Developing a Tool to Assess the Quality of Socio-Demographic Data in Community Health Centres

M. Laberge[1]; A. Shachak[1,2]

[1]Institute of Health Policy, Management and Evaluation, University of Toronto; [2]Faculty of Information, University of Toronto

Keywords

Data quality, assessment, primary care, electronic health records

Summary

Objective: The objectives of this study are to 1) create a quality assessment tool for socio-demographic data aligned with the needs of Community Health Centres (CHCs) and based on the data quality framework of the Canadian Institute for Health Information (CIHI), and 2) test the feasibility of the tool in CHCs.

Methods: The tool was developed based on both theoretical and practical knowledge. A review of the literature was performed to identify data quality frameworks and dimensions that could be employed. In addition, informal discussions with Community Health Centres staff members holding various positions were conducted and a team of subject matter experts was established. This approach supported the alignment between the tool (i.e., the indicators developed, the rating scale, and weighting system) and the setting for which it has been designed. The tool was pilot tested in five CHCs across Ontario.

Results: The decision to focus on socio-demographic data was based on findings from the discussions with staff members. The team established nine principles for the development of the tool, including the use of computer software, whenever possible, to query the data and ensure consistency of the measurement. Data quality scores ranged from 45 to 74 on a scale of 0 (lowest quality) to 100 (highest data quality), with one CHC that was not able to run all of the queries. The feedback from staff was positive and supports the feasibility of the tool as an application of the CIHI data quality framework in a local setting.

Conclusion: Pilot test results demonstrate the feasibility of the tool and an applicability of the CIHI framework as a basis for developing tools for data quality assessment in health care organizations.

Correspondence to:

Maude Laberge

Institute of Health Policy, Management and Evaluation
Faculty of Medicine, University of Toronto
Health Sciences Building, Suite 425
155 College Street
Toronto, ON M5T 3M7
Email: Maude.laberge@mail.utoronto.ca

Appl Clin Inf 2013; 4: 1–11
doi:10.4338/ACI-2012-10-CR-0041
received: October 1, 2012
accepted: December 6, 2012
published: January 9, 2013
Citation: Laberge M, Shachak A. Developing a tool to assess the quality of socio-demographic data in community health centres. Appl Clin Inf 2012; 4:1–11
http://dx.doi.org/10.4338/ACI-2012-10-CR-0041

1. Background

Electronic Medical Records (EMRs) have become an important tool for facilitating the retrieval and use of health care data in clinical practices and health care research [1, 2]. Their use has been associated with higher quality of care and cost-effectiveness [3, 4]. Some also suggest that EMRs may support health care system improvement through physician's effective use of clinical decision support tool embedded in them [5]. However, concerns have been raised about the quality of the data contained in EMRs [6] which limits realization of their potential benefits for information retrieval and analysis, and planning of health services [7]. Although there is a growing interest in data quality, methods to assess it are often limited. Quality is sometimes not defined, and there is lack of consistency in the way it is assessed across studies [8].

Improving data quality requires a theoretical framework on which to build assessments. Although a number of data quality frameworks have been developed [7, 9, 10, 11], many data quality studies still do not employ them and are limited to the comparison of electronic databases to paper charts and measuring the discrepancies between them [12, 13, 14, 15].

Although these studies are important, they do not provide health care organizations with specific, contextualized, tools to assess the quality of their data. As health care organizations are expected to hold high quality data for evidence-based decision making there is a growing need for computerized tools that enable data quality assessment. The present study attempts to fill in this gap by describing an assessment tool that is based on a multidimensional definition of data quality and a data quality framework, developed by the Canadian Institute for Health Information (CIHI) [16]. The data quality assessment tool (DQAT) was developed for a specific dataset that relates to clients' sociodemographic information in Community Health Centres (CHCs) in Ontario, Canada. Its feasibility and face validity were tested in five centres. Results from this pilot application are presented.

An important characteristic of a data quality assessment tool is not only that it determines whether the data are "fit for …actual use" [16] but also that it is based on an accepted definition and framework of data quality. For this study, data quality is defined as "the totality of features and characteristics of a data set, that bear on its ability to satisfy the needs that result from the intended use of the data" [7]. The CIHI Data Quality Assessment Framework was selected as a conceptual basis because of its appropriateness to the specific health care context and because it was developed by the national body responsible for the collection and analysis of health data, and has been accepted and used internationally [17]. The framework identifies five data quality dimensions [16]:

1. Accuracy, which refers to "how well information in or derived from the data holding reflects the reality it was designed to measure" [16];
2. Timeliness, i.e., how current or up-to-date the data are at the time of release;
3. Comparability, which refers to the extent to which databases are consistent over time and the use of standard conventions (such as data elements or reporting periods) or standard classifications or terminologies (that make them comparable to other databases;
4. Usability, which reflects the ease with which a database or data repository may be understood and accessed; and
5. Relevance, which is the degree to which a database meets the current and potential future needs of users.

The framework further divides each of these dimensions into a total of 19 characteristics and 61 criteria [16]. It was developed for large datasets from a number of health sectors and for various purposes. The objectives of this study are to

1. Provide an application of the CIHI framework; and
2. Test its feasibility and applicability to a practical, local, context in Ontario CHCs.

2. Case Report

The project took place in Community Health Centres (CHCs). CHCs are health care organizations which provide "primary health and health promotion programs for individuals, families and communities" [18]. They are non-profit and grass-root organizations established by the local community

and governed by a board of directors whose members are representative of that community [18]. Each CHC is unique in that it serves very specific and defined client populations that experience barriers to accessing primary care. The CHC model of care is centred around the needs of the target populations with a focus on health equity. To assess the extent to which CHCs meet the needs of the populations they serve, CHCs collect a defined set of data elements on their patients which includes gender, date of birth, postal code, preferred language of service, highest level of education, country of origin, ethnicity, religion, household composition, and the total household income level.

The development of the DQAT involved a number of steps including: collecting CHC information, establishing a team, developing the DQAT, and pilot-testing it.

Information from the sector was collected through group and one-on-one interviews at thirteen CHCs with executive directors, management staff, and data management coordinators (DMCs).

One of the aims was to ground the development of the DQAT in a team and community-based approach. A team of four CHC employees (three DMCs and one Regional Decision Support Specialist) and the Manager of Performance Management from the Association of Ontario Health Centres was established, with members selected based on years of employment and skills. Team members provided their expertise in the development of the tool and its fitness to the data and processes in CHCs.

The team reviewed the nineteen characteristics of the CIHI data quality framework and selected twelve based on applicability and measurability in the CHC context: coverage, capture and collection, completeness, measurement error, edit and imputation, processing and estimation, data currency at the time of release, documentation currency, standardization, historical comparability, accessibility, documentation, and value (▶Figure 1). Indicators for each of these characteristics were identified to assess each aspect of data quality. The team developed, reviewed and agreed on definitions and queries which would run on HummingbirdTM BI Query, the decision support software used by CHCs. The queries were also tested and reviewed by two other DMCs, who were not involved in the initial development process, to ensure they were correct and appropriately measuring the indicator. Some of the indicators were self-assessed questions using a Likert scale. These indicators were replicated into a separate questionnaire that could be accessed and answered on SurveyMonkey®. ▶Table 1 provides a summary of the dimensions, characteristics and indicators of the DQAT. In the next stage, each indicator was assigned a weight. The team selected two components – difficulty and impact – on which to weight each indicator. Difficulty is related to the work involved in achieving a high score on the indicator and impact is related to how achieving a better score on the indicator would affect overall data quality.

"The values for impact and difficulty range from one (lowest) to three (highest). These values were then multiplied to give the overall weight of the indicator (ranging from one to nine). Taking the first indicator – "Percentage of clients with only one socio-demographic encounter" – as an example of the decision making process, the problem identified was that, for patients who already have a socio-demographic profile in the EHR system, some health care providers might accidentally create a new socio-demographic encounter instead of updating the original one. The team considered the impact of this problem was average (2). However, the team considered that educating staff and having a protocol was quite simple. Hence, after discussing the indicator, a consensus was reached to assign a value of one for difficulty and two for impact. Multiplying these values gives the overall weight of two for this indicator.

The DQAT was developed in Microsoft® Excel 2007. This software was chosen for ease of entering weights and formulas for score calculation, which were built into the tool, as well as for its availability in all CHCs, which would facilitate its use. Thus, the DQAT is an excel spreadsheet to be populated with the results obtained from running the queries on HummingbirdTM BI Query and completing the questionnaire on SurveyMonkey®. Within the document, there is a separate tab with definitions of the indicators for reference.

Once the tool was completed, it was tested in five CHCs, who were provided with an electronic package. This package included a set of queries to run on the CHC database, a link to the SurveyMonkey® questionnaire, and the excel document with the DQAT to be populated.

3. Results

3.1 Create a quality assessment tool for socio-demographic data based on the needs of CHCs and on the CIHI data quality framework

CHC Executives and managers declared using the data collected to various degrees but in general, not to the extent that they would have liked. There was a general perception that data was of poor quality and particularly the socio-demographic data. This led the team to the decision of developing a tool to assess specifically the quality of socio-demographic data.

3.2 Test the feasibility of the tool in CHCs.

All five CHCs that were contacted to test the tool consented to participate and returned the completed DQAT, along with the reports from running the queries. In addition, they all completed the questionnaire online that could be accessed by the team for cross-checking the corresponding original answers to what was entered in the tool. The results of the data quality assessment, obtained by using the DQAT, for each CHC are shown in ▶Table 2. Final data quality scores varied from 45% (CHC 5) to 74% (CHC 1). For one of the CHCs, a final score could not be calculated due to a connectivity problem with the local databases. Only partial scores for this CHC are shown in ▶Table 2 (CHC 5).

4. Discussion

Socio-demographic information has rarely been studied and most previous data quality assessment research focused on diagnostic data [19]. As health care organizations are increasingly engaging in performance measurement, high quality data become necessary [20]. The tool described here provides an application of the CIHI data quality framework that is feasible and usable by health care organizations.

The DQAT integrates concepts for supporting data quality through indicators related to the use of the data, and the understanding of their importance by individuals involved in the collection and entry of data. This is aligned with findings suggesting that people perform their tasks more thoroughly when they understand their value to other members of the organization [21].

There has been much emphasis on data accuracy in the literature [6, 22], and this is consistent with the importance of the accuracy dimension in the CIHI framework and the assessment tool developed. The traditional data quality assessments are quantifications of "source-to-database error rate" [23]. In contrast, the DQAT provides a score on each of the quality dimensions and on the overall quality, which can be considered as a more complete assessment.

Data quality scores obtained from five CHCs ranged from 45% to 74%, suggesting that the DQAT is sensitive enough to detect differences between CHCs with high and low data quality, but perhaps not between CHCs with relatively similar scores. Face validity is supported by the fact that the tool was developed by a team of experts and having CHCs review the results and the measurements (queries and questions) for validation.

The motivation for the development of the DQAT was to improve the quality of socio-demographic data in CHCs. Although the initial testing of the DQAT raised participants' awareness of, and interest in data quality, further research is required to determine whether it will be successfully implemented and whether improvements in data quality can actually be measured and achieved by using it in CHCs. Study results were shared with all Ontario CHCs and a decision was made to apply the DQAT to the province. The expansion of the project could support its generalizability to the CHC sector and lead to improvements and refinements in the tool such as full computerization, and inclusion of further measures beyond socio-demographic data. Some aspects of data quality, such as error analysis, may still not be measureable through either computerized queries or questions. The interest in data quality generated by this study could potentially lead to improvements in overall EHR data.

However, since the DQAT was designed for the specific context of CHCs, its generalizability to other settings would require further development and adaptation of the indicators.

Finally, the development of this tool represented a challenge for the team members, as well as a learning process. Difficulties throughout the process included the identification of indicators for the characteristics and accepting that not all characteristics could be measured. Writing queries on the HummingbirdTM software resulted in enhanced skills of using the software, not only in relationship to the DQAT but also for other decision support activities. Another difficulty was in weighting the measures and deciding on an appropriate approach to do so. DMC team members tested the tool on their own databases (not reported in this study) and there was a sense that although quantification of data quality may remain an imperfect measure, it was still very informative and the results were aligned with their expectations.

5. Conclusion

This DQAT is a first application of the CIHI framework to a local database. The pilot test results demonstrate the feasibility of the tool and the applicability of the CIHI framework as a basis for developing local instruments for data quality assessment. Future research should be conducted to evaluate the effect of using the tool on data quality and decision making and planning processes, as well as its application to other settings.

Clinical Relevance

Health organization managers and providers need high quality data to make informed decisions in the planning of patients' care. Having a tool to assess data quality empowers health care organizations in using the data adequately to improve care and health services to patients.

Conflict of Interest

The authors declare that they have no conflict of interest in the research.

Human Subjects Protection

The project received ethics approval from the University of Toronto Research Ethics Board. All Executive Directors of participating CHCs agreed to have the results published. All individuals involved agreed to have their names published in the acknowledgements section.

Funding

There was no funding dedicated to the project. The tool was based on software tools that were already available in CHCs. Team members' hours were part of their respective roles.

Acknowledgments

This project received the support of Association of Ontario Health Centres (AOHC) and of the Ontario Community Health Centres Performance Management Committee (CHC-PMC). The authors also wish to thank the five CHCs who agreed to test the tool. The following individuals contributed to the design and development of the tool: Jeremy Irving, Anjali Misra, Christine Randle, and Jie Zhang. Nik Papanikolas contributed to the review and validation of the software-based queries.

Table 1 Dimensions, characteristics and indicators of the data quality assessment tool (DQAT).

Dimension	Characteristics	Indicators	Impact	Difficulty	Weight
Accuracy	Coverage	Percentage of clients with only one socio-demographic encounter;	2	1	2
		Percentage of clients with only one language;	1	2	2
		Percentage of clients with at least one encounter in three years.	2	2	4
	Capture and collection	There is a clear process including roles and responsibilities to collect data;*	3	2	6
		Staff are clear about what are the mandatory and required data fields that they should collect;*	3	1	3
		The intake forms(s) have all of the required data fields.*	3	1	3
	Completeness	Percentage of clients with complete socio-demographic information;	3	3	9
		Percentage of clients with ethnicity or religion data**;	1	2	2
		Percentage of clients with ethnicity data;	2	3	6
		Percentage of clients with religion data;	1	3	3
		Percentage of clients with country of origin data;	1	2	2
		Percentage of clients with education data;	1	1	1
		Percentage of clients with household income data;	2	3	6
		Percentage of client with number of people supported by household income data;	1	3	3
		Percentage of clients with current household composition data.	1	2	2
	Measurement Error	Percentage of clients with a valid age;	3	1	3
		Percentage of clients with a valid date of arrival in Canada;	2	1	2
	Edit and Imputation	Variance in number of clients in Live versus the Local Management Information System (LMIS)	3	3	9
	Processing and Estimation	Percentage of synchronizations with the Ministry of Health and Long Term Care (MOHLTC) performed successfully and as required.*	3	1	3

Table 1 Continued

Dimension	Characteristics	Indicators	Impact	Difficulty	Weight
Timeliness	Data currency at the time of release	Percentage of socio-demographic (SD) encounters entered within 2 weeks of the client's registration date.	1	2	2
		Percentage of extraction performed as guided (weekly).*	2	2	4
	Documentation currency	Percentage of clients who had their socio-demographic data updated in the last three years.	3	3	9
Comparability	Standardization	Staff collecting the data know the definition of all socio-demographic data fields as defined in the provincial standards.*	3	3	9
	Historical comparability	Changes in collecting or entering socio-demographic data are documented (e.g. version of new form designed and date that it was implemented).*	3	2	6
		Changes to queries associated with socio-demographic data are documented.*	3	2	6
Usability	Accessibility	There is documentation available on the data to be entered.*	3	2	6
	Documentation	Reports are generated regularly.*	2	2	4
Relevance	Value	Reports are requested on the socio-demographic data.*	3	3	9

*Data collected through Survey Monkey®; **The standards suggest that CHCs should collect either ethnicity or religion of clients and hence the indicator is meant to measure to which extent the standard is being met. However, it is considered better practice to request both ethnicity and religion to clients, which is the reason why there are also indicators for both these data elements separately.

Table 2 Results from the five CHCs testing the data quality assessment tool.

Note: Rates are in per cent	CHC 1		CHC 2		CHC 3		CHC 4		CHC 5	
	Rate	Score	Rate	Score	Rate	Score	Rate	Score	Rate	Score
Dimension: ACCURACY										
Percentage of clients with only 1 socio-demographic encounter	98.7	1.97	95.0	1.90	94.9	1.90		0.00	90.7	1.81
Percentage of client with only 1 language	99.2	1.98	99.0	1.98	97.0	1.94	86.6	1.73	93.8	1.88
Percentage of clients with at least one ISE or group attendance in 3 years	75.0	3.00	60.0	2.40	72.0	2.88	67.0	2.68	55.0	2.20
There is a clear process (including roles and responsibilities) to collect the clients' socio-demographic data	100	6.00	100	6.00	67.0	4.02	100	6.00	25.0	1.50
Staff know the mandatory and required socio-demographic data fields to collect.	100	3.00	75.0	2.25	100	3.00	100	3.00	50.0	1.50
The intake forms(s) have all of the required data fields	100	3.00	100	3.00	100	3.00	100	3.00	80.0	2.40
Percentage of clients with complete socio-demographic information										
Percentage of clients with ethnicity or religion data	75.2	1.50	82.4	1.65	88.2	1.76	60.7	1.21	20.4	0.41
Percentage of clients with ethnicity data	72.8	4.37	79.6	4.77	88.1	5.29	56.7	3.40	20.4	1.22
Percentage of clients with religion data	39.4	1.18	70.2	2.11	0.48	0.01	53.5	1.60	0.56	0.02
Percentage of clients with country of origin data	71.0	1.42	70.8	1.42	90.2	1.80	41.2	0.82	34.1	0.68
Percentage of clients with education data	73.0	0.73	81.8	0.82	83.0	0.83	54.2	0.54	33.7	0.34
Percentage of clients with household income data	69.0	4.14	76.0	4.56	79.0	4.74	20.0	1.20	31.0	1.86
Percentage of client with number of people supported by this income data	18.1	0.54	18.6	0.56	64.8	1.94	1.44	0.04	12.6	0.38
Percentage of clients with current household composition data	76.4	1.53	62.8	1.26	89.4	1.79	29.3	0.59	28.7	0.57
Percentage of clients with a valid age	100	3.00	99.9	3.00	99.8	2.99		0.00	98.9	2.97
Percentage of clients with a valid date of arrival in Canada	100	2.00	100	2.00	98.7	1.97		0.00	98.8	1.98
Percentage similarity between number of clients in Live and in LMIS	76.2	6.86	90.5	8.15	97.3	8.76		0.00	92.6	8.34
Last quarter synchronization was performed	100	3.00	100	3.00	100	3.00	100	3.00	100	3.00
TOTAL – ACCURACY	**48.79**		**52.98**		**51.66**		**28.88**		**33.05**	

Table 2 Continued

Note: Rates are in per cent	CHC 1		CHC 2		CHC 3		CHC 4		CHC 5	
	Rate	Score	Rate	Score	Rate	Score	Rate	Score	Rate	Score
Dimension: TIMELINESS										
Percentage of socio-demographic encounters entered within 2 weeks of the clients registration date	8.5	0.17	80.4	1.61	25.2	0.50		0.00	24.1	0.48
Data extractions are performed weekly	100	4.00	75.0	3.00	75.0	3.00	100	4.00	25.0	1.00
Percentage of clients who had their socio-demo data updated in the last three years	74.6	6.71	61.5	5.53	38.6	3.48		0.00	60.8	5.47
TOTAL – TIMELINESS	**10.88**		**10.14**		**6.98**		**4.00**		**6.96**	
Dimension: COMPARABILITY										
Staff collecting the data know the definition of all socio-demographic data fields as defined in the provincial standards	100	9.00	75.0	6.75	75.0	6.75	75.0	6.75	50.0	4.50
Changes in collecting or entering socio-demographic data are documented	75.0	4.50	75.0	4.50	50.0	3.00	100	6.00	50.0	3.00
Changes to queries associated with socio-demographic data are documented	75.0	4.50	75.0	4.50	50.0	3.00		0.00	25.0	1.50
TOTAL – Comparability	**18.00**		**15.75**		**12.75**		**12.75**		**9.00**	
Dimension: USABILITY										
Documentation is available for staff to collect and enter socio-demographic data	100	6.00	75.0	4.50	75.0	4.50	100	6.00	25.0	1.50
Reports that include socio-demographic data are generated and used	50.0	2.00	75.0	3.00	50.0	2.00	50.0	2.00	50.0	2.00
TOTAL USABILITY	**8.00**		**7.50**		**6.50**		**8.00**		**3.50**	
Dimensions: RELEVANCE										
Other staff at the centre request reports on socio-demographic data	50.0	4.50	75.0	6.75	50.0	4.50	50.0	4.50	50.0	4.50
TOTAL – RELEVANCE	**4.50**		**6.75**		**4.50**		**4.50**		**4.50**	
	CHC 1		**CHC 2**		**CHC 3**		**CHC 4**		**CHC 5**	
Total	**90.2**		**93.12**		**82.39**		**58.13**		**57.00**	
Final Score	**72%**		**74%**		**65%**		**46%**		**45%**	

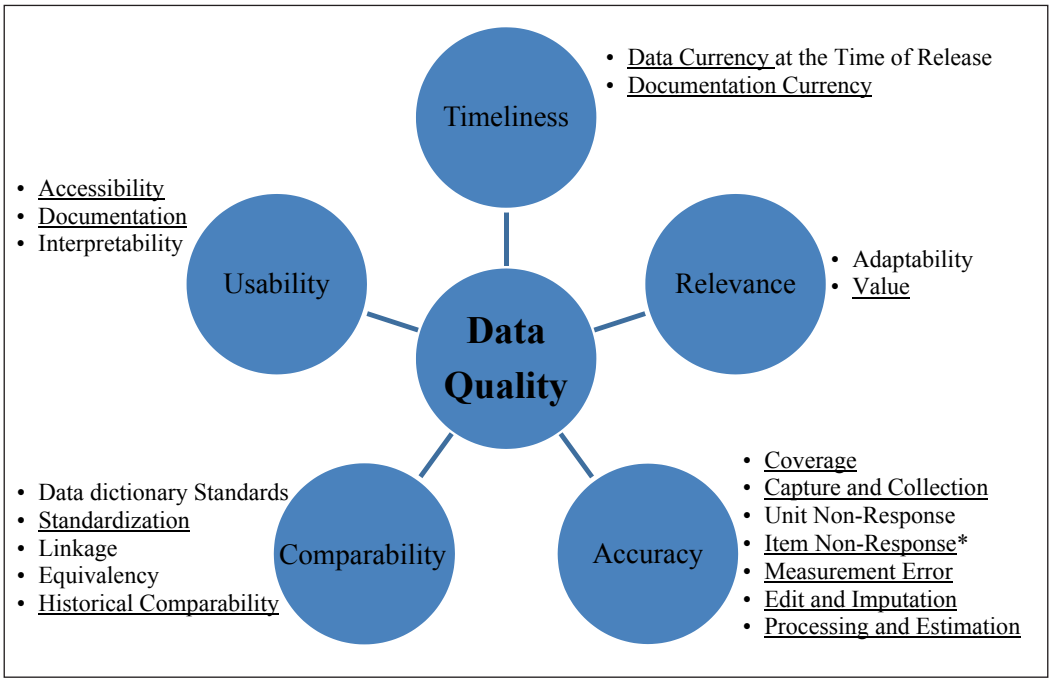

Fig. 1 CIHI data quality framework with five dimensions and 19 characteristics. Characteristics from the CIHI framework used for the DQAT are underlined. *The characteristic "Item Non-Response" was named "completeness in the DQAT to facilitate comprehension from users.

References

1. Dick RS, Steen EB. The computer-based patient record: An essential technology for healthcare. Revised Edition, National Academy Press, Washington, DC, 1997.
2. Hersh WR. Information retrieval: A health care perspective. Springer, New York, 1996.
3. Adams WG, Mann AM, Bauchner H. Use of electronic medical record improves the quality of urban pediatric primary care. Pediatrics 2003; 111: 626-632.
4. Chaudhry B, et al. Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care. Annals of Internal Medicine 2006; 144: 742-752.
5. Blumenthal D. Stimulating the adoption of health information technology. N Engl J Med 2009; 360: 1477-1479.
6. Hogan WR, Wagner MM. Accuracy of data in computer based patient records. J Am Med Inform Assoc 1997; 4: 342-355.
7. Arts DGT, De Keizer NF, Sheffer GJ. Defining and improving data quality in medical registries: A literature review, case study, and generic framework. J Am Med Inform Assoc 2002; 9: 600-611.
8. De Araujp Lima CR, et al. Saúde Pública. Review of data quality dimensions and applied methods in the evaluation of health information systems [Revisão das dimensões de qualidade dos dados e métodos aplicados na avaliação dos sistemas de informação em saúde]. Rio de Janeiro 2009; 25: 2095-2109.
9. Carson CS. Toward a framework for assessing data quality. IMF Working Paper, 2001;01/25.
10. Strong DM, Lee YW, Wang RY. Data quality in context, communication of the ACM. 1997; 40: 5.
11. Wang YW, Strong DM. Beyond accuracy: what data quality means to data consumers. Journal of Management Information Systems 1996; 12: 4.
12. Gimbel S, et al. An assessment of routine primary care health information system data quality in Sofala Province, Mozambique. Population Health metrics 2011; 9: 12.
13. Nielsen GL, et al. Analyses of data quality in registries concerning diabetes mellitus – a comparison between a population based hospital discharge and an insulin prescription registry. Journal of Medical Systems 1996; 20.
14. Scobie S, Basnett I, McCartney P. Can general practice data be used for needs assessment and health care planning in an inner-London district? Journal of Public Health Medicine 1995; 17: 475-483.
15. Armenti KR, et al. Improving the quality of industry and occupation data at a central cancer registry. American Journal of Industrial Medicine 2010; 53: 995-1001.
16. Canadian Institute for Health Information. The CIHI data quality assessment framework, 2005.
17. Kerr AK, Norris T, Stockdale R. The strategic management of data quality in healthcare. Health Informatics Journal 2008; 14: 259-266.
18. Ontario Ministry of Health and Long Term Care. http://www.health.gov.on.ca/english/public/contact/chc/chc_mn.html, consulted August 22nd, 2011.
19. Thiru K, Hassey A, Sullivan F. Systematic review of scope and quality of electronic patient record data in primary care. BMJ 2003; 326: 1070.
20. Schneider EC, et al. Enhancing performance measurement NCQA's road map for a health information framework. JAMA 1999; 282: 1184-1190.
21. Hammond WE, et al. Connecting information to improve health. Health Affairs 2010; 2: 284-288.
22. Pringle M, Ward P, Chilvers C. Assessment of the completeness and accuracy of computer medical records in four practices committed to recording data on computer. British Journal of General Practice 1995; 45: 537-541.
23. Nahm ML, CF Pieper, MM Cunningham. Quantifying data quality for clinical trials using electronic data capture. PLoS ONE 2008; 3: e3049.

