagnosis phase of this project detected the need of enhance the quality of administrative registers to be used as data source of the statistical, geographic and evaluation information system. The World Bank has supported the development and implementation of a standardized Tool for Quality Assessment of Administrative Registers for statistical use (HECRA in Spanish) in the State of Yucatan. The tool HECRA has been developed focusing on four quality aspects: data source, metadata, data and statistical product (drawn from administrative data). Each of these quality elements consist of dimensions or attributes that are groups of indicators about the same subject. The results of the tool HECRA are then used to discover improvement areas opportunities identifying problems and their root causes, implementing corrective actions to eliminate those causes. This is the basis of the improvement plan of administrative registers that is currently implemented by the State Secretary of Planning and Budget of Yucatan.

Keywords: enhancing access to information, decision making; quality assessment; quality indicators; administrative registers.

1. Introduction

Nowadays, politicians and administrators (decision makers) are increasingly entrusting on measured values and statistics to lead developments within economy and society. The credibility of public administrations and government agencies is enhanced if they base their decisions on indicators. The purpose of an indicator-based decision-making system is to provide objective evidence of facts and actions taken. Hence, the system should assure transparency and help to improve the effectiveness and efficiency of governmental programmes.

Consequently, the Mexican Constitution and its secondary laws prioritize Result-Based Management and Budgeting as a standardized strategic management approach for all levels of government. However, the lack of high quality, reliable and disaggregated information needed for effective decision-making has caused difficulties to accomplish this regulation, particularly at the state level. Therefore, a government project aimed to strengthen the state statistical, geographic and evaluation

information system (SIEGEY in Spanish) to support proof-based decision-making in the State of Yucatan and the establishment of the first State Committee for Statistical and Geographic Information in Mexico. The State of Yucatan is also a pioneer in implementing a State Monitoring and Evaluation System. Projects leaded by the State Secretary of Planning and Budget of Yucatan (SPP) with the cooperation of the National Institute of Statistics, Geography and Informatics of Mexico (INEGI) [1].

The diagnosis phase of this project detected the need of enhance the quality of administrative registers to be used as data source of the statistical, geographic and evaluation information system. Hence, a standardized method to evaluate the quality level of administrative data registers for statistical purposes should be implemented.

The World Bank has supported the development and implementation of a standardized Tool for Quality Assessment of Administrative Registers for statistical use (HECRA in Spanish) in the State of Yucatan. HECRA is mainly based on the quality framework for the evaluation of administrative data developed by Statistics Netherlands (CBS) [2], the quality assessment of statistical products of the National Statistical Institute of Uruguay and a research named "Review of assessment frameworks of statistical quality" by Alejandro Medina (World Bank).

2. Quality Framework

The scope of the tool HECRA is to assess the quality of administrative registers for statistical purposes i.e. that are useful to elaborate statistical products. The tool is not focused on the evaluation of administrative registers from the point of view of the administrative registration, operative, managing, maintaining or the accomplishment of administrative, fiscal or normative requirements, for which they were originally created.

The HECRA has been developed focusing on four relevant quality aspects: data source, metadata, data and statistical product (drawn from administrative data). Each of these quality elements consist of dimensions or attributes that are groups of indicators about the same subject. Most of the quality indicators are quantitative and they are measured through using a questionnaire that must be filled out by the data source keeper and the main user (primary) of the administrative data (for statistical purposes).

The four key quality elements of HECRA:
1. **Administrative data source**. It contains quality attributes related to the institutional environment, legal framework, information security, procedures, delivery and data treatment of administrative registers.
2. **Metadata**. This element is composed by quality attributes linked to information about data of administrative registers.
3. **Data**. Its objective is to assess the quality associated to the data stored into the administrative register file. Technical controls on data, coverage of the administrative register, record linkage methods, unit and item non response, missing data, measurement, identification keys, data processing, data precision, coding, data freshness and multiple records are some of the quality attributes to be assessed.
4. **Statistical product**. The last element determines the quality of statistics, tables and microdata files produced from administrative registers, either as frame of sample surveys, secondary information source, or as generating statistics directly by processing data of the statistical register.

---

[1] "MEXICO: Foundations for a M&E System. A Strategy for the Development of Statistics in the State of Yucatán, México" – Latin American and the Caribbean, Results in Action – World Bank.
[2] "Checklist for the Quality Evaluation of Administrative Data Sources". Discussion paper (09042), Statistics Netherlands.

Table 1. Elements, attributes and indicators of quality of administrative registers.

Quality elements	Quality attributes	Quality Indicators
Administrative data source	1. Relevance	1.1. Utility 1.2. Intended use 1.3. Demand for information 1.4. Satisfaction of primary users
	2. Information security and limitations on the use of the information.	2.1. Legal framework 2.2. Personal data protection 2.3. Limitations due to confidentiality regulations 2.4. Confidentiality agreements 2.5. Secure data transfer 2.6. Confidentiality, integrity and availability of information 2.7. Data protection 2.8. Data backup policies
	3. Data delivery commitment.	3.1. Costs associated with the delivery 3.2. Delivery agreements 3.3. Frequency of deliveries 3.4. Dates of last five deliveries 3.5. Punctuality 3.6. Risks due to lack of data 3.7. Alternative method to replace the lack of information 3.8. Means of data delivery 3.9. File format 3.11. Data selection
	4. Control and continuous improvement.	4.1. Data collection 4.2. Consistency control 4.3. Change control 4.4. Continuous improvement
	5. Data treatment.	5.1. Control of objective units 5.2. Control of variable content 5.3. Control of Outliers 5.4. Changes 5.5. Reasons for not changing 5.6. Changes according to procedure 5.7. Use of Database Management System 5.9. Database documentation 5.10. Database integrity
Metadata	1. Metadata documentation	1.1. Metadata documentation
	2. Completeness and clarity	2.1. Definition of population units 2.2. Description of variables 2.3. Communication of changes in definitions/concepts
	3. Use of unique keys	3.1. Identification keys 3.2. Comparability of identification keys 3.3. Unique combinations of variables
	4. Comparability	4.1. Comparability of the objective unit definition 4.2. Comparability of variable definitions
Data	1. Technical controls	1.1. Readable data 1.2. Redefinition of concepts and metadata in case of more than one data source 1.3. Correspondence between data and metadata 1.4. Record linkage method 1.5. Verification of effectiveness of the record linkage method
	2. Coverage	2.1. Over coverage 2.2. Classification errors
	3. Record linkage	3.1. Rate of record linkage
	4. Completeness	4.1. Rate of unit non response 4.2. Rate of item non response
	5. Measurement	5.1. External control (audit)

Quality elements	Quality attributes	Quality Indicators
	6. Identification keys	6.1. Rate of records with unique key
	7. Data processing	7.1. Data editing 7.2. Imputation
	8. Data accuracy	8.1 Data accuracy
	9. Coding	9.1. Use of standard coding 9.2. Verification of coding 9.3. Rate of coding errors 9.4. Rate of records without code
	10. Data freshness	10.1. More than 90% of the objective units created during year t have been registered before the end of the year t+1
	11. Multiple records	11.1. Rate of multiple records of the same unit
	12. Other controls	12.1. Rate of units with valid values into identification keys 12.2. Tables of the statistical operation have been validated through automatic procedures
Statistical product	1. Comparability	1.1. Length of comparable time series 1.2. Comparability of microdata along the time
	2. Relevance	2.1. Identification of users 2.2. Information about users 2.3. Rate of final user satisfaction 2.4. Utility (intended uses)
	3. Coherence	3.1. Coherence of statistics with different periodicity 3.2. Coherence of statistics with the same socio-economic scope
	4. Availability and clarity	4.1. Accessibility by Internet 4.2. Rate of completeness of metadata
	5. Accuracy	5.1. Coefficient of variation 5.2. Rate of unit non response 5.3. Rate of item non response 5.4. Rate of imputation 5.5. Rate of editing 5.6. Rate of over coverage 5.7. Rate of classification errors
	6. Timeliness and punctuality	6.1. Punctuality of statistical product dissemination 6.2. Length of time between its availability and the event or phenomenon it describes 6.3. Freshness of statistical register

The tool HECRA is composed by the following instruments: "Guide to self-assessment of the quality of administrative registers for statistical purposes"; "Self-assessment questionnaire of the quality of administrative registers for statistical purposes" and its "Guide to completing the questionnaire"; "Questionnaires for the assessment of satisfaction of the administrative register primary user and satisfaction of the final user of statistical product based on administrative registers".

The "Self-assessment questionnaire of the quality of administrative registers for statistical purposes" allows measuring the quality of administrative registers statistically exploited, by means of quality indicators calculated from the answers of the questionnaire.

The four quality elements (administrative data source, metadata, data and statistical product) correspond to chapters and modules of the questionnaire. Each chapter/module is mostly compounded of closed-ended questions with pre-coded responses. Then, the quality indicators are calculated from these pre-coded responses. The self-assessment questionnaire is available on spread-sheet format for easy filling and calculation of quality indicators.

3. Experiences implementing the tool HECRA

Secretary of Planning and Budget of the State of Yucatan has implemented an improvement plan of administrative registers for statistical purposes (generating indicators for the decision-making system) during 2011 and 2012.

The project began with the identification and characterization of the most relevant administrative registers regarding to the information needs for the decision-making system. Therefore, some of them have been selected (education, health, tourism and agricultural registers) to implement the tool HECRA in order to make a diagnosis of the administrative registers quality and execute an improvement plan.

Figure 2. Implementation process of HECRA.
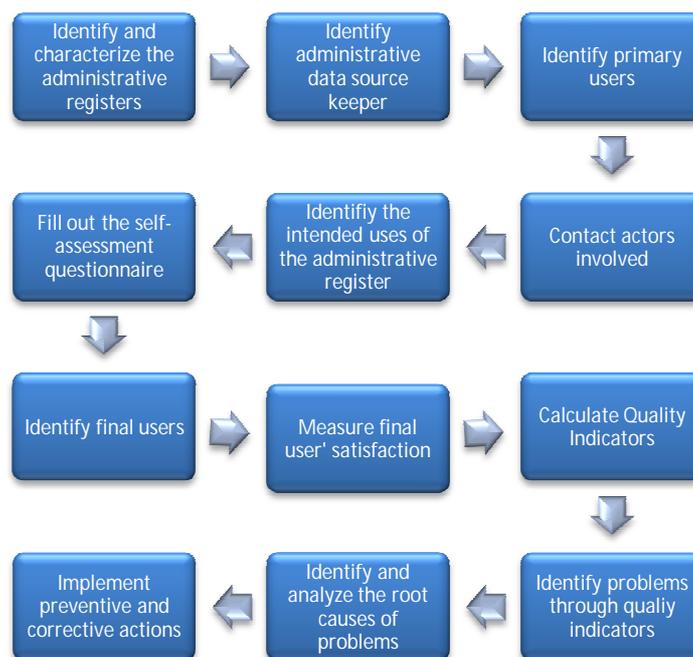
Despite HECRA includes a self-assessment questionnaire that has been designed to be self-administered, consultants expert on each government sector have been hired to apply the HECRA into the administrative records related to their sectors to assure a correct implementation during the first experiences.

Workshops on quality assessment of administrative registers for statistical purposes using HECRA have been organized to administrative data source keepers and primary users.

The pilot projects have allowed making some adjustments on the HECRA to improve its implementation. The final version of the HECRA is a consolidated tool founded on experiences from NSIs and quality experts, but mainly based on the lessons learned during the implementation of pilot projects into the government offices of the State of Yucatan, Mexico.

4. Lessons learned

The main conclusion is that HECRA is a very useful tool to self-assess and diagnose the quality of administrative registers for statistical purposes.

At the beginning, the identification and characterization of the administrative registers was a very difficult process because it was not easy to discover where all administrative registers are located or if they are statistically useful. The process of identification and characterization of the administrative registers should be planned based on the statistical indicators that compound the decision-making system. It means

that the planning of this process should start from the key result indicators of each government program to identify their information needs and lead to the administrative registers that provide that information.

The implementation of HECRA has helped to detect this problem and its solution. The assessment process of HECRA is focused on the statistical uses of administrative registers, thus it naturally guides the process' owners to identify the administrative registers (data source) from key result indicators of government programs (statistical use).

It is highly recommended to implement the quality assessment process step by step, starting from a pilot project with a few administrative registers. Implementing an assessment of all administrative registers at the same time should result in project failure. Priorities should be established based on the impact of the administrative data into the decision-making system defined from the government policy priorities.

It is also healthy to begin with a self-assessment approach and then advance to a higher level of evaluation through external audits. The self-assessment strategy facilitates to introduce the quality culture into the organizations (administrative data sources and statistic producers) and gives them an instrument to detect critical problems and improvement opportunities by themselves. External audits not always are welcome at the beginning, especially if there were no prior contacts to quality issues. Future plans include an external assessment of the quality of administrative registers for statistical purposes to give more objectivity and clarity to the process.

References

Albieri, S., Bianchini, Z., Muiños, R., Segui, F., Torres, N. and Valdivia, I. (2010). *Minimum Set of Standard Quality Indicators to being applied in the MERCOSUR*. Working group 8, Total Quality in Statistics. Project "Cooperation in Statistics CE-MERCOSUR II".

Daas, P., Arends-Tóth, J., Schouten, B., Kuijvenhoven, I. (2008). *Quality Framework for the Evaluation of Administrative Data*. Proceedings of the European Conference on Quality in Official Statistics Q2008. Istat and Eurostat. Rome, Italy.

Daas, P., Ossen, S. and Vis-Visschers R., Arends-T´oth, J. (2009). *Checklist for the Quality Evaluation of Administrative Data Sources*. Discussion paper (09042), Statistics Netherlands. The Hague/Heerlen, Netherlands.

EUROSTAT (2003). *Quality Assessment of Administrative Data for Statistical* Purposes. 6th Meeting, Working Group: "Assessment of quality in statistics". Luxemburg, 2-3 October 2003.

Medina Giopp, A. (2011). *Review of assessment frameworks of statistical quality – Basis for a guideline of quality audits of administrative registers of the State of Yucatan*. World Bank. Washington, DC.

Segui, F. (editor) (2008). *Standard Quality Report*. Quality Management Unit, National Statistical Institute (INE) – Uruguay.

Segui, F. (2009). *Implementing ISO 20252-2006 "Market, opinion and social research" in a statistical office*. Proceedings of the 57th ISI World Statistics Congress, Special Topic Paper. International Statistical Institute. Durban, South Africa.

Segui, F. (2011). *Certifying the Quality & Information Security Management Systems of the NSO according to the international standards ISO 9001, 20252 and 27001*. Proceedings of the 59th ISI-2011 WSC, Special Topic Paper. ISI Dublin, Ireland.

Wallgren, A., Wallgren, B. (2007). *Register-based Statistics: Administrative Data for Statistical Purposes*. John Wiley & Sons, Ltd.. Chichester, England.