Investing in E-Health: People, Knowledge and Technology for a Healthy Future
H. Grain et al. (Eds.)

doi:10.3233/978-1-61499-427-5-13

Methods for assessing the quality of data in public health information systems: A critical review

Hong CHEN [a, b], Ping YU [a,1], David HAILEY [a] and Ning WANG [c]
[a] *School of Information Systems and Technology, Faculty of Engineering and Information Sciences, University of Wollongong, Australia*
[b] *Jiangxi Provincial Centre for Disease Control and Prevention, China*
[c] *NCAIDS, China CDC, Beijing, China*

Abstract. The quality of data in public health information systems can be ensured by effective data quality assessment. In order to conduct effective data quality assessment, measurable data attributes have to be precisely defined. Then reliable and valid measurement methods for data attributes have to be used to measure each attribute. We conducted a systematic review of data quality assessment methods for public health using major databases and well-known institutional websites. 35 studies were eligible for inclusion in the study. A total of 49 attributes of data quality were identified from the literature. Completeness, accuracy and timeliness were the three most frequently assessed attributes of data quality. Most studies directly examined data values. This is complemented by exploring either data users' perception or documentation quality. However, there are limitations of current data quality assessment methods: a lack of consensus on attributes measured; inconsistent definition of the data quality attributes; a lack of mixed methods for assessing data quality; and inadequate attention to reliability and validity. Removal of these limitations is an opportunity for further improvement.

Keywords. Data quality, assessment, public health, information systems

Introduction

High-quality data and effective data quality assessment are vital for accurately measuring public health outcomes [1]. Public health information systems (PHIS), whether paper-based or electronic, are the repositories of public health data [2]. As the data stored in PHIS are frequently used to monitor and evaluate public health outcomes, the need for effective data quality assessment for PHIS has never been greater [3].

Data quality assessment is primarily based on the measurement theory which links data attributes to concrete measures [4]. Data quality attributes need to be precisely defined and representative to certain aspects or constructs of data quality. The corresponding measures are used to objectively or empirically examine the attributes [4]. Then effective data quality assessment methods are used to operationalise the attributes into measurable items which function appropriately and can be scored precisely [4].

[1] Corresponding Author: Ping YU, E-mail: ping@uow.edu.au (P.Y.)

Meanwhile, data quality assessment methods must be of adequate reliability and validity. Reliability refers to the degree to which a variable has consistent values when being measured several times and is free from random error. Measurement reliability can be determined by test-retest consistency, internal consistency and inter-and-intra-observer consistency [5]. A major indicator of reliability is consistency among repeated measures. We consider a measurement as reliable if the methods and procedures of data quality assessment are well defined with regard to repeatability. Validity is the degree to which a value obtained actually represents what it is supposed to represent and is similar to the true value. Measurement validity consists of internal validity (face/content validity, criterion-related validity/convergence, predictive validity) and external validity (generalisability) [5]. The most basic and first level of validity of a measurement is face/content validity.

Although methods for data quality assessment in PHIS have been reported, there is a lack of systematic reviews on which methods have been used in which context. Therefore, this systematic review aims to critically evaluate data quality assessment methods. We focus on which attributes were assessed, how the assessment was conducted, and the reliability and validity of each method. We believe this study is useful for the development of data quality assessment methods in public health.

1. Methods

A systematic search was performed between June 2012 and October 2013 by the first author in 7 databases like Scopus, IEEE Xplore and PubMed. The key words used for search included “data quality”, “information quality”, “public health”, “population health”, “information system*”, “assess*”, and “evaluat*”. They were searched individually or in combination. The inclusion criteria were peer-refereed empirical studies or institutional reports of data quality assessment for public health or PHIS during the period of 2001–2013. The exclusion criteria were narrative reviews, correspondence and commentaries in the topic area. Searches of well-known institutional websites, such as the World Health Organisation and Centres for Disease Control and Prevention in the USA, were also performed as they have leading roles in international or national data quality assessment efforts. Finally, 35 papers were included in our review [2,6-39], of which seven were published by institutions.

2. Results

2.1. Attributes of Data Quality

A total of 49 attributes were used in the studies to describe data quality. Completeness was the most-used attribute of data quality and measured in 24 studies [2,6,8-10,12,14-26,30-32,35,39]. This was followed by accuracy, assessed in 21 papers [2,8,10,15-22,25-27,29,32,34-37,39]. The third most-used attribute, timeliness, was measured in 9 studies [6-9,16,20,23,32,36]. The other attributes, such as relevance and validity, were assessed less than 3 times in the included methods.

Thirteen studies adopted the attributes that were previously developed [12-14,20,21,23-25,27,30,35,37,39]. The other 22 studies did not explain how the attributes were selected and why. The quality of data was frequently assessed in a positive manner by 38 desirable data quality attributes compared to 11 undesirable ones (e.g., "completeness" versus "missing data", "accuracy" versus "data errors").

The definition of the attributes was sometimes inconsistent. The same attribute could be given different meaning by different researchers. For example, in 16 studies, the attribute of completeness was defined as "standard data element entry" in the reporting forms [2,6,9,10,15,16,18,19,22,23,25,26,32,35,39]. Whereas the other studies defined completeness as either "the correctness of data collection methods" or "the proportion of facilities reporting in an administrative area" [12,16,17,21,24,30]. Also, different attributes could be given the same meaning. The attribute of accuracy was stated as "no missing data elements" in the reports, which was similar to the definition of completeness in other studies [19]. Whereas when accuracy was defined as "validity" in studies, it was similar to the definition of measurement validity [17].

2.2. Methods for Data Quality Assessment

The methods for data quality assessment are categorised into two types according to the type of study subjects. One is objective assessment, which examines the data values in PHIS. The other is subjective assessment, which explores perceptions of the users or stakeholders who use or manage the data, or assesses the quality of documentation for data management. Twenty-two studies only conducted objective assessment [2,6,8,10,11,13,18-20,22-24,26-31,33,34,37,38]. Four studies only conducted subjective assessment [7,14,36,39]. Nine used both [9,12,15-17, 21, 25,32,35].

Four approaches were used for objective assessment. The first was off-site desk review of the datasets [2,6,8,15,16,18-20,22-24,26,27,30,31,37,38]. The second was on-site follow-up assessment via field tracing and verification of the reported data [8,16,17,21,25,37]. The third was to conduct a cross-sectional survey or record-linkage to cross-check and testify the quality of PHIS [10,11,29,33,34]. The fourth presented the results without explicit information [9,12, 20,32].

Subjective assessment was often accomplished by two approaches. The first is open, semi-structured or structured interviews with data users or stakeholders. This was used in 12 studies [7,9,12,14-17,21,25,35,36,39]. The second, used in five studies, is reviewing documentation/guideline/instruction [16,17,21,32,35].

2.3. Sampling Methods for Data Quality Assessment

Of the 31 studies that conducted objective assessment, 19 used non-probability sampling methods [6,8,10,11,13,16-20,22,23,25,26,28,29,31,35,38]. Eleven used probability sampling methods, of which eight used multi-stage sampling [8,13,15,17,21,33,34,37] and three used systematic sampling [24,27,30]. The remaining four studies did not give the information about the sampling methods [2,7,9,12]. The sampling period of assessed datasets ranged from one month to 23 years and the sample size ranged from 72 entries of a reference syndrome to 7.5 million laboratory data items [11,24,28]. One study calculated the sample size and statistical power [34]. Another study predetermined the sample size but did not present a power calculation [27].

Of the 13 subjective assessment studies, three did not provide the sampling methods [16,17,36]. The remainder used purposive sampling to select interviewees. One study surveyed all users due to the small size of the population [14].

2.4. Measures of Data Quality Attributes

A common practice of measuring data quality attributes in the reviewed objective assessment studies was to determine the quality of a set of purposively selected data items. These data items served as measurement items for tracing and indicating the quality of data attributes and surrogates of data quality. The criteria used to select data items appeared to be relevance of data or practical considerations. Data items used included population demographics, such as age, gender, and birth date, or specific data such as laboratory testing results and disease code. The number of data items assessed varied from 1 to 30. Ratio (such as percentage, rate) was frequently used to measure each data quality attribute [2,6,8-13,15-35,37,38].

Seven studies reported the constructs of the questionnaire or checklist used in the subjective assessment studies [7,9,12,14,16,17,36]. They used Likert-type scales to measure data quality attributes. The measurement items were often grouped in themes in reporting the qualitative results [15,21,25,32,35].

2.5. Reliability and Validity of Methods for Data Quality Assessment

Of the 31 studies that used objective assessment methods, only two explicitly reported the measures of reliability. One used kappa statistics and the other used a Cronbach's alpha [28,34]. The measurement instruments reported in the seven institutional publications were applied in different public health contexts [2,6-8,16,17,36], suggesting the approval by the public health community of their reliability. Twelve studies used the existent data quality assessment instruments to measure data quality but did not report the process of measuring reliability [12-14,20,21,23-25,30,35,37,39]. However, 9 of these 12 studies did provide detailed description of the measurement process, which gave us confidence about the reliability of these studies.

Only 2 out of 13 articles that used the subjective assessment method gave information about reliability [14,25].

The validity of the reviewed methods was rarely reported. Only 4 out of 35 studies reported internal validity [14,25,28,33], 3 of which focused on the face/content validity [14,25,33]. The fourth conducted a principal-component analysis of measurement items [28]. Other types of validity were not found in the studies and the generalisability of the results was not clear.

3. Discussion

While few data quality assessment methods are infallible, a data quality assessment method in PHIS cannot be effective unless the measurement results represent and reflect the real world of public health practice. In order to conduct effective data quality assessment, measurable data attributes have to be precisely defined. Then reliable and valid measurement methods for data attributes must be used to measure each attribute.

In reviewing the literature, five themes emerged from the results. The first is a lack of consensus on which attributes should be measured for data quality in PHIS. This might be due to the lack of the criteria for selection of data attributes and measures. The second theme is a lack of clear, consistent definition of data attributes and measures in the included studies. This lack of definition clarity and consistency of the attributes and the corresponding measures could lead to confusion in interpretation and application of assessment results [4]. More importantly, this could reduce the generalisability of the results of the reviewed studies. The third theme is a lack of systematic procedures to measure data quality. Mixed-methods with both objective and subjective assessments are considered useful for assessing data quality [4]. The possible reasons for few studies adhering to this principle might be that public health practitioners were unaware of this approach or had not drawn attention to it. The fourth theme is little field verification of the quality of data. Therefore, the results of data quality assessment were not validated by a direct examination of the accuracy of data values in the relevant PHIS. A lack of assurance that data refer to the real world situation may undermine end-users' confidence in the data quality [27]. Last but not least, the reliability of the assessment methods was limited, and their validity was rarely assessed. Use of non-probability sampling methods might have been a sub-optimal choice for their representativeness and external validity. Therefore we assert that the importance of reliability and validity of data quality assessment methods is still a neglected area. We suggest that systematic, scientific data quality assessment needs to be conducted in the context of public health.

4. Conclusions

In this paper, we have systematically reviewed the extant literature on data quality assessment in PHIS and identified the major attributes and measurement methods. The results of this review demonstrate there is plenty of room for further development of effective data quality assessment methods in the context of public health. These include: improving the definition clarity in data quality attributes and associated measures, triangulation of mixed methods for assessment of data quality, and giving due attention to the reliability and validity of data quality assessment methods. Removal of these limitations is an opportunity for further improvement.

References

[1] Institute of Medicine, Future of Public Health. National Academies Press, Washington, DC, USA, 1988.
[2] CDC, Updated guidelines for evaluating surveillance systems: Recommendations from the guidelines working group. *Morbidity and Mortality Weekly Report*, **50** (2001), 1-35.
[3] M. Chan, et al., Meeting the demand for results and accountability: A call for action on health data from eight global health agencies. *PLoS Med*, **7** (2010), e1000223.
[4] L. Pipino, et al., Developing Measurement Scales for Data-quality Dimensions, in *Information Quality*, R.Y. Wang, E.M. Pierce, and S.E. Madnick, Editors. M.E. Sharpe, Inc., Armonk, NY, USA, 2005.
[5] A. Alkon, Understanding Research: An Overview for Health Professionals. 2014, www.coursera.org
[6] World Health Organization, *Guide to the Health Facility Data Quality Report Card*. 2013, World Health Organization: Geneva, Switzerland.
[7] World Health Organization, *Assessing the National Health Information System: An Assessment Tool Version 4.00.* World Health Organization, Geneva, Switzerland, 2008.
[8] World Health Organization, *The Immunization Data Quality Audit (DQA) Procedure*. World Health Organization, Geneva, Switzerland, 2003.
[9] E. Venkatarao, et al., Monitoring data quality in syndromic surveillance: Learnings from a resource limited setting. *Journal of Global Infectious Diseases*, **4** (2012), 120-127.
[10] N.A.H. Van Hest, et al., Record-linkage and capture-recapture analysis to estimate the incidence and completeness of reporting of tuberculosis in England 1999-2002. *Epidemiology and Infection*, **136** (2008), 1606-1616.
[11] L. Savas, et al., Mortality ascertainment of women veterans: A comparison of sources of vital status information, 1979-2002. *Med Care.*, **47** (2009), 125-8.

[12] K.M.I. Saeed, R. Bano, and R.J. Asghar, Evaluation of the national tuberculosis surveillance system in Afghanistan. *Eastern Mediterranean Health Journal*, **19** (2013), 200-207.
[13] O. Ronveaux, et al., The immunization data quality audit: Verifying the quality and consistency of immunization monitoring systems. *Bulletin of the World Health Organization*, **83** (2005), 503-10.
[14] S. Petter and A. Fruhling, Evaluating the success of an emergency response medical information system. *International Journal of Medical Informatics*, **80** (2011), 480-489.
[15] J.A. Pereira, et al., Exploring the feasibility of integrating barcode scanning technology into vaccine inventory recording in seasonal influenza vaccination clinics. *Vaccine*, **30** (2012), 794-802.
[16] MEASURE Evaluation. PRISM: Performance of Routine Information System Management. 2010; Available online: www.cpc.unc.edu/measure/tools/monitoring-evaluation-systems/prism.
[17] MEASURE Evaluation. Data Quality Audit Tool Guidelines for Implementation. 2008; Available online: www.cpc.unc.edu/measure/publications/ms-08-29?searchterm=data+quality+audit.
[18] K.S. Mate, et al., Challenges for routine health system data management in a large public programme to prevent mother-to-child HIV transmission in South Africa. *PLoS ONE*, **4** (2009), e5483.
[19] S.D. Makombe, et al., Assessing the quality of data aggregated by antiretroviral treatment clinics in Malawi. *Bulletin of the World Health Organization*, **86** (2008), 310–314.
[20] F. Litow and P. Krahl, Public Health Potential of a Disability Tracking System: Analysis of U.S. Navy and Marine Corps Physical Evaluation Boards 2005-2006. *Military Medicine*, **172** (2007), 1270-1274.
[21] W. Lin, et al., Is the HIV sentinel surveillance system adequate in China? Findings from an evaluation of the national HIV sentinel surveillance system. *Western Pacific Surveillance and Response Journal*, **3** (2012), 78-85.
[22] R.R. Lash, et al., Effects of georeferencing effort on mapping monkeypox case distributions and transmission risk. *International Journal of Health Geographics*, **11** (2012), 23.
[23] R.A. Hills, et al. Timeliness and data element completeness of immunization data in Washington State in 2010: A comparison of data exchange methods. *AMIA ... Annual Symposium Proceedings (1942-597X*), 2012, 340-349.
[24] S.L. Harper, et al., Improving Aboriginal health data capture: evidence from a health registry evaluation. *Epidemiology and Infection*, **139** (2011), 1774-1783.
[25] D. Hahn, P. Wanjala, and M. Marx, Where is information quality lost at clinical level? A mixed-method study on information systems and data quality in three urban Kenyan ANC clinics. *Global Health Action*, **6** (2013), 1-10.
[26] B.G. Frizzelle, et al., The importance of accurate road data for spatial applications in public health: Customizing a road network. *International Journal of Health Geographics*, **8** (2009), 24.
[27] D. Freestone, D. Williamson, and D. Wollersheim, Geocoding coronial data: Tools and techniques to improve data quality. *Health Information Management Journal*, **41** (2012), 4-12.
[28] M. Forster, et al., Electronic medical record systems, data quality and loss to follow-up: Survey of antiretroviral therapy programmes in resource-limited settings. *Bulletin of the World Health Organization*, **86** (2008), 939-947.
[29] J. Ford, et al., Using hospital discharge data for determining neonatal morbidity and mortality: a validation study. *BMC Health Serv Res*., **20** (2007), 188.
[30] M. Edmond, C. Wong, and S.K. Chuang, Evaluation of sentinel surveillance system for monitoring hand, foot and mouth disease in Hong Kong. *Public Health*, **125** (2011), 777-783.
[31] B.E. Dixon, J.J. McGowan, and S.J. Grannis. Electronic laboratory data quality and the value of a health information exchange to support public health reporting processes. *AMIA ... Annual Symposium Proceedings (1942-597X)*, 2011, 322-330.
[32] L. Dai, et al., Birth defects surveillance in China. *World Journal of Pediatrics*, **7** (2011), 302-310.
[33] M. Corriols, et al., The Nicaraguan Pesticide Poisoning Register: Constant Underreporting. *International Journal of Health Services*, **38** (2008), 773-787.
[34] H.B. Clayton, et al., The Florida investigation of primary late preterm and cesarean delivery: The accuracy of the birth certificate and hospital discharge records. *Maternal and Child Health Journal*, **17** (2013), 869-878.
[35] Y. Chiba, M.A. Oguttu, and T. Nakayama, Quantitative and qualitative verification of data quality in the childbirth registers of two rural district hospitals in Western Kenya. *Midwifery*, **28** (2012), 329-339.
[36] Canadian Institute for Health Information, *The CIHI Data Quality Framework*. CIHI, Ottawa, Ont, Canada, 2009.
[37] X. Bosch-Capblanch, et al., Accuracy and quality of immunization information systems in forty-one low income countries. *Tropical Medicine & International Health*, **14** (2009), 2-10.
[38] J.S. Ancker, et al. Root causes underlying challenges to secondary use of data. *AMIA ... Annual Symposium Proceedings (1942-597X*), 2011, 57-62.
[39] D. Lowrance, et al., Assessment of a national monitoring and evaluation system for rapid expansion of antiretroviral treatment in Malawi. *Tropical Medicine & International Health*, **12** (2007), 377-381.