International Journal of Medical Informatics 86 (2016) 104–116

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.ijmijournal.com

An assessment of data quality in a multi-site electronic medical record system in Haiti

N. Puttkammer[a,*], J.G. Baseman[b], E.B. Devine[c], J.S. Valles[d], N. Hyppolite[e], F. Garilus[f], J.G. Honoré[e], A.I. Matheson[b], S. Zeliadt[h], K. Yuhas[i], K. Sherr[g], J.R. Cadet[i], G. Zamor[e], E. Pierre[i], S. Barnhart[a]

[a] International Training and Education Center for Health, University of Washington, United States
[b] Department of Epidemiology, University of Washington, United states
[c] Department of Pharmacy, University of Washington, United States
[d] Division of Global HIV/AIDS, Centers for Disease Control and Prevention, United States
[e] International Training and Education Center for Health, Haiti Office, Haiti
[f] Population Division, Ministry of Public Health and Population, Government of Haiti, Haiti
[g] Department of Global Health, University of Washington, United States
[h] Department of Health Services, University of Washington, United States
[i] National AIDS Control Program (PNLS), Ministry of Public Health and Population, Government of Haiti, Haiti

ARTICLE INFO

Article history:
Received 6 May 2015
Received in revised form 30 October 2015
Accepted 4 November 2015

Keywords:
Data quality assessment
Electronic medical record
Health information system
Haiti

ABSTRACT

Objectives: Strong data quality (DQ) is a precursor to strong data use. In resource limited settings, routine DQ assessment (DQA) within electronic medical record (EMR) systems can be resource-intensive using manual methods such as audit and chart review; automated queries offer an efficient alternative. This DQA focused on Haiti's national EMR – iSanté – and included longitudinal data for over 100,000 persons living with HIV (PLHIV) enrolled in HIV care and treatment services at 95 health care facilities (HCF).
Methods: This mixed-methods evaluation used a qualitative Delphi process to identify DQ priorities among local stakeholders, followed by a quantitative DQA on these priority areas. The quantitative DQA examined 13 indicators of completeness, accuracy, and timeliness of retrospective data collected from 2005 to 2013. We described levels of DQ for each indicator over time, and examined the consistency of within-HCF performance and associations between DQ and HCF and EMR system characteristics.
Results: Over all iSanté data, age was incomplete in <1% of cases, while height, pregnancy status, TB status, and ART eligibility were more incomplete (approximately 20–40%). Suspicious data flags were present for <3% of cases of male sex, ART dispenses, CD4 values, and visit dates, but for 26% of cases of age. Discontinuation forms were available for about half of all patients without visits for 180 or more days, and >60% of encounter forms were entered late. For most indicators, DQ tended to improve over time. DQ was highly variable across HCF, and within HCFs DQ was variable across indicators. In adjusted analyses, HCF and system factors with generally favorable and statistically significant associations with DQ were University hospital category, private sector governance, presence of local iSante server, greater HCF experience with the EMR, greater maturity of the EMR itself, and having more system users but fewer new users. In qualitative feedback, local stakeholders emphasized lack of stable power supply as a key challenge to data quality and use of the iSanté EMR.
Conclusions: Variable performance on key DQ indicators across HCF suggests that excellent DQ is achievable in Haiti, but further effort is needed to systematize and routinize DQ approaches within HCFs. A dynamic, interactive "DQ dashboard" within iSanté could bring transparency and motivate improvement. While the results of the study are specific to Haiti's iSanté data system, the study's methods and thematic lessons learned holdgeneralized relevance for other large-scale EMR systems in resource-limited countries.

* Corresponding author at: Research and Evaluation Advisor International Training and Education Center for Health (I-TECH), Department of Global Health, University of Washington, 901 Boren Ave. #1100, Seattle, WA 98104, United States. Fax: +1 2062214945.
E-mail addresses: nputt@uw.edu (N. Puttkammer), jbaseman@uw.edu (J.G. Baseman), bdevine@uw.edu (E.B. Devine), hnn4@cdc.gov (J.S. Valles), nathaelfhyppolite@itech-haiti.org (N. Hyppolite), garilusfrance@gmail.com (F. Garilus), jeanguyhonore@itech-haiti.org (J.G. Honoré), alastair.matheson@gmail.com (A.I. Matheson), szeliadt@uw.edu (S. Zeliadt), yuhask@uw.edu (K. Yuhas), ksherr@uw.edu (K. Sherr), janwonal@yahoo.fr (J.R. Cadet), garryzamor@itech-haiti.org (G. Zamor), emapierre@yahoo.com (E. Pierre), sbht@uw.edu (S. Barnhart).

http://dx.doi.org/10.1016/j.ijmedinf.2015.11.003
1386-5056/

1. Introduction

In recent years, significant investment in health information systems in resource-limited settings has been made, including in clinic-level electronic medical record (EMR) systems. These systems offer the potential to improve quality of care and population health by informing clinical decision making [1–4], health facility management [2,5,6], and health-sector surveillance, policy-making, and resource allocation [2,5]. However, both in low- and high-resource settings, data quality (DQ) within routine health information systems can be problematic [7–9]. Poor DQ in EMRs can arise from systematic errors in software, non-intuitive user interfaces, transcription errors, lack of alignment between data collected and clinical practice, lack of provider clarity about how to use data fields, and many other EMR-specific issues. It can also originate from issues beyond the EMR, including measurement error from poorly calibrated or functioning equipment, limited provider training, inadequate numbers of personnel to collect data, or other underlying weaknesses in the broader healthcare delivery system.

Strong DQ is a precursor to strong data use, because consumers and users of the data must first trust the information [10]. Funders, including the US President's Emergency Plan for AIDS Relief (PEPFAR) and the Global Fund, have emphasized DQ at health care facilities (HCF) as central to producing valid routine reports for core indicators of service coverage and other strategic information [11,12], and have developed standard tools for routine DQ audits (DQA) involving spot checks of data in primary data sources [13,14]. There has also been support for special studies of DQ at HCF, using techniques such as comparisons across data sources, chart reviews, patient exit interviews, patient tracking, or even video-recordings of patient encounters [15–26]. These methods of DQA are resource-intensive and are not necessarily realistic for HCF to use on their own on a daily, weekly, or monthly basis to monitor and improve DQ over time in primary data sources. Because EMRs collect longitudinal patient data, it is possible to use automated database queries to assess completeness, timeliness, and accuracy of data, by checking for plausible data relationships [27,28]. This can be an efficient process for routine internal DQA of EMRs in resource-limited settings.

The purpose of this evaluation is to identify strengths, limitations, and strategies for strengthening DQ within the iSanté EMR, a multi-site system implemented by Haiti's Ministry of Health (MSPP) since 2005 [29,30]. As of March 2015, iSanté had been deployed in 124HCF in all geographic regions (Departments) of Haiti. The system holds longitudinal data for more than 550,000 patients, including more than half of all patients receiving HIV care and treatment services in Haiti.

While the system uses many structured data elements and protocols and standards for data capture and transmission, such as ICD9 codes for diagnoses and HL7 messages for data exchange with laboratory information systems, concerns with completeness, accuracy, and timeliness of the data can still arise due to how users interact with the system in the context of clinical care delivery. There is anecdotal evidence from system users that DQ is perceived as weak across many of the hundreds of data elements within iSanté, leading to sub-optimal data use for patient care, health program management, and program reporting [30]. Previous analysis found, for example, that a large percentage (>20%) of HCF typically failed to run various types of automated iSanté reports, including patient summary reports, program management reports, and data quality reports, even once during the month in recent years [30]. Our evaluation contributes to the growing literature on the implementation of large-scale multi-site EMRs in resource-limited settings by demonstrating efficient methods for conducting DQA within such systems, by providing a description of the status of DQ within a multi-site EMR, and by identifying HCF and EMR system-level factors associated with strong DQ.

2. Methods

2.1. iSanté system description

The architecture of the iSanté system is represented in Fig. 1. At the HCF level, the system can be used at point of care by clinicians, retrospectively by data clerks who enter data from paper-based encounter forms that match iSanté screens, or in a hybrid manner. Some HCF have local servers hosting their own iSanté databases, with access provided to multiple workstations via local area networks, while other HCF use the system via a Web server (ASP server). Data from the Web server and each HCF local server are securely replicated to a consolidated server housed within the Ministry of Health in Port-au-Prince for system-wide data backup and analytics, as Internet connectivity permits. iSanté is one of three EMR primary data sources for the MESI reporting system, used by PEPFAR and the Global Fund to report on Haiti's national HIV response [31], and also contributes HIV case reports to the national HIV/AIDS Surveillance System (HASS) [32].

Although the overall architecture of the iSanté system has remained consistent, the software has continued to evolve over time. Initially, iSanté focused on outpatient care for patients with HIV, but the system has since expanded to cover general outpatient primary care, women's health, and malaria and tuberculosis (TB) screening, prevention, and treatment [33]. Features to support data quality, such as in-built field validations, a color-coded feature flagging forms saved with errors, and automated data quality reports, were also added incrementally over time [30].

2.2. Study, methods overview

We used mixed methods for the iSanté data quality assessment (DQA). First, using a qualitative Delphi process, we established stakeholder consensus to determine the priority data elements to investigate. Then, we performed a retrospective quantitative analysis on the prioritized data elements. The analysis included only HIV patient data, because these data had the longest history within the system.

2.3. Qualitative assessment

Delphi processes are useful for obtaining and building consensus on strategic questions, where opinions matter, especially when competing options are present [34], and have been used for stakeholder input related to implementation of electronic health records [35]. The Delphi process we employed in this evaluation involved two groups of local Haitian staff: (1) experts in HIV clinical care; and (2) experts in strategic information and HIV program reporting. The two groups identified types of information crucial for competent clinical care and public health reporting, respectively. Respondents participated in one in-person session and two follow-up on-line surveys. During the in-person session, stakeholders provided an open-ended, unprompted listing of information for which they perceived high levels of completeness, accuracy and timeliness to be highly important. They also responded to open-ended questions about their attitudes toward the iSanté system. During subsequent surveys, participants rated and ranked the types of information, each time taking into account average group responses from the prior round. This method, which is described in greater detail elsewhere [36], produced consensus on a limited set of priorities for DQ improvement within iSanté.