Evaluation of Electronic Medical Record Implementation and Data Quality in HIV Programs in Kenya and Haiti

Beryne M. Odeny

A dissertation

Submitted in partial fulfillment of the

Requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Nancy Puttkammer, Chair

Steve Gloyd

Bradley H. Wagenaar

Program Authorized to Offer Degree:

Global Health

Abstract

Evaluation of Electronic Medical Record Implementation and Data Quality in HIV Programs in Kenya and Haiti

Candidate:

Beryne M. Odeny

Chair of the Supervisory Committee:

Nancy Puttkammer,

Global Health

The global HIV agenda is to end the AIDS epidemic. Robust health information systems (HIS) are essential for monitoring, evaluating, and strengthening universal efforts to achieve the Joint United Nations Programme on HIV/AIDS (UNAIDS) 95/95/95 targets by 2030 (to ensure that 95% of people living with HIV are diagnosed, 95% of those diagnosed are on ART and 95% of those on ART are virally suppressed) [1]. EMRs are considered essential building blocks for strong HIS by improving data capture, retrieval, reporting, and overall clinical decision making for enhanced patient care [2–4]. The World Health Organization's (WHO) global digital health strategy aims to "improve health for everyone, everywhere by accelerating the adoption of appropriate digital health" – this includes EMRs [5]. Funding agencies such as the United States (US) President's Emergency Plan for AIDS Relief (PEPFAR) have funded scale up and implementation of EMRs in HIV clinics within low- and middle-income countries (LMICS) with a high HIV burden like Kenya and Haiti [6,7]. The HIV programs in these two contexts are expanding considerably along with the volume of data and reporting requirements. Phased introduction of EMRs over the past decade has led to migration from paper-based health records to electronic records in HIV clinics to increase data availability, data quality, and data use. However, there is a paucity of robust observational studies that evaluate the evidence of impact of EMRs on data quality in routine health management information systems (HMIS) and disparities in data quality across different age groups. This three-part evaluation was aimed at: a) developing and comparing novel composite data quality scoring systems; b) using interrupted time series analysis to investigate the impact of EMR implementation on the quality of data uploaded to the District Health Information System version 2 (DHIS2) and facility correlates of data quality in HIV clinics in Kenya; and c) exploring variations and disparities in data quality of EMR records across different age groups in Haiti. We used the tools of implementation science (e.g., conceptual framework and quasi-experimental design) to explore and measure data quality gaps, and make causal inference on the effect of EMRs. We generated data-driven composite data quality scores that can be used at scale to assess national DHIS2 data quality remotely. We found no evidence of impact of EMRs on the accuracy of the data uploaded to the DHIS2 in Kenya. In Haiti, we found disparities in data quality across children, adolescents, and adult age groups with poorest data quality being demonstrated among children. Our research demonstrates the need for further evaluation and investment in EMRs to harness their full potential for strengthening HIS and clinical patient management.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my deepest gratitude to my professor and chair of my committee, Nancy Puttkammer, for her invaluable guidance, kindness, and patience. I also could not have undertaken this journey without my defense committee – Steve Gloyd, Jim Hughes, and Brad Wagenaar – who generously provided knowledge and expertise. Additionally, this endeavor would not have been possible without the support and contributions from the I-TECH, Palladium, MSPP, CHARESS, CDC and Ministry of Health teams in Kenya and Haiti.

I am also grateful to my peers and cohort members for their input, feedback sessions, and encouragement in this journey. Their moral support has been priceless. Thanks also go to the study staff and data assistants who supported this work.

I would be remiss in not thanking my wonderful spouse, children, parents, siblings and friends for their steadfast prayers and belief in me. Above all, I would like to offer exceeding thanks to God Almighty for giving me life, grace and motivation to complete this journey.

Chapter 1: Introduction

BACKGROUND

The global HIV agenda is to end the AIDS epidemic by 2030 [1]. In order for HIV programs to succeed, there is need to track progress toward achievement of global targets. High quality data is essential for tracking and evaluating local and global progress toward achievement of this ambitious goal. Data quality heavily depends on completeness, accuracy and timeliness of clinical information [8,9]. Compared to paper-based records, Electronic Medical Records (EMR) systems have been shown to not only improve data quality (DQ) but also enhance timely clinical decision making for improved health outcomes [10–13].

Figure 1: Performance of Routine Information System Management (PRISM) framework

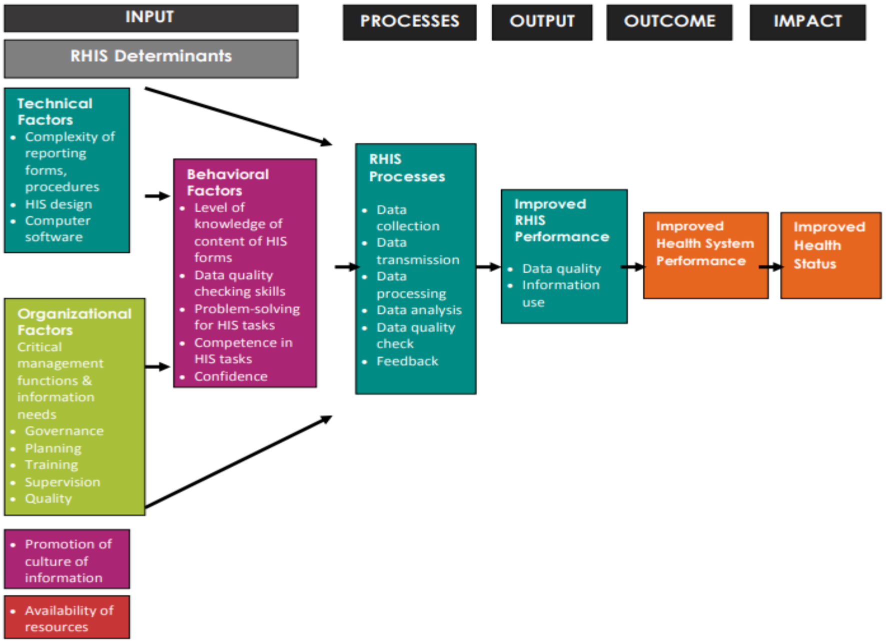


Figure 1.1 PRISM framework

Using the Performance of Routine Information System Management (PRISM) framework, we are able to conceptualize the mechanisms by which quality routine HMIS result in improved health service provision, performance, and overall patient health (figure 1.1) [14,15]. This framework illustrates the composition of a HMIS which includes inputs, processes and outputs which affect health system performance (outcomes) and lead to better health outcomes and impact [14]. This model offers an operational definition for HMIS performance which is defined as "improved data quality and continuous use of information". The desired outputs of quality HMIS include good data quality and continuous use of data which improve health system performance and overall health outcomes.

EMRs have tremendous potential to improve data quality and quality of patient care by relaying actionable information on each step of the HIV cascade [2,16–18]. EMRs with high quality data promote use and synthesis of HIV programmatic data for patient management and broader monitoring of the HIV epidemic [18–21]. Additionally, when carefully designed for usability and implemented with robust technologies, EMRs have been shown to ease healthcare providers' documentation workload [19]. In order to leverage the benefits of EMRs as conceptualized within the PRISM framework, the data captured should be complete, accurate, consistent, and reliable [14].

The past decade has seen a surge in EMR implementation within national HIV programs in LMICs including Kenya and Haiti [22–24]. As EMR use expands in these contexts, there is increasing interest in assessment of HIV programmatic DQ to identify and intervene on data issues. Poor data quality remains a formidable challenge to fully harnessing the utility of routine HMIS data in low resource settings [25,26]. The use of EMRs coupled with data quality assessments is associated with increased data quality [10,11,27,28]. Facility-based DQ assessments typically involve abstraction of information from patient charts and comparisons with EMR records or data registries. In practice, this may be time-consuming and resource intensive. DQ assessment is a familiar exercise in LMICs and different approaches have been used to quantify DQ; however, there is limited evidence on comparisons of composite DQ scores generated by different scoring approaches.

In this evaluation, we generated and compared composite DQ scores based on attributes of DQ such as completeness, plausibility, and consistency. We demonstrated the utility of assessing online national DHIS2 data summaries – which were conveniently updated using facility-based EMRs – to assess DQ in HIV programs in Kenya. In Haiti, we leveraged facility-level EMR data to investigate and demonstrate disparities in data completeness across age groups; specifically, children, adolescents, and adults receiving HIV care.

SPECIFIC AIMS

In the next three chapters we demonstrate how this evaluation was done by specific aims. In the first aim we leveraged electronic data from the District Health Information System version 2 (DHIS2) to perform DQAs at scale. For Aim 1, we generated and compared the correlation of three, data-driven composite data quality scores which include a completeness score, a consistency score, and a discrepancy (plausibility) score. Next, we used these composite scores to identify problematic HIV indicators, poorly performing health facilities, and we determined whether the scores produced consistent results in facility ranking.

For aim 2, we explored the longitudinal association between EMR implementation and HIV programmatic DQ in Kenya. Using an interrupted time series design, we investigated whether introduction of EMRs improved DQ of the aggregated HIV data uploaded in the DHIS2. Specifically, we assessed data quality as captured by a "composite discrepancy score" (a plausibility check) which encapsulates the degree of discrepancy or deviation from the expected values across groups of indicators and related data checks [29]. We evaluated the evidence for effects of EMR implementation on discrepancy scores for HIV data across 187 health facilities in Kenya from January 2011 to June 2018; and explored facility and EMR implementation correlates of data quality (i.e., the composite discrepancy score) using a cross-sectional facility survey conducted in 2017.

For Aim 3, we applied the composite score development methodology described in Aim 1 to develop Composite Completeness scores (CCS score) based on completeness of preselected clinical and care process variables in Haiti. We used this system to profile individual characteristics and facility structural factors associated with data completeness in Haiti's iSanté EMR system. We investigated whether there were disparities in completeness of HIV-related clinical and care processes data across the different age groups (children, adolescents, and adults).

Our results shed light on approaches for assessing DQ at national scale using data-driven approaches, impact of EMRs, vulnerabilities and gaps in HIS data quality, and potential areas for research, public health focus and policy formulation to strengthen routine HIS.

SIGNIFICANCE AND INNOVATION

To our knowledge, we conduct the first evaluation in SSA to characterize and compare DQ using simplistic and complex data-driven approaches. We add to the base of evidence on efficient and robust approaches for assessing data quality at scale without time-consuming data validation exercises involving comparing EMR data with external sources like registers and client charts. This evaluation also meets the need for rigorous longitudinal evaluations on the impact of EMRs on data quality. We demonstrate that facility-level EMR implementation may not have had the anticipated "spread effect" of improved data quality in other parts of the HIS. Other investments in automating linkage between EMRs and other parts of the HIS are needed. We also conduct the a comparative evaluation of data completeness of clinical and care process variables across the different age groups – children, adolescents, and adults. We demonstrate inequities and disparities in EMR data quality that were associated with age groups. This goes against the anticipated effect of digital health interventions to promote equity in healthcare access and health outcomes and illustrates the need for metrics to assess equity as digital health interventions like EMRs expand. Another significant piece of this evaluation its nation-wide scope, and rigorous quasi-experimental designs in low-resource settings. Existing literature on the effects of EMRs on data quality or quality of care have focused on small sample sizes, cross-sectional designs and pre-post designs based on a few clinics, and limited geographical regions. We evaluated 187 HF in Kenya and 58 facilities in Haiti. The results could motivate policy makers and health providers to invest more in EMR platforms to improve DQ and quality of care. Ultimately, we aim to provide evidence that will provide a framework for further evaluation of other service areas beyond HIV care in LMICs.

Chapter 2: Development of novel composite data quality scores to evaluate facility-level data quality in electronic data in Kenya: a nationwide retrospective cohort study

ABSTRACT

Background: In this evaluation, we aim to strengthen Routine Health Information Systems (RHIS) through digitization of data quality assessment (DQA) processes. We leverage electronic data from the District Health Information System version 2 (DHIS2) to perform DQAs at scale. We provide a systematic guide to developing composite data quality scores and use these scores to assess data quality in Kenya.

Methods: We evaluated 187 HIV care facilities with electronic medical records across Kenya. Using quarterly, longitudinal DHIS2 data from January 2011 to June 2018 (total N = 30 quarters), we extracted indicators encompassing general HIV services including services to prevent mother-to-child transmission (PMTCT). We assessed the accuracy (the extent to which data were correct and free of error) of these data using three data-driven composite scores: 1) completeness score; 2) consistency score; and 3) discrepancy score. Completeness refers to the presence of the appropriate amount of data. Consistency refers to uniformity of data across multiple indicators. Discrepancy (measured on a Z-scale) refers to the degree of alignment (or lack thereof) of data with rules that defined the possible valid values for the data.

Results: A total of 5,610 unique facility-quarters were extracted from DHIS2. The mean completeness score was 61.1% [standard deviation (SD)=27%]. The mean consistency score was 80% (SD=16.4%). The mean discrepancy score was 0.07 (SD= 0.22). A strong and positive correlation was identified between consistency score and discrepancy score (correlation coefficient = 0.77), whereas the correlation of either score with the completeness score was low with a correlation coefficient of -0.12 (with consistency score) and -0.36 (with discrepancy score). General HIV indicators were more complete, but less consistent, and less plausible, than PMTCT indicators.

Conclusion: We observed a lack of correlation between the completeness score and the other two scores. As such, for a holistic DQA, completeness assessment should be paired with measurement

of either consistency or discrepancy to reflect distinct dimensions of data quality. Given the complexity of the discrepancy score, we recommend the simpler consistency score, since they were highly correlated. Routine use of efficient data-driven composite scores on DHIS2 data could enhance efficiencies in DQA at scale as digitization of health information expands and could be applied to other health sectors besides HIV.

INTRODUCTION

High quality data is imperative to tracking progress to achieving UNAIDS 95-95-95 targets by 2030 [30]. In Kenya, the expansion of HIV services over the past two decades has been attended by a complementary surge in paper-based registries for documentation of care processes along the HIV care cascade. These processes include HIV testing and diagnosis, linkage to care and anti-retroviral treatment (ART), clinical and virological monitoring, Prevention of mother-to-child Transmission (PMTCT), and infant prophylaxis, among others. Inadvertently, numerous treatment indicators and registries, and their storage locations, have increased the documentation workload shouldered by a severely understaffed healthcare workforce [31–33]. These conditions, alongside other factors like erratic data entry, poor data transmission, duplicate registry of information, numerator-denominator incompatibility and infrequent use of data for decision-making, increase the likelihood for error-prone data entry and poor data quality overall [32–35].

In response to the need to bolster data management efforts in HIV treatment facilities, the Kenyan Ministry of Health (MOH), through support from the President's Emergency Plan For AIDS Relief (PEPFAR) began employing Electronic Medical Record (EMR) systems at large scale beginning in 2012 [28,36]. By 2019, at least 700 facilities were using an EMR system. In spite of the introduction of EMRs, poor data quality poses formidable obstacles to effective data utilization – a challenge that many low- and middle-income countries (LMICs) grapple with [16,33,37–40]. Systems for routine data quality assessments (DQAs) were successfully implemented by the Kenyan MOH and partners to monitor and strengthen data quality as EMRs expanded throughout the country [28].

In HIV programs in Kenya, Routine DQAs have been instrumental in identifying and resolving data quality concerns while enhancing the data's usefulness for national health programming.[41] They are vital for accurate evaluation of public health programs and interventions. DQAs assess various elements of data quality including completeness, consistency, and discrepancy (which has to do with plausibility) which map onto data accuracy (see table 2.0 for definitions) [29,41,42]. DQAs can be conducted routinely or periodically as in-person audits and/or remotely using electronic databases at facility-, regional-, or national-levels[27,28,34,41,43–52]. In-person audits primarily require physical visits to health facilities to review and cross-check documents.

Table 2.0: Data quality elements of interest and definitions [29,41,42]

Data quality element	**Definition/ Description**
Completeness	Refers to "having the appropriate amount of data present."
Consistency	Consistency referred to uniformity of data across multiple indicators. For example, number of males and females in care should be equal to number of adults and children in care.
Discrepancy	Refers to validation rules defining the possible valid values for the data element, e.g., 1) maximum and minimum acceptable value, and 2) degree of deviation from expected values.

Table 2.0 DQ elements & definitions 1

Electronic health records and databases, such as the District Health Information System version 2 (DHIS2) have also been used to assess data quality across LMICs [52]. DHIS2 is the official MOH data repository in Kenya, and it contains aggregate data on HIV and non-HIV related health indicators, with the health facility as the reporting unit. In Kenya, HIV clinics with EMRs can use aggregate EMR data to populate the DHIS2. DQAs of the web-based DHIS2 are important because this is a national surveillance and planning tool. DQAs can be conducted remotely without the need for in-person facility visits or retrieval of paper records. As such, remote DQAs of DHIS2 can be used to determine the quality of facility-level data used to populate them. While in-person facility-level DQAs are highly detailed and effective, they can be inefficient, costly (e.g., large surveys), cumbersome (involving retrieval of paper records or patient charts), and time intensive for health care workers. These inefficiencies are compounded by contextual factors in LMICs like frequent provider turnover, understaffing, competing workplace demands, multiplicity of data collection tools and indicators, and lack of provider training in DQA, limited funds, rapidly changing donor priorities, and verticalized programs without centralized data systems [32,33,53,54]. Considering these barriers, in-person assessments may not be feasibly conducted at the scale and frequency needed for growing HIV programs.

We propose retrospective analysis of DHIS2 data in Kenya to perform DQAs at scale via automated queries as a complement or alternative to time-intensive in-person DQAs. The aims of this research are three-fold: (1) to generate and compare the correlation of three, data-driven composite data quality scores which include a completeness score, a consistency score, and a discrepancy score. These dimensions of data quality were determined a priori and have been shown to be of high importance to health workers [47,55]; (2) to use these composite scores to identify individual problematic HIV indicators by clinical service department; (3) to utilize these composite scores to rank high, fair, or low performing health facilities and determine whether the scores produce consistent results in facility ranking. Our hypothesis was that these composite scores would help identify and prioritize areas (indicators and health facilities) for data quality improvement. We also hypothesized that the composite scores would be consistent in ranking poor-, fair-, and high-performing facilities.

METHODS

Study design and setting

This was a retrospective cohort study of DHIS2 data from 187 facilities implementing EMRs (specifically, KenyaEMR and IQCare EMR platforms) across HIV programs throughout Kenya (Figure 2.1). There are three main EMRs used in Kenya supporting HIV programs and we focused on DHIS2 data from facilities operating two of the most widely used: KenyaEMR and IQCare.

Figure 2.1: Map of county HIV burden in Kenya - National AIDS Control Council estimates [56]

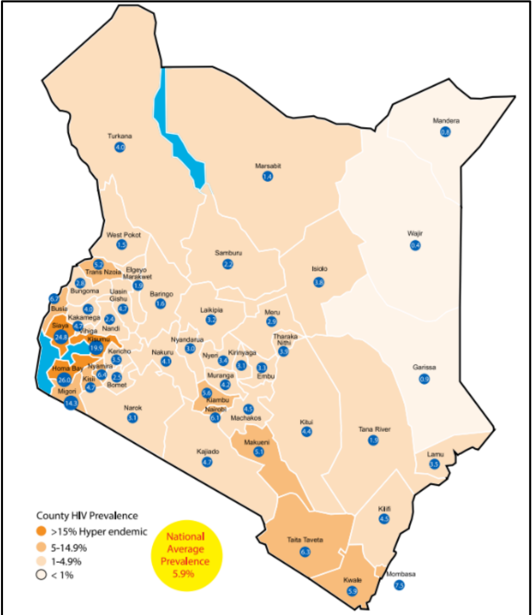


Figure 2.1 Map of county HIV burden in Kenya

Sampling

From a pool of approximately 700 facilities with EMRs, we randomly selected 129 health facilities and purposively selected an additional 58 facilities from five high HIV burden counties (Kisumu, Homabay, Siaya, Nairobi, Migori), for a total of 187 facilities across Kenya. All facilities with EMRs in high burden counties were included a priori, with the knowledge that the heavy investments in HIV programming in these regions would potentially undergird data management processes and lead to better data quality. Investments would potentially consist of staffing, computer hardware, software support, technical, and supervisory support.

Data source

DHIS2 is an internationally recognized, web-based, open-source platform used by national governments to aggregate, track, and report population data [57,58]. Ideally, health facility personnel use aggregated data from their EMRs to report on health service delivery indicators to the DHIS2 system, thus DHIS2 data quality would reflect facility-level data quality. For efficiency, DHIS2 data can be remotely assessed and evaluated for quality, as a proxy for in-person DQAs.

Data collection

Quarterly data were obtained from DHIS2 from January 2011 to June 2018. This contributed a total of 30 quarters for 187 facilities resulting in a total of 5,610 observations.

Data checks and HIV-related indicators: We used HIV-related indicators – encompassing general adult and pediatric HIV care, antenatal care (ANC), Labor & Delivery care (L&D), and Prevention of mother-to-child Transmission of HIV (PMTCT) – captured in the DHIS2. The data were aggregate numbers of service utilization by service department and were uploaded quarterly. Appropriate data checks were determined a priori, and these checks primarily summarized relationships between indicators to ensure the data were complete, consistent, and plausible. For example, one data check compared the total number of patients in HIV care in a specific quarter versus the total number on ART in that quarter. The difference between the two indicator values was expected to be zero or greater (the logic being that those enrolled would always be more or equal to those receiving ART). A series of data checks were used to construct composite scores for each unique facility. Table 2 below summarizes the data checks and indicators explored in this analysis. A list of HIV-program indicators available in DHIS2 is shown in Appendix 1.

As outlined in table 2.1, four ANC and PMTCT data checks (#1 – #4) were computed as differences between related indicator values. Similarly, four general HIV care data checks (#5 – #8) were computed as differences between general HIV care indicator values. The difference for all data checks were expected to be greater than or equal to zero (except data check # 8 which was expected to only be 0, see table 2.1).

Table 2.1 Data checks 1

Table 2.1: Data checks and composite scores for completeness, consistency, and discrepancy (plausibility)

Data check		**Completeness score***	**Consistency score***	**Discrepancy score †**
1. Number of mothers tested for HIV at ANC **minus** number of newly HIV positive mothers at ANC	Any value >=0	Score=1 if value is present, score is zero (0) if value is missing	Score= 1 if value is >=0. Score is zero (0) if not	Z-score based on extent of deviation of observed value from expected. Increasing score means higher deviation/discrepancy thus worse data quality
2. Number of mothers tested for HIV at (L&D and postnatal ward) **minus** number of newly HIV positive mothers				
3. Number of HIV-positive mothers at ANC **minus** number of mothers receiving Infant prophylaxis at ANC				
4. Number of HIV-positive mothers at (L&D and postnatal ward) **minus** HIV-positive mothers receiving Infant prophylaxis				
5. Cumulative (total ever) number of patients in HIV care **minus** Cumulative (total ever) number of patients on ART				

6. Cumulative (total ever) number of patients on ART **minus** total number currently on ART				
7. Current number of patients in HIV care **minus** current number on ART				
8. Total number of patients currently on ART **minus** sum of patients on ART across all age groups	Must be equal to zero (0)	Same as above	Score= 1 if value is = 0. Score= 0 if value is not = 0	Same as above

* Continuous score – percent complete vs percent consistent
† A measure of plausibility – small discrepancy scores are more plausible
ANC- Antenatal; L&D- Labor & Delivery, ART – Antiretroviral therapy

Description of composite data quality score generation

Three data-driven composite scores were created as follows:

1. Completeness score

The completeness score was based on the proportion of data checks which were complete for each observation i.e., each unique facility quarter. A binary score of 1 or 0 was assigned to each data check based on the presence or missingness of the data (1=present, 0= missing). The completeness score was a continuous score computed as the proportion (percent) of data checks with complete data for each observation (i.e., unique facility-quarters). The minimum possible score was 0% (0 complete out of 8 checks) and the maximum possible was 100% (8 complete out of 8 checks).

2. Consistency score

The consistency score was a continuous score based on the proportion of data checks that were consistent (i.e., scored “1”, among data checks that were not missing). A binary (0/1) score was given to individual data checks: “1” was assigned for a consistent value, and “0” assigned for an inconsistent value. The values of data checks #1-7 were expected to be equal or greater than zero, and that of data check #8 (Total number of patients currently on ART **minus** sum of patients on ART across all age groups) was expected to be equal to zero only. On this basis, values for data checks #1-7 which were less than zero (<0), and data check #8 values which were not zero, were scored “0” for being inconsistent. Conversely, values for data checks #1-7 which were greater than or equal to zero, and data check # 8 values which were equal to zero were designated a score of "1".

3. Discrepancy (Plausibility) score (further description of methodology is provided in supplementary material. Appendix 2)

This approach was based on the actual deviation or discrepancy (as a continuous measure) of observed values from the expected value i.e., how far observed values were from zero. To allow for comparison of values across facilities with varying patient volumes, we standardized the data checks by dividing them by respective facility patient volumes (i.e. general HIV care and treatment data checks were divided by the "total number of patients on ART", and ANC and PMTCT data checks were divided by the "total number of women accessing antenatal HIV testing"). Due to heterogeneity across facilities over time, deviations from the expected values were standardized using *Z*-score transformations in order to objectively compare the degree of discrepancy [59]. Specifically, the mean and standard deviation were first computed for each observed value (across all facilities) and the deviation from "0" standardized to form individual Z-scores for each data check. For data check # 1-7, individual Z-scores >= 0 were considered plausible and accurate regardless of how large or small the positive values were, thus all positive scores (> 0) were assigned a *Z*-score of 0. For data check # 8, values > 0 and values < 0 were both considered implausible, and only a value of 0 was considered accurate. For uniform handling of deviations, all negative individual *Z*-scores were multiplied by a factor of -1 and thus converted to the positive scale with the lowest possible *Z*-score value being "0". Higher *Z*-scores meant greater discrepancy or deviation from the expected values, thus poorer data quality. Lower *Z*-scores meant less discrepancy, thus better data quality. The composite discrepancy score was computed as an average of all the individual *Z*-scores for each unique facility quarter. This approach was less stringent than the consistency score as discrepant values that did not considerably deviate away from the expected value were not excessively penalized compared to larger deviations or outliers.

Ethical considerations

The program protocol was reviewed by the University of Washington Institutional Review Board (IRB) and determined to be non-research. It was also reviewed under the United Sates (US) CDC's Center for Global Health Office of the Associate Director for Science (ADS) approved protocol

(#2018-528) and also given local IRB clearance by the AMREF Ethical Scientific Review Committee (ESRC).

Statistical analysis

Descriptive statistics were used to summarize baseline characteristics. Categorical variables were summarized as counts and percentages. Continuous variables such as the composite scores were summarized with means (standard deviation) and medians (interquartile range). A descriptive bar graph and map was used to illustrate the extent of missing data across the data checks and highlight the most incomplete data. To determine how similar or correlated the three score profiles were, the Spearman rho correlation test was used. Panel plots of individual *Z*-scores for discrepancy were used to identify the high and low performing HIV indicators by service department.

For each of the three scoring profiles, facilities were ranked as high, fair, or low performing. These categories were defined using percentile cut offs; those below the 30th percentile were considered low performing, those between the 30th and 60th percentile were considered fair performing, and those above the 60th percentile were considered high performing. Sankey diagrams were used to illustrate whether the three composite scores were consistent in ranking facilities as high, fair, or low performing.

RESULTS

Baseline characteristics:

Characteristics of facilities are summarized in table 2.2.

Table 2.2: Baseline characteristics of health facilities based on 2017 facility survey

Table 2.2 Facility characteristics 1

N =187 facilities	
Facility characteristics	**N (%)**
Facility type	
County hospital	4 (2)
Sub-county hospital	69 (37)
Health center	77 (41)
Dispensary	37 (20)

High-volume facility (>500 patients)	
Yes	118 (63)
No	69 (37)
EMR type	
KenyaEMR	112 (60)
IQCare	75 (40)
High HIV-burden county	
Yes	58 (31)
No	129 (69)

Summary of composite scores:

The mean completeness score was 61.1% [standard deviation (SD)= 27.0%], illustrating that on average, approximately sixty percent of individual data checks were complete. Of 5,610 observations, 953 were missing consistency scores and discrepancy scores because they were missing all data check values for that facility-quarter (Table 4). The mean consistency score was 80% (SD=16.4%), illustrating that an average eighty percent of complete data checks were consistent. The mean discrepancy (plausibility) score was 0.07 (SD= 0.22). Please see table 2.3.

Table 2.3: Descriptive statistics of composite scores by approach

Table 2.3 Descriptive statistics- scores 1

		Total N= 5,610			
Score approach	Mean (SD)	Median (IQR)	Minimum	Maximum	Missing N (%) **
Completeness score*	61.1 (27.0)	75.0 (50.0 -75.0)	12.5	100	N/A
Consistency score†	80.0 (16.4)	83.3 (66.7-100)	0	100	953 (17)
Discrepancy score ++	0.07 (0.22)	0.02 (0 - 0.06)	0	6.16	953 (17)

*Completeness Score (%) is based on the proportion of data checks which are complete for each facility by quarter.

†Consistency score (%) was derived by calculating the proportion of complete data checks that had consistent data.

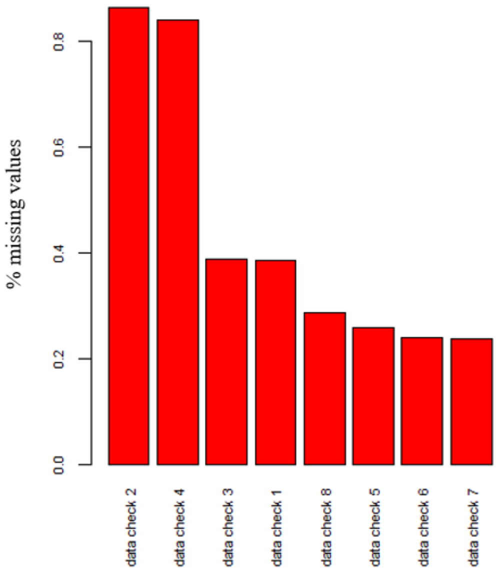
++ Discrepancy score has a lower bound of 0 and is based on *Z*-scores which depict degree of discrepancy with expected values.

SD – Standard Deviation; IQR – interquartile Range
**Refers to the number of unique facility quarters that had missing values for all 8 data checks.

Missing data:

Figure 2.2 is a bar graph and map of the missing data checks. The bar graph is ordered by magnitude of missing data across data checks. Labor and delivery HIV testing, positivity, and infant prophylaxis data checks (#2 and #4) were the most incomplete, whereas general HIV care data checks had the lowest proportions of missing data (#5, #6, #7, #8).

Figure 2.2: Percent of missing data for each data check



The y-axis of the histogram represents proportion missing

The x-axis has the following (ordered by data checks with the highest missing values to the least missing):

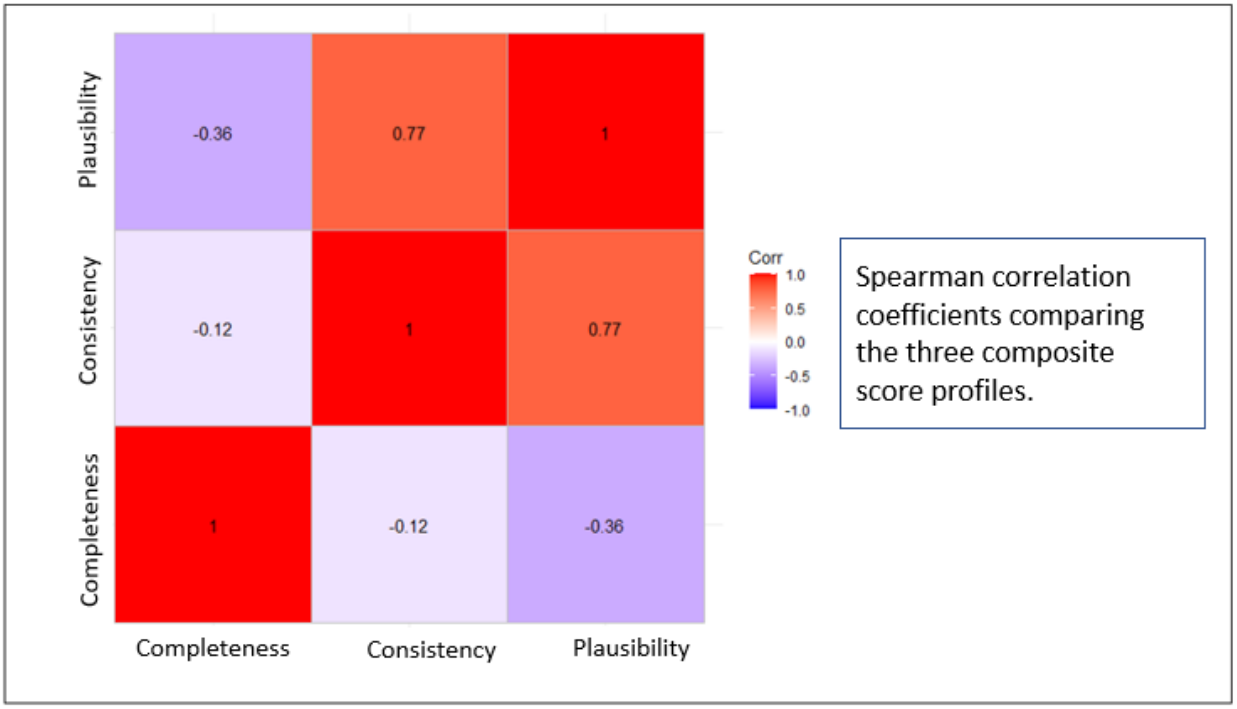
Data check 1 – ANC HIV tests vs HIV test positive;
Data check 2 – L&D HIV tests vs HIV test positive (86% missing);
Data check 3 – ANC HIV test positive & receipt of infant prophylaxis;
Data check 4 – L&D HIV test Positive & receipt of infant prophylaxis (84% missing);
Data check 5 – Total in HIV care vs total on ART;
Data check 6 – Total ever on ART vs currently on ART;
Data check 7 – Total current in HIV care vs current on ART;
Data check 8 – Total current on ART vs sum of those on ART by age group

Figure 2.2 Missing data by data checks 1

Correlation results:

The correlation plot below (figure 2.3) illustrates the strength of the correlation coefficients between the different composite scores. The completeness score was weakly correlated with the consistency score and the discrepancy (plausibility) score in the negative direction with correlation coefficients of -0.12 and -0.36, respectively. On the other hand, the consistency score and the discrepancy (plausibility) score were strongly and positively correlated (correlation coefficient = 0.77).

Figure 2.3: Correlation of the three composite score approaches



Spearman correlation coefficients comparing the three composite score profiles.

Figure 2.3 Correlation of scores

<u>High and low performing data checks</u>:

Figure 2.4 provides an overview of the discrepancy (plausibility) assessment using individual *Z*-scores and illustrates the distribution of individual scores over time. Z-scores were capped at a minimum of zero and larger/positive values were associated with greater discrepancy and deviation from expected values, thus poorer data quality and, potentially, clinical care. Most ANC- and L&D-related data checks (# 1, # 3, # 4) and one general HIV care data check # 8 (see Table 3 above) had the most favorable discrepancy (plausibility) scores, i.e., lower values closer to zero. On the other hand, most general HIV care data checks (# 5, # 6, and # 7) and one ANC data check #2 had unfavorable discrepancy (plausibility) scores with considerably larger discrepant values.

Figure 2.4: Trends for individual discrepancy (plausibility) Z-scores

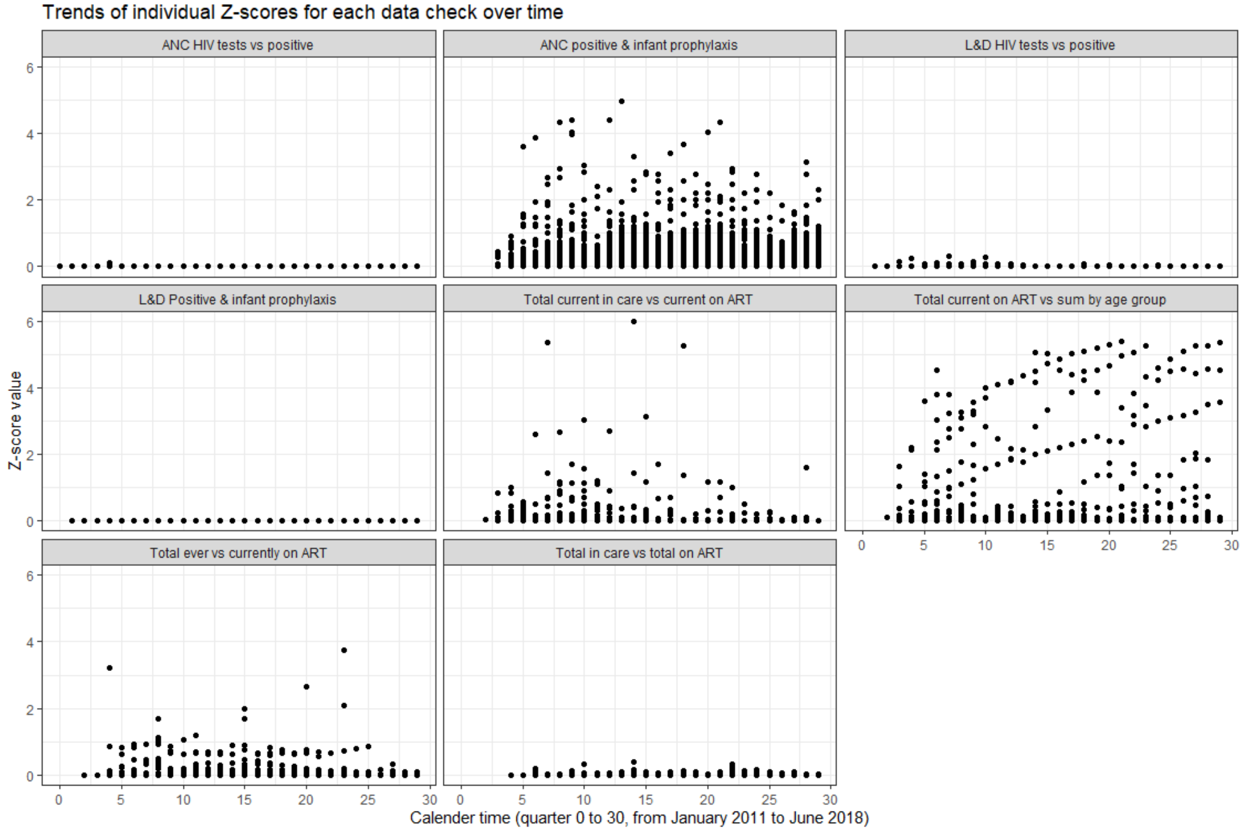


Figure 2.4 Trends of scores- data checks 1

The scatter plots above illustrate the varying degrees of discrepancy for individual data check values over time.

Table 2.4: Summary of facility ranking (across all quarters) by composite score

Table 2.4 Facility ranking 1

		Ranking by composite score (N/%)		
		Completeness (N=5610)	**Consistency** (N=4,657)	**Discrepancy** (N=4,657)
Performance N (%)	**High**	736 (13)	1221 (26)	1309 (28)
	Fair	3268 (58)	2861 (61)	2904 (62)
	Low	1606 (29)	575 (13)	444 (10)

N=unique facility quarters
Ranking is based on cut offs at the 30th and 60th percentile for each composite score. Values below the 30th percentile score are considered low performing. Values between the 30th & 60th percentile are considered fair performing, while those above the 60th percentile are high performing.

Table 2.4 provides a summary of facility ranking across all quarters by each composite score profile. While ranking was based on percentiles cut offs, the actual proportion of unique facility quarters in each of the three ranks was not commensurate to the expected proportion (i.e., ~30%) in each percentile rank because some scores clustered at the cut off points. For example, if some scores clustered at the 30th or 60th percentile, this translated to fewer scores being in the <30th or >60th percentile ranks, and so forth. We also noted considerable uniformity, across the three composite scores, in ranking/ categorization of facilities into high, fair, or low performing. More than 90% of facility observations ranked in the same category across at least two composite scores and >50% ranked in the same category across all three composite scores. Please see table 2.5 for details. Only four percent (4%) ranked differently across all the scores.

--- INSERT TABLE 2.5 ---

Table 2.5: Uniformity of ranking (high, fair, or low) across various score profiles (N=5,610)

Table 2.5 Uniformity of ranking by score 1

	All facility quarters N (%)
Uniform ranking across three composite scores	1,805 (32)
Uniform ranking across two composite scores	3602 (64)
Does not rank consistently across any of the scores	203 (4)

DISCUSSION

The DHIS2 is an important HMIS used by health sector decision makers to monitor health program implementation, to inform resource allocation in the health sector, and to understand burden of disease in the country. As such, the study of data quality in this system will be of relevance to the health sector. Kenya has a rapidly expanding national HIV program that relies on the DHIS2 system to track, monitor, and evaluate programmatic progress. We successfully developed a systematic approach for examining large volumes of data by computing three data-driven composite scores reflecting completeness, consistency, and discrepancy (plausibility) which were instrumental in determining overall data quality, and identification of high, fair, or low performing facilities. Individual completeness, consistency, and discrepancy (plausibility) scores – focusing on individual data checks – were used to detect high versus low performing indicators.

We discovered a low and slight negative correlation between the completeness score and the other two consistency and discrepancy (plausibility) scores. This lack of correlation illustrates that completeness assessment should be paired with measurement of either the consistency or discrepancy (plausibility) to reflect distinct dimensions of data quality. Contrary to our pre-specified analysis plan, we were unable to combine the completeness dimension and the consistency or discrepancy (plausibility) dimensions into one composite metric, for two reasons. First, the high degree of incomplete data would inadvertently be the primary driver of the overall score and give little weight to the other two dimensions. Second, the option of considering a weighted composite score, which combined the three dimensions, was not appropriate given the negative correlation between the completeness score and both consistency and discrepancy (plausibility) scores. We opted to explore these dimensions separately.

The completeness score was strikingly low with 40% missing data on average. The missing data graph revealed missingness as primarily driven by incomplete ANC- and L&D-related data check values. This graph could be used to chart low performing indicators to bolster precise targeting of service departments that need strategies to improve completeness. In contrast with current literature, smaller scale studies have reported higher levels of completeness of data in maternal child health indicators including PMTCT, and general HIV clinics in SSA [46,60]. In a small study of DHIS data by Garrib et al., from 10 clinics in South Africa, clinic data were 97% complete and 75% accurate or valid [31]. Another observational study of 57 South African facilities by Nicol et al. found completeness levels as high as 96% in PMTCT clinics [60]. Further, a study of three facilities in South Africa by Jamieson et al., revealed different levels of completeness (ranging from 22% to 89%) for various HIV and TB treatment indicators [61]. Muthee et al., in an evaluation of 27 clinics in Kenya found 87% completeness of HIV data - an improvement of 17% following implementation of routine DQA [28]. In a cross-sectional single-site study in Ethiopia by Abiy et al., completeness of EMR data in the HIV clinic was 76% [37]. In Malawi, a study by O'Hagan et al. demonstrated completeness was high across service departments including HIV testing and counselling, though data accuracy varied across service areas [54]. Our nationwide evaluation consists of a larger, longitudinal sample and thus offers a broader snapshot of the level of data completeness in Kenya – capturing both high and low performing areas. Of note, this may be an underestimate of the level of completeness of facility-based paper registers as providers may be more likely to complete paper registers compared to EMR.

In comparison to the completeness scores, the consistency and discrepancy scores we observed in this evaluation performed better. We found an average of 80% consistency of complete data checks. Furthermore, the mean discrepancy score suggested that most scores were within 1 SD of the expected value. These high scores of consistency and discrepancy (plausibility) map on to the broader concept of accuracy and may be comparable to high performing reliability and concordance measures examined in other settings [46,51]. For instance, a study by Endriyas et al. of 163 facilities in Ethiopia found that approximately 85% of facilities reported maternal child health data that was accurate or within the acceptable range [51].

In our evaluation, most individual discrepancy (plausibility) scores for the general HIV care data checks were scattered considerably in the larger positive direction implying greater discrepancy in the data, whereas majority of ANC- and L&D-related data checks had lower discrepancy scores, likely indicative of better performance. This surprising finding, that general HIV indicators were more complete, but less consistent and plausible than PMTCT indicators, underpins the need for in-depth exploration of the reasons behind disparate performance across departments in order to tailor interventions appropriately. A mixed picture of high completeness with low accuracy (depicted by consistency and concordance), and vice versa, was observed when exploring EMR or DHIS2 data in other SSA contexts including Malawi, Ethiopia, South Africa [31,37,40,51,54,60]. In South Africa, accuracy of PMTCT data, based on concordance across various databases and organizational levels, ranged from 51% to 84% [60]. Contrary to this finding of disparity between completeness score and consistency or discrepancy (plausibility) performance, an assessment of four-year ANC data from 495 facilities in Rwanda, revealed high performance for both completeness and consistency across data sources (absence of extreme outliers) at 98% and 83%, respectively [46]. A study in South Africa revealed relatively high accuracy and concordance of data in HIV/TB clinics ranging from 85-88% accuracy [61]. While the measures used in various studies are not similar or comparable, they provide a foundation for understanding data quality performance across SSA. Our evaluation continues to broaden this understanding as studies from other settings may be disadvantaged by shorter follow-up periods and/or small sample sizes.

With regard to uniformity, the three composite scores were considerably uniform in their ranking of facilities into high, fair, or low performing categories. A third of all facilities maintained their rank – high, fair, or low – across the three scores. Most facilities maintained their rank across at least two of the scores, primarily consistency and discrepancy. This uniformity supports the use of one of the scores to rank facility performance. We recommend the simpler consistency score over the more complex discrepancy score.

Regarding feasibility, we demonstrate that the use of composite scores which may be highly efficient compared to in-person DQAs. These composite scores are instrumental, not only with timely flagging of low performing facilities for targeted interventions, but also bring attention to high performing facilities from which program managers could learn. This study had several strengths beyond efficiency. One of the strengths was the large sample and longitudinal design which allowed repeated measurements of data quality over an eight-year time frame – an added advance over studies that have assessed the data quality cross-sectionally or over shorter timeframes [31,37,47,51,54]. Further, the use of locally accessible programmatic data adds to the relevance of this study to the Kenyan context specifically and may thus instigate policy or practice changes that improve routine health information systems in the country. Finally, with regard to external validity, the methodology used in this evaluation can be reused in other countries that use the DHIS2 system [57].

This study had some limitations. We noted higher levels of consistency and discrepancy (plausibility) in data checks that had less complete data. This could have been driven by systematic bias as less complete data would mean less opportunity to fail consistency or discrepancy (plausibility) data checks. We were unable to compare this DHIS2-based DQA with facility-based EMR or paper registry DQAs to give a holistic picture of the data quality landscape. We recommend using DHIS2 DQA for expeditious identification of facilities that need more intense DQAs which incorporate facility-level DQAs. Another limitation is a multiplicity of EMRs used throughout Kenya and we focused on DHIS2 data from facilities operating two of the most widely used: KenyaEMR and IQCare. This would limit the generalizability of our findings to reflect facility-level data quality in settings using other less commonly used EMR platforms or paper-based data systems. Due to lack of similarity of metrics across studies, it was not possible to directly compare our findings with those from other settings. We also primarily focused on HIV-related indicators which would reduce the application of our findings to other health service departments. We recommend future expansion of these composite scores to assess data quality in other service departments beyond HIV care.

CONCLUSIONS

Routine DQAs are essential for optimizing HMIS and efficient mechanisms for conducting assessments are urgently needed. The state of data quality in the DHIS2 is of great relevance to national health policy makers and thus with this evaluation, we demonstrate that national-level DHIS2 data can be harnessed to rapidly assess both data quality in DHIS2 and as a proxy, facility-level data quality. For a holistic DQA, data completeness score assessment should be paired with measurement of either the consistency or discrepancy (plausibility) score to reflect distinct dimensions of data quality. As health systems – including routine information systems – weather the challenges of the COVID-19 pandemic, routine use of efficient-data driven composite scores on DHIS2 data may be a feasible approach to monitoring and upholding data quality in LMICs.

AIM 1 APPENDIX

Table of HIV -indicators captured in DHIS2

Table 2.6 DHIS2 HIV indicators 1

Antenatal Testing for HIV
Labour and Delivery Testing for HIV
Postnatal (within 72hrs) Testing for HIV
Antenatal Positive to HIV Test
Labour and Delivery Postive to HIV Test
Postnatal (within 72hrs) Postive to HIV Test
Known positive status (at entry into ANC)
Prophylaxis - interrupted HAART
Prophylaxis â€“ HAART
Issued in ANC (Infant ARV prophylaxis)
Labour and Delivery (Infant ARV prophylaxis)
PNC (<72hrs) (Infant ARV prophylaxis)
PCR (within 2 months) Infant Testing (Initial test only)
PCR (from3 to 8 months) Infant Testing (Initial test only)
PCR (from 9 to 12 months) Infant Testing (Initial test only)
Total Enrolled in Care
Total Starting on ART
HIV Currently in Care - Total
Total Revisit on ART
Total currently on ART
Total Ever on ART
HIV Care visit Scheduled
HIV Care visit- unscheduled
Currently on ART - Female Below 15 years
Currently on ART - Female above 15 years
Currently on ART - Male above 15 years
Currently on ART - Male below 15 years
Currently on ART - below 1 year

Chapter 3: The effect of introducing Electronic Medical Records on data quality and factors associated with data quality across 187 HIV clinics in Kenya, 2011-2018: An interrupted time series analysis

ABSTRACT

Background

The implementation of health facility-level electronic medical records (EMRs) may improve quality of data in the District Health Information System version 2 (DHIS2). The objective of this evaluation was to investigate the evidence for effect of EMR implementation on HIV program data quality in the DHIS2.

Methods

This was an interrupted time series analysis of DHIS2 data. Data was extracted from 187 Kenyan health facilities from January 2011 to June 2018 (i.e., spanning 30 quarters). We assessed data quality as captured by composite discrepancy scores based on pre-selected HIV-related indicators. Facility and EMR implementation factors associated with data quality were also explored. The primary exposure was presence of EMR, and the primary outcome was quarterly composite discrepancy scores. The composite discrepancy score depicted the extent of deviation of observed values from expected values for HIV-related indicators. A cross-sectional facility survey done in 2017 was used to assess factors associated with data quality. Linear regression analyses of the cross-sectional survey were used to determine factors associated with data quality.

Results

There were no statistically significant changes in level and slope, thus no effect of EMRs on composite discrepancy scores in the post-EMR period. The level change was 0.0118 points (95% CI -0.0248 – 0.0484, p value 0.528). The slope change was 0.0004 points (95% CI -0.0038 – 0.00347, p value 0.844). However, facility type, facility volume, and EMR type were significantly associated with data quality. In adjusted analyses, health centers had an average composite discrepancy score that was 0.093 points higher than that of dispensaries (95% CI 0.026-0.159, p value = 0.007); facilities with ≥ 500 patients had a score that was 0.084 points higher than that of facilities with <500 patients (95% CI 0.033-0.136, p value <0.001); and facilities using

KenyaEMR had an average composite discrepancy score that was 0.082 lower compared to facilities using IQCare (-0.082, 95% CI -0.132- 0-0.031, p value 0.002).

Conclusions

EMR implementation did not show evidence of benefit on DHIS2 data quality as captured by the composite discrepancy score. However, facility type, facility volume, and the type of EMR used in a facility were associated with data quality. These findings may be explained by the fact that presence of EMRs alone is not insufficient to achieve high DHIS2 data quality. Implementation science involving qualitative research is needed to understand the reasons behind lack of improvement in data quality. We demonstrate that facility-level EMR implementation may not have had the anticipated "spread effect" of improved data quality in other parts of the HIS. Other investments in automating linkage between EMRs and other parts of the HIS are needed. provide a framework for further evaluation of other service areas beyond HIV care in LMICs.

INTRODUCTION

Robust routine health information systems (RHIS) are essential for optimal health system evaluation, quality improvement, governance, and health management [62–64]. The WHO's global digital health strategy aims to "improve health for everyone, everywhere by accelerating the adoption of appropriate digital health" – this includes EMRs [3,4]. EMRs are considered essential building blocks for strong HMIS [2,65]. EMRs streamline data capture, data retrieval, and reporting which facilitate data use for clinical decision making and enhanced patient care. Furthermore, EMRs have also been found to reduce data recording errors [11,64,66] as well as time/ monetary costs of data management [67–70]. EMRs have a range of purposes crucial to health system strengthening including data reporting, aggregation and management, supporting clinical decision making, and interlinking health departments). This paper aims to add to the base of evidence on the utility of EMRs as a tool for strengthening data management – specifically, data quality in aggregate data reporting [75,76].

EMR utilization in healthcare in low- and middle-income countries (LMICs) has dramatically increased over the past two decades. The United States (U.S.) President's Emergency Plan for AIDS Relief (PEPFAR) has funded scale up and implementation of EMRs specifically for HIV care in health facilities throughout Kenya – one of the high-burden HIV countries in sub-Saharan Africa (SSA). Kenya was among the first high HIV burden countries to publish national strategies for EMR integration in the health system [3,4]. In Kenya, the District Health Information System (DHIS2) – a national electronic health information system – has been used for over a decade to house aggregated facility-level data. Version 2 of the program is an open-source platform (www.DHIS2.org) [77]. Health facilities collate patient-level data to prepare summaries that are uploaded to the DHIS2 routinely on a monthly basis. The data in DHIS2 are used in planning health service delivery, including planning of health personnel, supply chain management, among other functions, and as such high standards of data quality need to be maintained [45,57,58].

Before the introduction of EMRs in Kenya, facility-level HIV data were captured on paper charts and registries. These paper records were used to generate HIV-specific facility summaries for the DHIS2. Phased introduction of EMRs in 2012 led to progressive migration from paper records to a hybrid implementation of paper and electronic records in most HIV facilities. According to programmatic records in mid-2018, about 23% of target facilities have fully transitioned to EMR and 75% are using both paper and EMRs. Collation of electronic data was found to simplify the aggregation of facility-level HIV data, thus use of EMRs to generate HIV-specific facility reports for the DHIS2 expanded as EMRs became widespread. While the DHIS2 hosts data from other sectors beyond HIV service delivery, EMRs are primarily tailored for HIV data management, thus EMR reports uploaded to DHIS2 are HIV-specific. While EMRs are associated with better data quality [28,78], they are not immune to error [75], thus routine data quality assessments are essential to uphold EMR data quality [79]. Improved data quality in EMRs would potentially lead to improved quality of aggregate data uploaded to the DHIS2.

To our knowledge, high quality evidence on the longitudinal effects of EMRs on data quality in the DHIS2 aggregate data are limited or non-existent in LMICs. Using an interrupted time series design, we aim to evaluate whether introduction of EMRs improved data quality of the aggregate HIV data in the DHIS2. Specifically, we assessed data quality as captured by a "composite discrepancy score" (a plausibility check) which encapsulates the degree of discrepancy or deviation from the expected values across groups of indicators and related data checks [29]. The objectives of this evaluation are two-fold: 1) to evaluate the evidence for effects of EMR implementation on discrepancy scores for HIV data across 187 health facilities in Kenya from January 2011 to June 2018; and 2) to assess facility and EMR implementation correlates of data quality (i.e., the composite discrepancy score) using a cross-sectional facility survey conducted in 2017.

METHODS

Setting

This was a quasi-experimental study utilizing time series data from the national DHIS2 HMIS in Kenya. Data was extracted quarterly from January 2011 to June 2018 for 187 facilities implementing EMRs – specifically, KenyaEMR and IQCare EMR software.

Data sources

Data sources were the DHIS2 database and I-TECH and Palladium program records on EMR deployment and implementation process. I-TECH and Palladium were the two predominant EMR technical assistance providers in Kenya during the timeframe of this analysis. Program records capture the date of KenyaEMR or IQCare deployment at each health facility. The primary exposure is presence of EMR (a binary pre or post variable). The primary outcome is composite discrepancy scores, as defined below.

Computing the composite discrepancy score

The composite discrepancy score was based on assessment of plausibility across indicators and eight predefined data checks. Individual data checks were derived by comparing groups of indicators. For example, one data check was derived by comparing the total number of patients currently on ART (an indicator) minus sum of patients on ART across all age groups (another indicator) and this was ideally equal to zero; the further away the data check value was from the expected value (zero in this case), the higher the individual discrepancy score (Appendix 1). These individual scores were computed for each data check and were based on *Z*-score deviations that depicted the extent of discrepancy of observed values i.e., how far observed values were from expected values. High discrepancy scores indicated poorer data quality. The composite discrepancy score was a continuous variable computed as an average of all the *Z*-scores for individual data logic checks, for each unique facility quarter. Given the composite discrepancy score is *Z*-score standardized, we expect that ideally 68% of scores would be ≤ 1 *Z*-score. For additional context, a score that is 0.01 lower, is comparatively small but represents a small improvement in data quality performance, and vice versa. Additional details on the development of this score are provided in the Appendix 2 and in a paper detailing the composite score development [80].

Cross-sectional survey

A cross-sectional facility survey done in October to December of 2017 was used to collect data on non-time varying facility and EMR implementation factors which are potentially correlated with data quality such as facility type, facility volume, and EMR type. Survey data on EMR implementation characteristics was only available for 115 health facilities.

Ethical considerations

This evaluation was approved by the AMREF Ethical Scientific Review Committee (ESRC), Kenya, United States (US) Center for Disease Control and Prevention (CDC), and the University of Washington (UW) Human Subjects Review committee.

STATISTICAL ANALYSIS

Descriptive statistics

An analysis plan was prepared prospectively prior to data analysis. Descriptive statistics were used to analyze the distribution of health facility characteristics. We reported the number of facilities included in the analysis by facility type (national hospital, county hospital or sub-county hospital, health center, dispensary), facility volume ($\geq$ 500 versus <500 HIV patients in care), region/county (46 counties classified into low and high-HIV burden counties), and EMR type (KenyaEMR, IQCare or other). Median and interquartile ratios were used for continuous data, whereas proportions were used for binary and categorical data. Baseline characteristics were summarized and presented in Table 3.2. The autocorrelation function (ACF) and partial autocorrelation function (PACF) plots were assessed to determine the degree of autocorrelation (Appendix).

Proposed impact model [81]

The proposed outcome, composite discrepancy score, was hypothesized to be responsive to the intervention of EMR implementation on a gradual basis. Based on the literature, we hypothesized a 6-quarter (i.e., 18 months) time lag after EMR introduction to account for the "wash-in period" after EMR introduction. It was expected that EMR implementation would improve over time until optimal functioning was achieved [72,78]. Therefore, a gradual change in the composite discrepancy score was anticipated over the 6-quarter "wash-in period", followed by a plateauing of the intervention effect after 6 quarters as demonstrated in Figure 1(a) (impact model 1). The final impact model which informed the statistical model and analysis, excluded the "wash-in period", thus only two segments were modeled: pre-EMR, and post-EMR (starting at 7 quarters post-intervention). Eliminating the "wash-in" period from the model would make any potential level change – in the outcome post-EMR – more discernable. This is illustrated in Figure 3.1 (final impact model 2).

Figure 3.1: Impact model

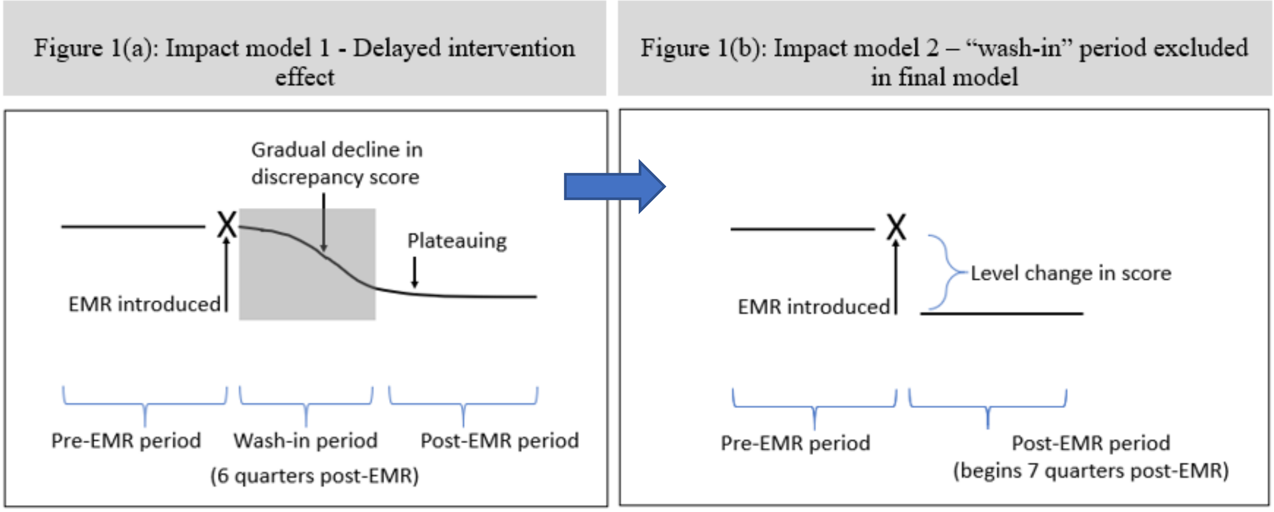


Figure 1(a): Impact model 1 - Delayed intervention effect

Figure 1(b): Impact model 2 – "wash-in" period excluded in final model

Figure 3.1 ITS impact model 1

Inferential analyses using Interrupted Time Series (ITS)

Multilevel models were used to control for clustering at the health facility level. Changes in the data quality scores pre- and post- EMR implementation were estimated using generalized linear models (GLM) with Gaussian distribution and identity link and bootstrapped standard errors. Random effects for facilities (including random intercepts and random slopes for EMR effect) were used to account for intra-cluster correlation. Data was assumed to be missing at random. We included data points beginning at 14 quarters pre-EMR implementation.

The primary model included dummy variables for quarter (i.e., season) to account for seasonality and random intercept to account for facilities having different baseline DQ scores.

$$Yj\text{t} = \beta_0 + \sum_{m=2}^{m=4} \beta_m Season + \beta_4 Time + \beta_5 EMRjt + \beta_6 \text{Time\_after\_EMRjt} + \text{u}j + \text{e}jt$$

where j=facility, t= quarter, EMR_{jt} = Intervention introduced, u_j = random intercept for facility, e_{jt}= error term (Table 3.1 defines coefficients further).

Table 3.1 Definition of coefficients 1

Table 3.1: Definition of coefficients in the final ITS model

Intercept: baseline mean DQ score at time=0 (i.e., Quarter 1, 2011) for an average facility.

Time: the mean change in DQ score per quarter before EMR implementation for an average facility.

Season indicator: the difference in mean DQ score comparing each Quarter X to Quarter 1 (reference) for an average facility.

EMR: mean change in DQ score after the 6-quarter "wash-in" period compared to DQ score immediately before EMR introduction for an average facility. This is the level change.

Time_after_EMR: Slope in the post-EMR period. This is the mean change in DQ score each quarter post-EMR implementation compared to quarterly trend before EMR implementation for an average facility. This is the slope change.

Cross-sectional survey analysis:

Linear regression models were used to assess the relationship between EMR and facility characteristics and the composite discrepancy scores (as a measure of data quality). We used the composite discrepancy scores from quarter 4 of June 2017, i.e., October to November 2017. By this time, all facilities had an EMR and the survey was conducted post-EMR. A priori, based on programmatic relevance, we selected EMR type, facility volume, and being in a high-burden county to be included in the multivariable model. Health facilities missing data on facility and EMR implementation characteristics were excluded from the analysis, i.e., complete case analyses. All statistical analyses for this evaluation were done using R studio version 3.6.2 (2019-12-12).

RESULTS

Based on the 2017 cross-sectional survey, most facilities (73%) in the sample were health centers and sub-county hospitals (table 3.2). Fifty six percent (56%) of facilities were high volume, attending to more than 500 patients in their HIV clinics. Thirty six percent (36%) of facilities were in high-HIV burden counties. Most facilities in this sample implemented KenyaEMR software (56%). With regard to EMR implementation characteristics, 55% facilities used retrospectively entered data after completing paper forms. Most facilities (55%) reported at least one routine data quality assessment exercise that year. Only 17% of facilities reported equipment failure in the preceding week, and most facilities had at least one or two people trained to use EMR. Assessment of ACF and PACF plots revealed minimal autocorrelation – autocorrelation dropped from 0.4 in ACF plot to 0.2 in the PACF plot at lag=1.

Table 3.2: Health Facility (n=187) and EMR implementation characteristics (n=115)

Table 3.2 Baseline characteristics 1

Health facility characteristics (n=187)	**n (%)**
Facility type	
Dispensary	45 (24)
Health Center	68 (36.5)
County/ Sub-County Hospital	68 (36.5)
Missing	6 (3)
Facility volume	
Low volume (<500)	105 (42)
High volume (≥500)	79 (56)
Missing	3 (2)
High HIV burden county	
No	120 (64)
Yes	67 (36)
EMR type	
IQCare	82 (44)
KenyaEMR	105 (56)
EMR implementation characteristics (data from N=115 facilities)	
Implementation mode	
Hybrid[†]	23 (20)
Point of care entry	26 (23)
Retrospective data entry	64 (55)
Uncategorized	2 (2)
Annual number of RDQAs[‡]	
0	27 (23)
1	63 (55)
≥2	9 (8)
Missing	16 (14)
Equipment failure (past week)	
No	95 (83)
Yes	20 (17)

[†] Hybrid mode – facility practices both point of care entry and retrospective data entry

[‡] RDQAs- Routine Data Quality Assessments within the year preceding the survey

Table 3.3 provides a descriptive summary of the discrepancy score which had a mean (SD) of 0.07 (0.22), and a median (IQR) of 0.02 (0-0.06).

Table 3.3: Summary of composite discrepancy scores

Table 3.3 Summary of discrepancy scores 1

		Total N= 187 facilities over 30 quarters		
Score approach	Mean (SD)	Median (IQR)	Minimum	Maximum
Discrepancy score	0.07 (0.22)	0.02 (0 - 0.06)	0	6.16
SD – Standard Deviation; IQR – interquartile Range				

The interrupted time series graph, centered on time of EMR implementation, shows an initial drop in discrepancy scores between 14 and 13 quarters pre-EMR (figure 3.2); however, scores have remained stable over the duration of the evaluation signaling minimal changes to the trend of data quality before and after EMR implementation.

Figure 3.2: Time series plot of the average composite discrepancy scores across all 187 facilities

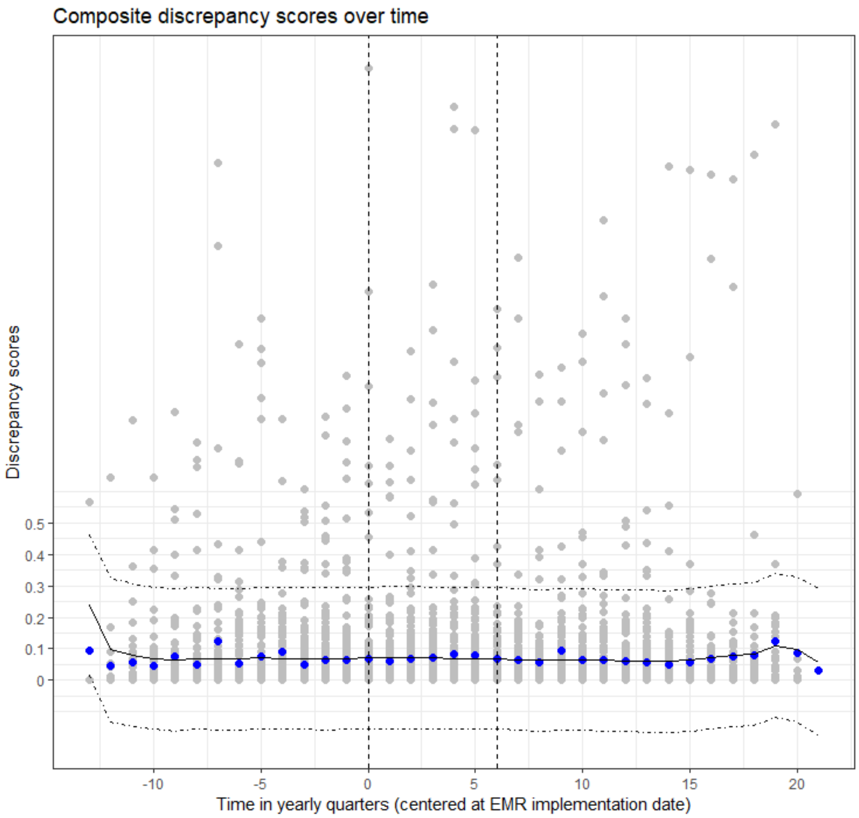


Figure 3.2 Time series graph of scores

Figure 3.2 footnote: The data centered on the date of EMR implementation and adjusted for seasonality and autocorrelation: Blue dots in figure 1 represent the average discrepancy score for all the facilities by quarter; gray dots represent individual facility discrepancy scores; the solid line represents the time series plot; the dashed lines represent the prediction intervals

Table 3.4 summarizes the parameters of the time series model. The average baseline composite discrepancy score at EMR implementation (centered time=0) was 0.0753 [95% confidence interval (CI): 0.0482 – 0.1025]. In the pre-EMR period, the average quarterly discrepancy score decreased by 0.0012 points (95% CI: -0.0045 – 0.0020, p-value =0.458), i.e., there was no significant change in trend.

Regarding level change, comparing the values post-EMR, and pre-EMR period; the average composite discrepancy score in the first quarter post-EMR was 0.0118 points higher (95% CI: (-0.0248 – 0.0484, p-value =0.528) than the average score in the quarter preceding EMR implementation for the average facility. This change was not statistically significant. Comparing the trends post- and pre-EMR implementation; the average change in the composite discrepancy score each quarter post-EMR was 0.0004 points higher (95% CI: -0.0038 – 0.0047, p-value = 0.844 than the average quarterly trend pre-EMR implementation for the average facility. This slope change was non-significant (Table 3.4).

Average discrepancy scores were compared across the four seasons in the pre-EMR period. Season did not predict level of discrepancy score when comparing other seasons with the first calendar quarter (Table 3.4).

Table 3.4 Time trends- discrepancy score 1

Table 3.4: Crude time trends in discrepancy scores across 187 health facilities with EMRs, 2011 – 2018, Kenya

	β*(95% CI)	**p-value**
Discrepancy score		
Average change in pre-EMR period (slope)	-0.0012 (-0.0045 – 0.0020)	0.458
Level change following EMR implementation	0.0118 (-0.0248 – 0.0484)	0.528
Average change in post-EMR period (slope)	0.0004 (-0.0038 – 0.0047)	0.844
Baseline score at EMR implementation†	0.0753 (0.0482 – 0.1025)	--
Average quarterly change in reference to season 1 (January to March) pre-EMR		
Season 2 - April to June	-0.0173 (-0.0353 – -0.0007)	0.059
Season 3 - July to September	-0.0097 (-0.0287 – 0.0092)	0.315
Season 4 - October to December	-0.0104 (-0.0292 – 0.0084)	0.278

Computation of Intracluster correlation coefficient (ICC)

Between-cluster variation	0.015
Within- cluster variation	0.034
ICC	0.306

β*represents unadjusted quarterly change in mean DQ score using generalized linear mixed models with random intercept for clinics. Time was centered at date of EMR deployment. Time was modeled with a knot 6 quarters after EMR introduction.

With regard to the post-EMR cross-sectional survey of health facility and EMR implementation factors associated with data quality, significant findings in unadjusted and adjusted analysis have been described below and in Table 3.5. In unadjusted analysis, facility type, facility volume, being in a high HIV-burden county, and EMR type were associated with data quality. After adjustment, all these factors remained significant except for the association with high HIV-burden county. Health centers, on average, had a composite discrepancy score that was 0.066 (95% CI: 0.002-0.130) points higher than that of dispensaries, thus a higher degree of discrepancy indicating poorer data quality (p-value = 0.045). High volume facilities (>500 patients on ART,) on average, had a score that was 0.090 (95% CI: 0.043-0.138) points higher than that of low volume facilities, thus a higher degree of discrepancy indicating poorer data quality (p-value <0.001).

Facilities with KenyaEMR, on average, had a score that was 0.058 (95% CI: -0.107- -0.008) points lower than that of facilities with IQCare, thus a lower degree of discrepancy indicating better data quality (p-value = 0.024). Of note, being in a high HIV burden county was not associated with the discrepancy score. Implementation mode (whether point of care, retrospective data entry, or hybrid), annual number of RDQAs, and having experienced equipment failure in the preceding week had no significant association with the discrepancy score.

Table 3.5: Facility and EMR implementation correlates of data quality post-EMR – cross-sectional analysis of survey done in 2017, quarter 4 (October – December)

Table 3.5 Correlates of data quality 1

N=187					
Characteristic	**N (%) Median (IQR)**	**β*-unadjusted (95% CI)**	**p-value**	**β** - adjusted (95% CI)**	**p-value**
Facility type					
Dispensary	45 (24)	Ref		Ref	
Health Center	68 (36.5)	0.066 (0.002-0.130)	**0.045**	0.093 (0.026-0.159)	**0.007**
County/Sub-county hosp	68 (36.5)	0.042 (-0.022-0.106)	0.197	0.045(-0.25-0.114)	0.212
Missing	6 (3)				
Facility volume					
<500	105 (42)	Ref		Ref	
>500	79 (56)	0.090 (0.043-0.138)	**<0.001**	0.084 (0.033-0.136)	**0.001**
Missing	3 (2)				
High HIV burden county					
No	120	Ref		Ref	
Yes	67	0.065 (0.015-0.116)	**0.011**	0.042 (-0.011-0.094)	0.122
EMR type					
IQCare	82 (44)	Ref		Ref	

KenyaEMR	105 (56)	-0.058 (-0.107- -0.008)	**0.024**	-0.082 (-0.132- 0-0.031)	**0.002**
EMR implementation characteristics (data from N=115 facilities)					
Implementation mode					
Hybrid[†]	23 (20)	Ref		Ref	
Point of care entry	26 (23)	-0.024 (-0.137-0.090)	0.684	-0.012 (-0.121-0.097)	0.827
Retrospective data entry	64 (55)	0.016 (-0.079-0.111)	0.739	-0.012 (-0.109-0.085)	0.808
Uncategorized	2 (2)	-0.037 (-0.326-0.252)	0.801	-0.013 (0.033-0.184)	0.945
Annual number of RDQAs[‡]					
0	27 (23)	Ref		Ref	
1	63 (55)	-0.033 (-0.131-0.065)	0.511	-0.039 (-0.135-0.057)	0.426
≥2	9 (8)	0.009 (-0.154-0.171)	0.916	-0.008 (-0.172-0.156)	0.922
Missing	16 (14)				
Equipment failure (past week)					
No	95 (83)	Ref		Ref	
Yes	20 (17)	-0.020 (-0.118-0.078)	0.684	-0.034 (-0.128-0.060)	0.474

β*represents the crude relationship between DQ score and key facility and EMR implementation factors

β** represents the adjusted relationship with DQ score adjusting for being in a high vs low HIV burden county, facility volume, and EMR type

[†] Hybrid mode – facility practices both point of care entry and retrospective data entry in parallel

[‡] RDQAs- Routine Data Quality Assessments

DISCUSSION

In this evaluation we examined the association of EMR implementation on data quality, as captured by composite discrepancy scores in the DHIS-2 system across 187 facilities in Kenya from 2011 to 2018. Overall, changes in the data quality score comparing different periods pre- and post-EMR were very small (most estimates were smaller than 0.01) and thus within 1 *Z*-score i.e., 1 standard deviation (SD). We found no evidence that EMR implementation was associated with DHIS2 data quality as captured by the discrepancy score. This could be attributed to the already low (i.e., good or favorable) baseline discrepancy scores.

A cross-sectional facility survey done in 2017 unveiled health facility and EMR implementation factors that were associated with data quality. While the differences in Z-score deviations were mostly less than 0.1, the differences were significant enough to reveal associated factors. Health centers were associated with higher discrepancy scores (thus worse data quality) compared to dispensaries. This could be because of higher patient volumes without commensurate increase in number of health workers including data managers. When comparing county/ sub-county hospitals with dispensaries there was no difference in the average discrepancy scores. Higher facility volumes, with patient numbers >500, were associated with higher discrepancy scores thus poorer data quality. This could be an indicator that facility volume is a stronger factor than number of staff and high workload among health providers is a significant barrier to accurate documentation and integration of EMR to facility workflows [31,82].

Operating KenyaEMR software was associated with lower discrepancy scores – thus better data quality – compared to using IQCare. The underlying reasons could be differences in technical support (i.e., implementation partners working with ministry of health) and usability profiles of the two EMR systems. Other factors which were examined included number of annual RDQAs and mode of EMR implementation. Having one RDQA was associated with better data quality compared to no RDQA; however, this was non-significant. Having two or more RDQAs was associated with poorer data quality. While this association was non-significant, the finding could be explained by reverse association if recurrent RDQAs were implemented for remedial purposes. We found that mode of implementation (i.e., whether point-of-care, retrospective or hybrid data entry) was not associated with data quality.

There is existing research describing data quality in the context of electronic data systems, but less research examining the evidence for causal effects of EMR implementation on data quality. Previous studies in sub-Saharan Africa have shown a mixed picture in terms of levels of data accuracy for HIV-related indicators. Several studies have shown high data quality of EMR records [11,31,37,46,47,54,60,83], while others have demonstrated overall poor EMR data quality [51]. Notably, there were no studies that compared data quality in EMRs and paper records longitudinally. A pre-post study in Iran found that vocal-electronic documentation was associated with higher quality data than paper documentation [84]. Two cross-sectional studies compared electronic and paper records concurrently. These studies found that data quality was comparable across the electronic and paper records [37,85]. A one-month randomized controlled study in Ethiopia demonstrated better data quality and efficiency with electronic compared to paper-based survey data [86].

Our null findings suggest that presence of EMRs alone is insufficient to achieve strong aggregated data quality used in reporting; optimization of EMR usage and data exchange is a necessary requirement for enhancing data quality [40]. Some barriers to improved data quality during EMR implementation include sub-optimal EMR use which may be associated with insignificant changes in data quality post-EMR implementation [31]. Recent studies evaluating the actual use of EMR implementation have demonstrated concerningly low EMR usage and data exchange [31,87]. Ngugi et al, and colleagues measured seven EMR usage indicators and found gaps in active use of EMRs and data completeness in Kenya [87]. Factors affecting EMR usage need further exploration and remedial interventions [33,37,51,79,88]. An important consideration is variability in the approach health workers use to generate EMR aggregate summaries that are reported in DHIS2. It is likely that some health workers may manually transcribe EMRs summaries for retrospective upload to DHIS2, while others may not use the EMR and thus the EMR and DHIS2 would run parallel to each other.

Dual/ parallel documentation in EMR and paper records is huge barrier to EMR usage and good data quality. In this evaluation, we found that 75% of facilities had a hybrid or retrospective data entry system – they performed both point-of-care and retrospective data entry, which would contribute to inefficiencies and errors in data management. A study in Malawi found that retrospective data entry by clerks was associated with more errors than point of care entry by clinicians [89]. Assisting facilities to channel efforts toward point-of-care EMR usage would be a possible solution for improving data quality [66,89,90]. A study in Kenya established the benefit of RDQAs in improving EMR data quality [28]. While the sub-analysis of factors in the present evaluation showed that there was no association between RDQAs and data quality, it is acknowledged that the cross-sectional survey may have had inherent bias and insufficient power to detect all significant associations.

This evaluation points to a constellation of factors beyond the presence of EMR that may impact DHIS2 data quality. There are other factors that may be at play that can affect data quality such as facility type, facility volume, and EMR type, as demonstrated by the cross-sectional survey. The survey demonstrated that a suboptimal proportion of facilities (23%) maintained fidelity to the intended mode of EMR use, i.e., point-of-care entry. Another important consideration is variability in the approach health workers may use to generate EMR aggregate summaries that are reported in DHIS2. It is likely that some health workers may manually transcribe EMRs summaries for retrospective upload to DHIS2, while others may not use the EMR summaries, thus the EMR and DHIS2 would run parallel to each other. In resource limited settings, it can take a considerable duration of time to implement EMR software – this should be kept in mind when evaluating EMRs

[13]. A parallel focus on EMR usage, data quality, and associated facility or EMR system factors is a prerequisite for optimizing HMIS in LMICs [91].

EMR implementation is feasible; however, the challenges of streamlining use of EMR, and lack of interoperability with DHIS2, are considerable. Other noteworthy implementation challenges include operationalizing EMRs, health system constraints, and upfront cost-related issues [65,91–93]. More ongoing work is needed to understand systemic deficits that hinder health provider acceptance, uptake, and meaningful use of EMRs [94]. It can be hypothesized that poor uptake and use of EMR would limit realization of the benefits of EMRs for DHIS2 data reporting. This work should be coupled with continuous quality improvement efforts and design of contextually appropriate implementation strategies to strengthen the uptake of EMRs. Strategies such as training health providers, infrastructure support, modifying electronic tools, designing usable and interoperable systems, data audits, and sustainable technical support from ministries of health are critical for optimal implementation of EMRs [2,95–97].

Strengths:

The DHIS2 is an important HMIS used by health sector decision makers to monitor health program implementation, to inform resource allocation in the health sector, and to understand burden of disease in the country. As such, the study of data quality in this system will be of relevance to the health sector. This evaluation serves as a blueprint for future longitudinal evaluations of mature health information systems. To our knowledge this is the first nationwide longitudinal analyses of data quality that spans pre-EMR and post-EMR periods that assesses the causal relationship between EMR implementation and aggregate data quality. Prior studies in LMICs have considerable limitations including small sample sizes, had shorter follow-up periods or cross-sectional designs, use of less robust designs, or lack long-term comparative evaluations of EMRs and paper systems. In light of this, the present evaluation is a much-needed addition to the literature. Another strength includes the use of the time series models which eliminate confounding by time invariant factors. The stepped/ differential introduction of EMRs across multiple facilities by default creates a multiple baseline design which is useful in controlling for possible time-varying confounders [98].

Limitations:

While we present a strong quasi-experimental design, it is not without limitations. First, while we aim to obtain evidence for causal effects, we recognize that this robust study design is still susceptible to confounding by time varying factors and that some of the assumptions for the time series analysis may not have been fulfilled. Examples of time varying factors include recurrent staff turnovers and program-level changes to data documentation and tools. Second, we were not able to obtain granular monthly data, but were restricted to quarterly data which limits our understanding of the month-to-month variations in data quality. Third, the cross-sectional survey design, for understanding facility and EMR implementation correlates of data quality, is vulnerable to selection bias and residual confounding inherent to the design. We were also not able to understand the level of interoperability between EMR and DHIS2 and how providers navigated these technical issues to generate DHIS2 summaries. Given power calculations were not used to determine the sample size for the survey, the cross-sectional analysis may have been underpowered to detect all significant associations between facility or EMR implementation factors and data quality. Finally, this evaluation focused on HIV-specific aspects of DHIS2, thus findings cannot be generalized to other health service departments, and further research would be needed to assess effects beyond HIV service delivery.

CONCLUSION

We found no evidence to suggest that EMRs in HIV facilities in Kenya improved aggregated data quality as captured by the discrepancy score. This does not detract from the multiple mechanisms through which EMRs could have effects on quality of care and health outcomes. We found that facility type, facility volume, and EMR type were associated with DHIS2 data quality. These findings suggest that presence of EMRs alone is insufficient to achieve high DHIS2 data quality. We investigated how implementing EMRs affects data quality in the routine aggregate data reporting system, thereby studying whether implementing a electronic system for client level data has a “spread effect” on DQ in other parts of the HIS ecosystem. A lot of digital health investments are relying on the premise that if we improve data systems at individual level, then these data can be leveraged for all types of public health purposes, including disease surveillance, public health reporting (as happens through DHIS2), and others. These negative findings are really important at demonstrating that we have to invest in not only the facility-level data systems, but also in automated linkages between the facility level and aggregate data systems in order for EMR implementation to impact data quality.

AIM 2 APPENDIX

Development of composite discrepancy (plausibility) score using z-scores

This approach utilized *Z*-scores to estimate the extent of deviation or discrepancy of observed values i.e., how far observed values were from expected values. To allow for comparison of values across facilities with varying patient volumes, we standardized the data checks by dividing them by respective facility patient volumes (i.e. general HIV care and treatment data checks were divided by the "total number of patients on ART", and ANC and PMTCT data checks were divided by the "total number of women accessing antenatal HIV testing"). Due to heterogeneity across facilities over time, values were standardized using *Z*-score transformations so that a common scale was used to compare degree of discrepancy [59]. Sites with greater discrepancy of observed values were penalized more than those with less discrepancy. The untransformed unadjusted Z-score was defined as:

$$z = (y - 0)/s_D$$

where y is the observed 'standardized data check value', 0 is the expected value if there is no difference between indicators, and S_D is the standard deviation of y. If the observed data check values are centered around 0, then z has mean 0 and SD 1. We assumed a normal distribution of data check values centered around the difference value, 0, the threshold for discrepancy (plausibility) and good care. An individual discrepancy (plausibility) *Z*-score was computed for each data check and the S_D was based on data from all facilities and quarters.

The composite discrepancy (plausibility) score was computed as an average of all the individual *Z*-scores for each unique facility quarter. Missing data check values were dropped from Z-score computation because the goal was to assess discrepancy (plausibility) of available values regardless of completeness. The discrepancy (plausibility) score was a continuous variable, centered on zero. For data check # 1-7, composite scores >= 0 were considered plausible and accurate regardless of how large or small the positive values were, thus all positive scores (> 0) were assigned a *Z*-score of 0. For data check # 8, values > 0 and values < 0 were both considered implausible. For consistency of handling deviations in the negative and positive direction, all negative individual *Z*-scores were multiplied by a factor of -1 and thus converted to the positive scale. As such, the lowest possible *Z*-score value was "0". Increasing *Z*-scores meant greater discrepancy or deviation of observed values, thus poorer data quality with regard to discrepancy (plausibility). The lower *Z*-scores were associated with less discrepancy, thus better data or greater discrepancy (plausibility). This approach was less stringent than the consistency score as discrepant values which did not considerably deviate away from the expected value were not excessively penalized compared to larger negative values or outliers.

Appendix 2 table: Individual data checks and composite discrepancy scores[80]

Data check and related indicators	**Expected value**	**Discrepancy score †**
Data check 1. Number of mothers tested for HIV at ANC **minus** number of newly HIV positive mothers at ANC	Any value >=0	Z-score based on extent of deviation of observed value from expected. Increasing score means higher deviation/discrepancy thus worse data quality
Data check 2. Number of mothers tested for HIV at (L&D and postnatal ward) **minus** number of newly HIV positive mothers		
Data check 3. Number of HIV-positive mothers at ANC **minus** number of mothers receiving Infant prophylaxis at ANC		
Data check 4. Number of HIV-positive mothers at (L&D and postnatal ward) **minus** HIV-positive mothers receiving Infant prophylaxis		
Data check 5. Cumulative (total ever) number of patients in HIV care **minus** Cumulative (total ever) number of patients on ART		
Data check 6. Cumulative (total ever) number of patients on ART **minus** total number currently on ART		
Data check 7. Current number of patients in HIV care **minus** current number on ART		
Data check 8. Total number of patients currently on ART **minus** sum of patients on ART across all age groups	Must be equal to zero (0)	Same as above
* Continuous score – percent complete vs percent consistent † Refers to plausibility ANC- Antenatal; L&D- Labor & Delivery, ART – Antiretroviral therapy		

Development of the composite discrepancy score

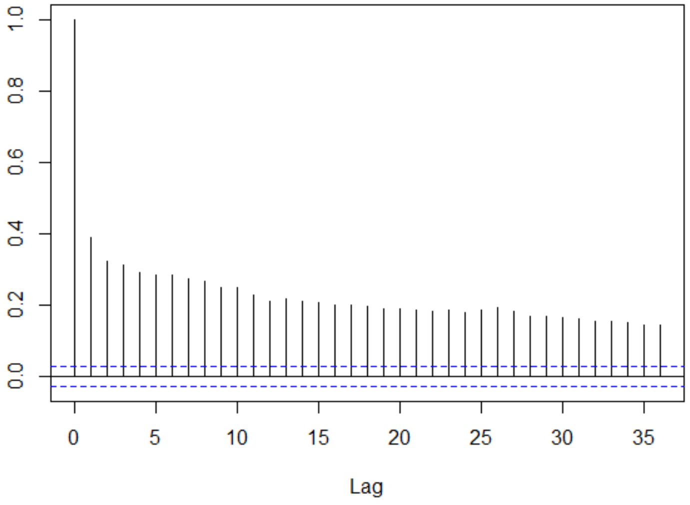
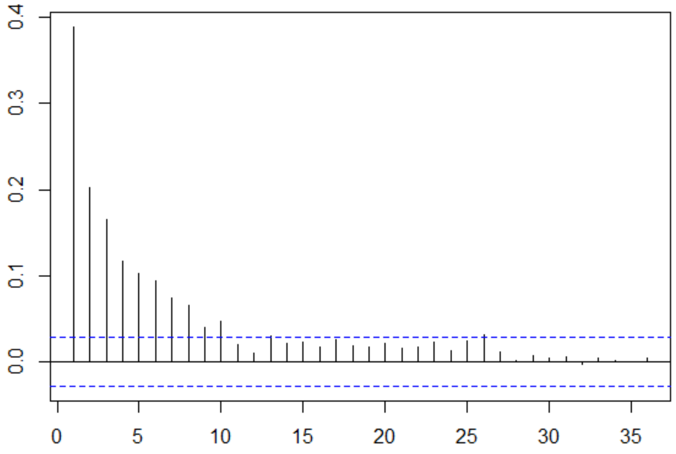
This approach utilized *Z*-scores to estimate the extent of deviation or discrepancy of observed values i.e., how far observed values were from expected values. To allow for comparison of values across facilities with varying patient volumes, we standardized the data checks by dividing them by respective facility patient volumes (i.e. general HIV care and treatment data checks were divided by the "total number of patients on ART", and ANC and PMTCT data checks were divided by the "total number of women accessing antenatal HIV testing"). Due to heterogeneity across facilities over time, values were standardized using *Z*-score transformations in order to objectively compare degree of discrepancy. Sites with greater discrepancy of observed values were penalized more than those with less discrepancy.

An individual discrepancy *Z*-score was computed for each data check. The composite discrepancy score was computed as an average of all the individual *Z*-scores for each unique facility quarter. The discrepancy score was a continuous variable. For data check # 1-7, composite scores >= 0 were considered plausible and accurate regardless of how large or small the positive values were, thus all positive scores (> 0) were assigned a *Z*-score of 0. This capping controlled large positive values such that they did not influence the average composite scores unduly. For data check # 8, values > 0 and values < 0 were both considered implausible. For uniform handling of deviations, all negative individual *Z*-scores were multiplied by a factor of -1 and thus converted to the positive scale and the lowest possible *Z*-score value was "0". Increasing *Z*-scores meant greater discrepancy or deviation of observed values, thus poorer data quality. Lower *Z*-scores were associated with less discrepancy, thus better data quality.

Autocorrelation plots

ACF and PACF plots revealed the time autocorrelation was insignificant thus, this was not adjusted for in the multilevel models. The ACF plot showed lingering autocorrelation for all observations and timepoints. The PACF plot showed that most of the autocorrelation is on lag 1, however, the correlation coefficient is ≤0.2 for subsequent observations, thus not significant (Figure 2).

Assessment of autocorrelation

Autocorrelation function (ACF) plot	Partial Autocorrelation function (ACF) plot
	

Chapter 4: Comparative analysis of HIV data completeness in Haiti's Electronic Medical Record System across children, adolescents, and adults: A retrospective evaluation

ABSTRACT

Background

We sought to assess whether there were disparities in completeness of documentation of HIV-related variables by age group in Haiti's Electronic Medical Record systems (EMR) system.

Methods

This was a retrospective study of data completeness from 58 facilities in Haiti's iSanté EMR system. We assessed 36,965 enrollment visits and 123,608 return visits from June 2016 to March 2022. We preselected routinely collected variables important for clinical management including sex, weight, height, WHO stage, TB screening, cotrimoxazole and isoniazid prophylaxis, adherence assessment, and first viral load. Completeness of individual variables was assessed. These variables were used to develop Composite Completeness Scores (CCS scores) as a measuer of data completeness. Generalized multivariable linear models were used to investigate patient and facility factors associated with CCS scores and timely first viral load documentation.

Results

At the enrollment visit, a majority of patients were adults (81.6%) and female (56.7%). Most of the facilities were health centers (75.9%). The overall mean enrollment visit CCS (completeness) score was 47.0% - lower CCS scores mean lower data completeness. The overall mean return visit CCS score was 39.6%. At enrollment, in multivariable analyses, being a child (-5.9 %, 95% CI: 6.36 – -5.43, p <0.001) and newer/ recent year of enrollment (-0.55%, 95% CI: -0.65 – -0.46, p <0.001), were significantly associated with lower data completeness; while being male (0.76%, 95 CI: 0.43 – 1.09, p <0.001) was significantly associated with higher data completeness. At return visit, both children (-8.43%, 95% CI: -8.70 – -8.17, p <0.001) and adolescents (-1.10%, 95% CI: -1.55 – -0.66, p <0.001) had significantly lower completeness scores than adults. Newer enrollment (i.e., more recent year of enrollment) was associated with lower data completeness (-0.41%, 95% CI: -0.47 – -0.35%, p <0.001). Male patients had significantly higher completeness scores

compared to women (0.50, 95% CI: 0.33 – 0.68, p <0.001). For timely first VL documentation, children had 80% lower odds of completeness compared to adults (AOR = 0.21, 95% CI: 0.16 – 0.28, p<0.001). A more recent year of enrollment was associated with 24% higher odds of first VL documentation completeness compared to the preceding year (AOR = 1.204 95% CI: 1.21 – 1.28, p<0.001). Each additional year of EMR use (i.e., duration of EMR use) was associated with 21% higher odds of first VL documentation completeness (AOR = 1.21, 95% CI: 1.09 – 1.34, <0.001).

Conclusions

Documentation of important clinical data was poor particularly among children at both enrollment and return visits. Adolescents had comparable data completeness scores to adults at return visits. Concerningly, data completeness declined overtime and among newer patient enrollments for all variables except first viral load documentation. We demonstrate that inequities can undermine the benefits of digital health technologies. Further evaluation and health provider support are needed to close the documentation gap and ensure complete data for comprehensive HIV management, especially among children, adolescents, and new patients in expanding HIV programs.

INTRODUCTION

Routine health information systems, including EMRs, can play an important role in improving patient health outcomes by improving the quality and availability of clinical data used in patient management and by reinforcing compliance with care guidelines [62]. In HIV programs, a notable challenge over the past decade has been the disparity in access to HIV treatment and health outcomes among children, adolescents, and adults. In 2021, globally, 52% of children living with HIV (CLHIV) had access to HIV treatment compared to 75% of people living with HIV (PLHIV) – a gap that continues to widen [99]. Identification of HIV infection status, delayed treatment initiation, low engagement and suboptimal retention in care are more challenging in pediatric and adolescent populations [100]. HIV program data is crucial for disease surveillance, monitoring quality of care, and timely intervention to address adverse patient-level and program-level outcomes [101]. It has been shown that poor data quality negatively impacts clinical decision making and is associated with poor patient outcomes including unsatisfactory retention in care [101].

Haiti is the highest HIV-burden country among the Caribbean countries with total of 150,000 [130,000 – 170,000] PLHIV and a HIV prevalence of 1.8% in adults aged 15-49 years; and 5,700 [5,000 – 6,700] CLHIV aged <15 years [102]. In 2016, eighty five percent (85%) of PLHIV were on antiretroviral therapy (ART) [102], and approximately 48% of PLHIV had suppressed viral loads [103]. In a review of retention in the HIV care and treatment cascade in Haiti as of 2019, it was noted that only 84% of PLHIV were aware of their HIV status, 86% of them were on treatment, 74% of the treated had a viral load, and 80% among those with a VL were virally suppressed (i.e. 48% of all PLHIV) [103]. In Haiti, children are less likely than adults to receive and adhere to life-saving HIV treatment with only 52% of children receiving treatment in 2020 compared to 86% of adults [3]. Furthermore, children have significantly lower retention and VL suppression rates than adults (59% vs 80% respectively) [103,104]. The highest losses to follow up were observed among adults aged 20-39-years and among children <10 years of age.

In 2016, along with other high HIV burden countries, Haiti adopted the “Test and Start” strategy for ART, which dramatically increased the number of individuals eligible for and initiated on ART [105]. Alongside this was the integration of routine VL testing which was introduced as part of the package of care for people living with HIV. Ongoing in-country efforts by the MOH to improve same day initiation of ART and routine VL testing are significant programmatic strides toward achieving the 95-95-95 HIV targets. In response to the growing national HIV response and the need for improved data quality for disease monitoring and surveillance, the iSanté EMR system was developed and rolled out in 2005. This large-scale, multi-site EMR was developed by the Haitian Ministry of Health (MSPP) and CDC in conjunction with I-TECH [23]. As of 2018, 126 facilities were using iSanté and this number grew to 157 facilities in 2022. The systems houses > 1,050,173 unique total patient records and 144,319 records of PLHIV [106].

Despite these great strides, suboptimal engagement of children and adolescents compared to adults raises concern as to whether these disparities are also reflected in data quality among other factors that impact quality of care. There is a paucity of research comparing disparities in data quality by age group, specifically completeness of documented HIV care processes among children, adolescents, and adults in low-and middle-income countries (LMICs like Haiti. Our aim is to apply a composite data quality score based on completeness of preselected clinical variables and use this system to profile individual characteristics and facility structural factors associated with data completeness in Haiti's iSanté EMR system. We aim to investigate whether there are differences in completeness of HIV-related data across the different age groups and hypothesize that there will be disparities across the different age groups.

METHODS

Study design: This was a retrospective study of data completeness in Haiti's iSanté EMR system.

Sample: 58 facilities, spread across various regions/ departments in Haiti, that upload data to iSanté EMR and the central server were included in this analysis. Fifty-eight facilities provided enrollment visit data while 57 facilities provided return visit data. All facilities used the EMR system; however, they had different lengths of time with EMR implementation. Data from June 2016 to March 2022 were included in the analysis. Data extraction was restricted to start from June 2016 because this marks the landmark roll out of the Test-and-Treat model in which all diagnosed patients received ART, and VL testing became the gold standard of monitoring HIV treatment. These facilities are primarily public clinics supported by non-profit organizations to provide HIV care. Facilities that did not use the iSanté EMR system were excluded. Program data from children (0-9 years), adolescents (10-19 years), and adults (≥20 years) enrolled in HIV care since 2016 were extracted. Observations without date of birth (<1%) were excluded.

Setting: With regard to clinic structure / workflow, most facilities integrate children, adolescent, and/or adult visits, without separate clinic days or clinics, but a few clinics have special adolescent clinics. According to national HIV treatment guidelines, once patients initiate ART, children are seen within one month of initiation, thereafter they are reviewed every 3 months and the initial viral load (VL) test is done after 3 months. Adult patients are seen within one month of ART initiation, then after 3 months for VL testing. If VL is undetectable, patients can return to clinic after 3 to 12 months based on the differentiated care model. Adolescent patients <15 years follow the pediatric visit schedule, whereas those >15 years old follow the adult visit schedule. Patients receive the full supply of ART that will last them the until their subsequent visits e.g., 3-, 6-, or 12-month supplies, thus there are no refill/ pharmacy-only visits. In between visits, patients receive home visits or check-in phone calls. Home visits are captured in the EMR as a clinical encounter, but phone calls are not.

Pre-selected variables that are needed for clinical care were examined for completeness. We assessed data on these variables that should be documented at initial versus return visits for all pediatric, adolescent, and adult patients. A uniform list of variables was selected for all age categories. The following is an outline of the variables of interest:

- Initial enrollment visit:
    - Sex
    - Weight
    - Height
    - TB screening
    - WHO stage
    - Cotrimoxazole prophylaxis administration
    - Isoniazid prophylaxis
- Return visit:
    - Weight
    - Height
    - Adherence assessment

  - WHO stage
  - TB screening
  - Isoniazid administration
  - Cotrimoxazole administration
  - First viral load at 3-9 months

All variables of interest were assessed for completeness of documentation and contributed equal weight to the completeness score (see below for further description) for data collected at the initial enrollment visit ("enrollment completeness") and return visits ("return completeness"). Data from home visits were deleted as these are documented differently and facilities lack consistency in the handling of these data. Specifically, there are elements of the home visit that are not documented and only a few variables to do with dispensation of ART are documented. Including home visits would result in misleadingly higher levels of missing data.

Exposure variables are age group and facility characteristics such as ownership (private or public), level (dispensary, health center or hospital), volume or number of patients, and duration of EMR implementation. Primary outcome variables are the Composite Completeness Score (CCS) at enrollment and return visit, and first viral load completeness.

Assessment of completeness of individual variables

Enrollment completeness assessment: The "enrollment completeness" percent for each clinical variable, was computed using "the number of patients with complete data on that variable" as the numerator, and "the unique number of patients" as the denominator multiplied by 100. The CCS score for each patient-visit was computed as the % completeness of the clinical variables.

Return completeness assessment: Individual observations consisted of unique clinic visits. To compute the "return completeness" percent for each variable, the numerator was "the number of visits with the complete data on a variable" and the denominator was "the total number of return visits" multiplied by 100. The CCS score for each patient-visit was computed as above.

Timely first Viral Load (VL) completeness: Completeness of the first viral load test variable was modelled as binary (1/0) outcome based on available documentation of a VL result among patients eligible for their first VL at a return visit, i.e., 6 -12 months after ART initiation. Eligible patients with the first VL load documented within 12 months of enrollment were assigned "1", while those missing a documented result were assigned “0”. This 12-month window allowed for capturing of VL results whose documentation may have been delayed.

Statistical Analysis

We used descriptive statistics to analyze the distribution of CCS scores and individual patient and facility characteristics. We summarized continuous data such as CCS scores as mean and standard deviation (SD), while we used proportions to summarize binary and categorical data such as variable completeness, age groups, facility type and facility volume. We assessed percent completeness across the various indicator variables by age group and overall. We used chi-squared test with continuity correction to compare the proportions of completeness among; a) children vs adolescents, b) children vs adults, and c) adolescents vs adults.

We assessed both individual-level enrollment and return visit CCS scores, and first VL test completeness. We calculated average CCS scores by age group (children, adolescents, and adults) and compared them with cluster-adjusted t-test (using linear regression with cluster robust standard errors) i.e., children vs adolescents, children vs adults, and adolescents vs adults.

To explore factors associated with enrollment, return visit CCS scores, and timely first VL documentation, all the patient-level and facility-level variables were included in multilevel multivariable linear mixed models. We employed a generalized linear mixed models (GLMM) with Gaussian distribution and identity link, with Restricted Maximum Likelihood variance estimator for continuous outcomes i.e., enrollment and return visit CCS scores. Specifically, for timely first VL documentation (binary outcome), we used a GLMM with binomial distribution and logit link. For enrollment and first viral load test completeness, we accounted for intra-cluster correlation using random intercepts for facilities. For return visit completeness, we accounted for intra-cluster correlation using random intercepts for patients and facilities.

All statistical analyses for this evaluation were done using R studio version 3.6.2 (2019-12-12).

RESULTS

A total of 36,965 patients from 58 facilities were captured for enrollment visits. Among these facilities, 57 facilities contributed return visit data for analysis. At enrollment, the median age was 32 years. Fourteen percent (14.6%) of the patients were children, 3.8% were adolescents, and 81.6%% were adults. The majority of patients were females (56.7%). Among facilities, 75.9%

were health centers (n=44), 19.0% were hospitals (n=11), and 5.1% were dispensaries (n=3). Most facilities were public (39.7%) or private (39.7%). Facilities had a median of 3,823 clients, and the median duration of EMR use was 15 years. The distribution of patient and facility characteristics was similar across both enrollment and return visit data (see table 4.1).

Table 4.1: Patient and facility characteristics across 58 facilities at enrollment and 57 facilities at return visits

Table 4.1 Baseline characteristics 1

Patient characteristics	**Enrollment visit N (%)/ Median (IQR) Total N= 36,965**	**Return visit N (%)/ Median (IQR) Total N=123,608**
Median age	32 (23-42)	34 (25-44)
Age groups		
Children (0-9)	5383 (14.6)	14827 (12.0)
Adolescent (10-19)	1407 (3.8)	4617 (3.7)
Adult (>20)	30175 (81.6)	104164 (84.3)
Sex		
Female	20963 (56.7)	73560 (59.5)
Male	14890 (40.3)	50006 (40.5)
Missing	1132 (3.0)	42 (0.0)
WHO stage		
Complete	2734 (7.4)	5405 (4.4)
Missing	34231 (92.6)	118203 (95.6)
	Facility N=58	**Facility N=57**
Facility level		
Hospital	11 (19.0)	11 (19.3)
Health center	44 (75.9)	43 (75.4)
Dispensary	3 (5.1)	3 (5.3)
Facility ownership		
Private	23 (39.7)	23 (40.3)
Public	23 (39.7)	22 (38.6)
Mixed	12 (20.6)	12 (21.1)
Total patient volume	3823 (787 – 16613)	3828 (795 – 17509)
Duration of EMR (years)	15 (11 – 17)	15 (11 – 17)

Enrollment visit completeness

All observations, regardless of age group, had >95% documentation of sex. Thirty one percent (31.9%) of all observations had documented weight with the lowest percentage documentation among children (30.0%) compared to adolescents (38.0%) and adults (32.0%). Twenty nine percent (29.1%) of observations had documented height with children having the lowest documentation percentage of 23.4% compared to adolescents (35.2%) and adults (29.8%). Regarding WHO stage, overall documentation was low at 7.4%; more specifically, 0.2%, 5.4%, and 8.8% of children, adolescents, and adults had documented WHO stage, respectively. Overall documentation of TB screening and isoniazid prophylaxis documentation was similar at 62.0%; for both variables, documentation was similar across all age groups at 61.0%, 63.6% and 62.1% for children, adolescents, and adults, respectively. Overall complete documentation for cotrimoxazole prophylaxis was 17.7% with 9.4%, 18.5%, and 19.2% complete documentation among children, adolescents, and adults, respectively (see table 4.2).

Table 4.2: Completeness proportions for variables at enrollment visit, overall and by age group, across 58 HIV clinics in Haiti (N= 36,965)

Table 4.2 Enrollment variables 1

Documented variables	**All patients Completeness N/ Total N=36,965 N (%)**	**Children (0-9) Completeness N/ Total N=5,383 N (%)**	**Adolescents (10-19) Completeness N/ Total N=1,408 N (%)**	**Adults (>20) Completeness N/ Total=30,174 N (%)**	**Estimate (95% CI, p-value) *** [children vs adolescents] [children vs adults] [adolescent vs adults]
Enrollment visit					
Sex	35833/ 36965 (97.0)	5249/ 5383 (97.6)	1405/ 1408 (99.8)	29179/ 30174 (96.7)	-2.3 (-2.8 – -1.8, p**<0.001)** 0.8 (0.3 – 1.3, p**=0.002**) 3.1 (2.7 – 3.4, p**<0.001**)
Weight	11789/ 36965 (31.9)	1610/ 5383 (29.9)	535/ 1408 (38.0)	9644/30174 (32.0)	-8.1 (-10.9 – -5.2, **p<0.001**) -2.1 (-3.4 – -0.7, **p=0.003**) 6.0 (3.4 – 8.7, **p<0.001**)
Height	10751/ 36965 (29.1)	1261/ 5383 (23.4)	496/ 1408 (35.2)	8994/ 30174 (29.8)	-11.8 (-14.6 – -9.0**, p<0.001)** **-**6.4 (-7.6 – -5.1, **p<0.001**) 5.4 (2.8 – 8.0**, p<0.001)**
WHO stage	2734/36965 (7.4)	10/ 5383 (0.2)	76/ 1408 (5.4)	2648/ 30174 (8.8)	-5.2 (-6.4 – -4.0**, p<0.001)** -8.6 (-8.9 – -8.2, **p<0.001)** -3.4 (-4.6 – -2.1, **p<0.001)**
TB screening	22913/ 36965 (62.0)	3282/ 5383 (61.0)	896/ 1408 (63.6)	18735/ 30174 (62.1)	-2.7 (-5.5 – 0.2, p=0.072) -1.1 (-2.5 – 0.3, p=0.123) 1.5 (1.1 – 4.2, p=0.254)
Isoniazid Prophylaxis	22913/ 36965 (62.0)	3282/ 5383 (61.0)	896/ 1408 (63.6)	18735/ 30174 (62.1)	-2.7 (-5.5 – 0.2, p=0.072) -1.1 (-2.5 – 0.3, p=0.123) 1.5 (1.1 – 4.2, p=0.254)
Cotrimoxazole prophylaxis	14586/ 36965 (39.5)	1138/ 5383 (21.1)	634/ 1408 (45.0)	12814/ 30174 (42.5)	-23.9 (-26.8 – -21.0**, p<0.001)** -21.3 (-22.6 – -20.1**, p<0.001)** 2.6 (-0.1 – 5.3, p=0.061)
* estimate, 95% Confidence Interval (CI) , and p-value of the chi-squared tests					

Composite completeness score at enrollment

The overall mean CCS score was 47.0% for enrollment visits. The mean CCS score was 42.0% for children, 50.1% for adolescents, and 47.7% for adults (see figure 1). After adjusting for clustering, adolescents had significantly higher average CCS score compared to children (difference = 8.1%, 95% CI: 2.2 – 14.0, p= 0.007)]; but no significant difference in average CCS score compared to adults (difference = 2.4%, 95% CI: -0.5 – 5.4, p <0.110); and adults had notably higher average CCS scores compared to children though this was not statistically significant (difference = 5.7%, 95% CI: -0.5 – 11.8, p=0.070).

Multivariable analyses of enrollment CCS scores

Table 4.3 provides the details of inferential analysis. In multivariable analysis that accounted for clustering at facility level, age group was significantly associated with CCS scores at the enrollment visit. Children had significantly lower CCS scores compared to adults i.e., 5.90 percentage points lower on average [-5.44 (95% CI: 6.36 – -5.43, p <0.001]. On average, male patients had CCS scores that were significantly higher than female patients, i.e., 0.43 percentage points higher (0.76, 95% CI: 0.43 – 1.09, p <0.001). The year of enrollment into care was also significantly associated with completeness of documentation, with each additional year being associated with a 0.55 percentage points lower score compared to the preceding year (-0.55, 95% CI: 0.65 – -0.46, p <0.001]. Compared to public facilities, facilities with mixed ownership (joint private and public ownership) had an average of 14.77 percentage points lower CCS scores (-14.77, 95% CI: -28.58 – -0.97), p=0.036) compared to public facilities. Facility level, total patient volume, and duration of EMR use were not associated with completeness of documentation at the enrollment visit.

Table 4.3: Patient and facility factors associated with Composite Completeness Score at enrollment, accounting for clustering by facility (N=36,965)

Table 4.3 Enrollment CCS score factors 1

	N (%)/ Median (IQR)/ Range	Unadjusted† β (95% CI)	P value	Adjusted* β (95% CI)	P value
Patient characteristics					
Age groups					
Adult	30175 (81.6)	Ref (baseline CCS: 46.8)		Ref	
Child	5383 (14.6)	-5.44 (-5.94 – -4.94)	**<0.001**	-5.90 (-6.38 – -5.43)	**<0.001**
Adolescent	1407 (3.8)	0.07 (-0.96 – 0.65)	0.884	-0.80 (-1.62 – 0.03)	0.058
Sex					
Female	20943 (56.7)	Ref (baseline CCS: 47.6)		Ref	
Male	14890 (40.3)	0.43 (0.10 – 0.76)	**0.011**	0.76 (0.43 – 1.09)	**<0.001**
Year of enrollment	2016-2022	-0.57 (-0.67 – -0.47)	**<0.001**	-0.55 (-0.65 – -0.46)	**<0.001**
Facility characteristics N=58 facilities					
Facility level					
Hospital	11	Ref (baseline CCS: 47.1)		Ref	
Health center	44	1.82 (-11.83 – 15.47)	0.794	-1.07 (-15.48 – 13.34)	0.884
Dispensary	3	-14.65 (-41.13 – 11.83)	0.278	-14.52 (-40.45 – 11.40)	0.272
Facility ownership					
Public	23	Ref (baseline CCS: 50.1)		Ref	
Private	23	1.94 (-9.57 – 13.44)	0.742	1.62 (-9.81 – 13.05)	0.781
Mixed	12	-15.01 (-28.90 – -1.11)	**0.034**	-14.77 (-28.58 – -0.97)	**0.036**
Total patient volume**	3823 (787-16613)	-0.02(-0.37 – 0.34)	0.923	0.06 (-3.40 – 3.51)	0.975
Duration of EMR (years) median (IQR)	15 (11-17)	-0.02 (-1.17 – 1.12)	0.968	0.03 (-1.08 – 1.14)	0.955

†Analysis accounting for clustering at facility level
*Multivariable models included all patient and facility-level characteristics and accounted for clustering.
IQR- interquartile range; CI – confidence interval
**Coefficients represent change in completeness for each 1000 person increase in facility volume.

Return visit completeness

Table 4.4 provides a summary of completeness of variables at the return visits. Overall, 123,608 return visits which included 14,827 pediatric return visits, 4,617 adolescent return visits, and 104,164 adult return visits were analyzed. For return visits, documentation of weight and height data at return visits was complete for 33.2% and 23.0% of the patients, respectively; children had the lowest percentage of complete weight and height documentation (28.2% and 22.6%, respectively) compared to adolescent patients (34.5% and 23.4%, respectively) and adults (33.9% and 23.0%, respectively). Overall completeness of weight was higher in return visits compared to enrollment visits (33.2 % vs 28.4%), but height data was less complete at return visits compared to enrollment visits (23.0% vs 25.9%). Similar to enrollment visits, data on WHO staging was poorly documented with completeness percent of 0.3%, 3.9%, and 5.0% among children, adolescents and adults, respectively. More than half (53.5%) of all return visits had documentation of ART adherence assessment; documentation of adherence was lowest among children (18.7%), higher among adolescents (57.8%), and highest among adults (58.3%). Among return visits, documentation of TB screening and isoniazid administration was 73.1% overall, and highest among children at 76.2%, while adolescents and adults had 70.1% and 72.7% completeness, respectively. This was higher than the 60.0% completeness of TB screening and isoniazid administration data at enrollment visits. Overall, complete documentation of cotrimoxazole was lower in return visits (16.6%) compared to documentation in enrollment visits (40.3%); lowest among children (12.3%), lower among adolescents (14.8%), and documentation was highest among adults (17.3%). Overall, timely documentation of the first viral load within 12 months of enrollment among patients eligible for VL testing was low at 17.0%; it was lowest among children (6.8%), but higher among adolescents (25.8%) and adults (21.6%) – table 4.4.

Table 4.4: Completeness proportions for variables at return visits, overall and by age group, across 57 HIV clinics in Haiti (N= 123,608)

Table 4.4 Return visit variables 1

Documented variables	**All patients Completeness N/ Total N=123,608 N (%)**	**Children (0-9) Completeness N/ Total N=14,827 N (%)**	**Adolescents (10-19) Completeness N/ Total N= 4,617 N (%)**	**Adults (>20) Completeness N/ Total N=104,164 N (%)**	**p-value (95% CI))** [children vs adolescents] [children vs adults] [adolescent vs adults]
Return visit					
Weight	41093/ 123608 (33.2)	4175/ 14827 (28.2)	1593/ 4617 (34.5)	35325/ 104164 (33.9)	-6.3 (-7.9 – -4.8**, p<0.001)** -5.8 (-6.5 – -5.0**, p<0.001)** 0.6 (-0.8 – 2.0, p=0.416)
Height	28408/ 123608 (23.0)	3350/ 14827 (22.6)	1080/ 4617 (23.4)	23978/ 104164 (23.0)	-0.8 (-02.2 – 0.6, p=0.268) -0.4 (-1.1 – 0.3, p=0.253) 0.4 (-0.9 – 1.6, p=0.569)
WHO stage	5405/ 123608 (4.4)	49/ 14827 (0.3)	179/ 4617 (3.9)	5177 / 104164 (5.0)	-3.6 (-4.1 – -3.0, **p<0.001)** -4.7 (-4.8 – -4.5, **p<0.001**) -1.1 (-1.7 – -0.5, **p=0.001**)
Adherence assessment	66183/ 123608 (53.5)	2777/ 14827 (18.7)	2667/ 4617 (57.8)	60739/ 104164 (58.3)	-39.1(-40.6 – -37.5, **p<0.001)** -39.6 (-40.3 – -38.9, **p<0.001)** -0.5 (-2.0 – 0.9, p=0.471)
TB screening	90306/ 123608 (73.1)	11297/ 14827 (76.2)	3235/ 4617 (70.1)	75774/ 104164 (72.7)	6.1 (4.6 – 7.6, **p<0.001)** 3.5 (2.7 – 4.2, **p<0.001)** -2.6 (-4.0 – -1.3, **p<0.001)**
Isoniazid Prophylaxis	90306/ 123608 (73.1)	11297/ 14827 (76.2)	3235/ 4617 (70.1)	75774/ 104164 (72.7)	6.1 (4.6 – 7.6, **p<0.001)** 3.5 (2.7 – 4.2, **p<0.001)** -2.6 (-4.0 – -1.3, **p<0.001)**
Cotrimoxazole prophylaxis	20580/ 123608 (16.6)	1829/ 14827 (12.3)	683/ 4617 (14.8)	18068/ 104164 (17.3)	-2.5 (-3.6 – -1.3, **p<0.001)** -5.0 (-5.6– -4.4, **p<0.001)** -2.5 (-3.6 – -1.5, **p<0.001)**
First viral load*	2426/ 14,291 (17.0%)	69/ 1016 (6.8)	78/ 302 (25.8)	2279/ 10547 (21.6)	-19.0 (-24.4 – -13.6, **p<0.001)** -14.8 (-16.6 – -13.0**, p<0.001)** 4.2 (-0.9 – 9.4, p=0.092)

*Timely documentation of the first viral load after VL test was ordered among eligible patients who have been in care for at least 6 months. Denominators are based on the number of individuals who were eligible for their first viral load.

The overall mean return visit CCS score was 39.6% (SD= 22.7); 33.5% for children, 39.2% for adolescents and 40.4% for adults. After adjusting for clustering, children had a significantly lower average CCS score compared to adolescents (difference = -5.7%, 95% CI: -10.6 – -0.8, p=0.023); and adults (difference = -6.9%, 95% CI: -11.9 – -2.0, p=0.006); and adolescents had lower average CCS score compared to adults (difference = -1.2%, 95% CI: -3.4 – -1.0, p <0.001).

Multivariable analysis of return visit CCS scores and first VL documentation

Table 4.5 provides the details of the analysis based on return visit CCS scores and first VL documentation. In the full multivariable model, both children and adolescents, on average, had significantly lower CCS scores than adults i.e., 8.43 points lower (-8.43, 95% CI: -8.70 – -8.17, p <0.001), and 1.10 percentage points lower (-1.10, 95% CI: -1.55 – -0.66, p <0.001), respectively. The year of enrollment into care was also significantly associated with completeness of documentation, with each additional year being associated with 0.41 percentage points lower CCS score compared to the preceding year (-0.41, 95% CI: -0.47 – -0.35, p <0.001). The aforementioned findings were similar to enrollment visit findings. On average, male patients had CCS scores that were significantly higher than women, i.e., 0.50 percentage points higher (0.50, 95% CI: 0.33 – 0.68, p <0.001). This association was also observed in the enrollment visit. Similar to the enrollment visit, facility level (i.e., hospital, health center, dispensary), facility ownership (public, private, mixed), total patient volume, and duration of EMR use were not significantly associated with completeness of documentation at the return visit., i.e., after including variables that were significant in bivariable analysis (age group, and year of enrollment).

With regard to timely first VL documentation completeness, in the full multivariable model, being a child was significantly associated with 80% lower odds in completeness of documentation compared to adults (adjusted OR (AOR) = 0.21, 95% CI: 0.16 – 0.28, p<0.001). A more recent year of enrollment was significantly associated with 24% higher odds of first VL documentation completeness compared to the preceding year (AOR = 1.204 95% CI: 1.21 – 1.28, p<0.001). Each additional year of EMR use (i.e., duration of EMR use) was also associated with 21% higher odds of first VL documentation completeness (AOR = 1.21, 95% CI: 1.09 – 1.34, <0.001). Sex, facility level, facility ownership, and total patient volume were not associated with complete first VL documentation. See table 4.5.

Table 4.5: Patient and facility factors associated with return visit Composite Completeness Score (CCS score) and first viral load completeness

Table 4.5 Return visit CCS score factors 1

		Return visit CCS scores				**First Viral load (VL) completeness**			
	N / Median (IQR)	**Unadjusted† β (95% CI)**	**P value**	**Adjusted* β (95% CI)**	**P value**	**Unadjusted OR (95% CI)**	**P value**	**Adjusted OR* (95% CI)**	**P value**
Patient characteristics (N=123,608)						**Patient N =14,291**			
Age groups									
Adult	104164	Ref (baseline CCS: 39.0)		Ref		Ref		Ref	
Child	14827	-8.35 (-8.62 – -8.09)	**<0.001**	-8.43 (-8.70 – -8.17)	**<0.001**	0.25 (0.19 – 0.32)	**<0.001**	0.21 (0.16 – 0.28)	**<0.001**
Adolescent	4617	-1.02 (-1.47 – -0.57)	**<0.001**	-1.10 (-1.55 – -0.66)	**<0.001**	1.20 (0.92 – 1.56)	0.171	1.10 (0.84 – 1.43)	0.496
Sex									
Female	73560	Ref (baseline CCS: 38.2)		Ref		Ref		Ref	
Male	50006	0.04 (-0.15 – 0.24)	0.679	0.50 (0.33 – 0.68)	**<0.001**	0.99 (0.91 – 1.10)	0.958	1.05 (0.96 – 1.16)	0.307
Year of enrollment⁑		-0.41 (-0.47 – -0.35)	<**0.001**	-0.46 (-0.52 – -0.40)	**<0.001**	1.20 (1.16 – 1.25)	**<0.001**	1.24 (1.21 – 1.28)	**<0.001**
Facility characteristics N=57 facilities									
Facility level									
Hospital	11	Ref (baseline CCS: 36.6)		Ref		Ref		Ref	
Health center	43	2.85 (-10.03 – 15.73)	0.665	2.19 (-10.69-15.07)	0.739	0.54 (0.19 – 1.57)	0.261	0.88 (0.32 – 2.45)	0.808
Dispensary	3	-8.85 (-33.87 – 16.18)	0.490	-9.06 (-34.07-15.95)	0.478	0.32 (0.03 – 2.98)	0.315	0.95 (0.12 – 6.82)	0.960
Facility ownership									
Public	22	Ref (baseline CCS: 39.1)		Ref		Ref		Ref	
Private	23	3.58 (-7.44 – 14.60)	0.525	3.08 (-7.95 – 14.11)	0.584	0.42 (0.17 – 1.06)	0.065	0.60 (0.23 – 1.59)	0.306
Mixed	12	-10.85 (-24.15 – 2.44)	0.110	-10.81(-24.12 - 2.50)	0.111	1.02 (0.35 – 3.03)	0.964	1.16 (0.42 – 3.17)	0.772

Total patient volume**	3828 (795 – 17509)	3.44 (-29.80 – 36.68)	0.839	4.94 (-2.83 – 3.81)	0.770	1.00 (0.99 -1.00)	0.283	1.00 (0.99 – 1.02)	0.963
Duration of EMR (years) median (IQR)	15 (11 – 17)	-0.09 (-1.22 – 1.03)	0.870	-0.06 (-1.19 – 1.06)	0.912	1.22 (1.11 – 1.35)	**<0.001**	1.21 (1.09 – 1.34)	**<0.001**

†Analysis accounting for clustering at facility level and patient level
*Fully adjusted multivariable models were used.
⁑ Year of enrollment from 2016-2022
OR: Odds ration; IQR- interquartile range; CI – confidence interval

DISCUSSION

In this retrospective study of EMR data from HIV clinics in Haiti, we found that overall documentation of care processes and clinical variables was less than 50% at both enrollment and return visits. Children had significantly lower data completeness in the EMR compared to adolescent and adults. Further, newer enrollments had less complete data, and overall data completeness declined; however, timely first viral load documentation completeness increased over time and with longer duration of EMR use. Important care processes and clinical variables like WHO staging, adherence assessment, cotrimoxazole prophylaxis, and first viral load results were poorly documented, especially in children when compared to adolescents and adults.

Factors associated with overall documentation completeness at enrollment and return visits were age, sex, and year of enrollment in HIV care; however, the difference in completeness scores between sexes was very small and may not be scientifically meaningful. Age, year of enrollment, and duration of EMR use were significantly associated with timely first VL completeness. Facility ownership was the only facility characteristic that was associated with the overall completeness score, albeit weakly. Duration of EMR use was also associated with first VL completeness but this could be due to maturation of the national VL program more than EMR use. Nevertheless, documentation completeness was lowest for the first VL and cotrimoxazole prophylaxis; however, this could be due to a variety of reasons beyond documentation, and it is difficult to parse out the factors resulting in this. More specifically for VL, we posit that, in addition to poor documentation, patient hesitancy to access VL tests, low retention in care, shortage of lab supplies, and long turnaround times for VL results could have contributed to low documentation.

To our knowledge, this is the first large-scale study to assess EMR data quality and compare documentation completeness across age groups in a low-resource national HIV program. A study of Haiti's EMR system from 2005-13, revealed higher, albeit variable levels of completeness across varied indicators such as height, TB status, and ART eligibility. The prior study demonstrated improved data quality overtime, unlike our present evaluation which only showed reducing data completeness over the years except for timely first VL documentation that[107]. The prior study demonstrated that facility factors (such as facility level, and maturity of EMR) were associated with data completeness – a finding that was not observed in our more recent evaluation except for facility ownership which was statistically significant at enrollment and had a large association with completeness (albeit not statistically significant) at return visit.

Other studies have assessed documentation completeness in different domains of routine health programs in LMICs such as TB and HIV care, and prevention of mother-to-child transmission of HIV (PMTCT), and national Health Management Information Systems (HMIS), without comparing completeness across age groups. An evaluation of Rwanda's national HMIS assessed data completeness, measured as percentage of non-missing reports, and found increasing levels of data completeness overtime [46]. A cross-sectional study of an ART clinic in Ethiopia found 76% completeness of EMR data, which was slightly lower compared to paper-based data [37]. Conversely, our evaluation found decreasing levels of completeness with each increasing year when assessing CCS scores; however, the timely completeness of first viral load increased significantly with each additional year in our evaluation. This likely reflected the scale up of VL availability rather than improved documentation practices, though we can't disentangle this.

A study evaluating routine data quality in PMTCT programs across three large, high HIV-prevalence Districts in South Africa showed that data was neither complete nor accurate [108]. An evaluation of a TB Surveillance System which used Electronic TB Register in South Africa revealed low levels of completeness and reliability of data with likely implications on the notification and treatment of patients newly diagnosed with TB [109].

This present evaluation demonstrates that documentation continues to be a challenge within routine HMIS in resource-constrained settings; particularly, less complete documentation for children. Good data quality is associated with better health system performance and improved health outcomes [62]. Conversely, poor documentation and data quality will potentially be associated with suboptimal care and poor health outcomes among historically vulnerable groups like CLHIV [110]. For example, lower documentation of care processes e.g., weight, height, adherence, and WHO stage would result in inappropriate ART dosage, inability to detect failure to thrive, and delayed detection of poor adherence and drug resistance in children with declining clinical status. Poor tracking of care could inadvertently contribute to dismal clinical outcomes and low retention among CLHIV and thus fuel the vicious cycle of poor follow-up and retention in care [101,111].

Another finding of this evaluation is the deteriorating data completeness over time. Return visits and newer enrollments, by year, were associated with poorer data documentation. This could partly be due to the COVID-19 pandemic which spanned more than one third of the evaluation's time frame. The disruptions in service delivery may have contributed to documentation failure as demonstrated in other settings. A multi-country assessment found that completeness of data reporting on maternal and child health indicators reduced after the pandemic [112]. Other factors include increased documentation burden programs scale up – the numbers of patients increases and outpaces the capacity of HIV programs information systems[101,113]. This paradox of successful scale up and deteriorating data quality was described in Nigeria [113]. Additionally, a study of the SmartCare EMR system in Zambia showed that increasing data demands, staffing shortage and lack of supplies may have contributed to declines in data quality [114]. Despite declining documentation completeness for a majority of care processes, completeness of timely viral load documentation increased with time and with longer duration of EMR use. This could be as a result of maturation over time alongside mounting global and national investment in viral load monitoring, registries and reporting as a requirement for PEPFAR and other funding agencies [115].

Other programmatic modifications, such as differentiated service delivery and challenges scaling up monitoring and evaluation systems, could potentially contribute to lack of synchronization of data collected from clinic- and community-based encounters [116]. Further, staff turnover and lack of refresher training may contribute to documentation gaps in the EMR. All these are factors that need further exploration and comprehension. This declining trend in completeness has potential ramifications on the care of newly enrolled patients, who are more vulnerable to loss of follow-up and poor retention. Poor documentation from the outset puts patients at an initial disadvantage as the pattern is likely to be carried forward. Further, interventions that rely on complete and timely documentation may be administered late if not all. For example, lack of weight data for appropriate ART dosage in children and lack of WHO stage for timely clinical intervention and assessment of treatment failure among other concerns.

Missing data is one of the most prevalent challenges of routine HMIS [117,118]. The impact of incomplete data goes beyond compromising patient management - it undermines: 1) utility of data for research and evaluation of policy implementation; 2) upstream organizational and national operations required for forecasting programmatic needs; and 3) informed planning of the national – and global – HIV response [115]. National health systems and supporting partners should continue to put data quality at the center of health system strengthening efforts [119]. Various strategies to consider include use of more robust data quality evaluation designs, data quality audits, staff training and feedback, supportive supervision, and technical support including use of unique patient identifiers to allow for data linkages across routine data sets [28,54]. A study in Viet Nam that assessed the completeness of death records in the HMIS showed highly levels of data completeness which could have further been augmented by triangulation with data from various sources [120].

Our evaluation had a number of strengths including the use of large-scale routine clinical data that is widely spread geographically. As such, the findings are highly applicable and relevant to Haiti. This offers findings that could potentially prompt change in the HIV program and HMIS. The methods used in this evaluation may also be generalizable to other similar settings and may shed light on documentation challenges in other low resource settings, particularly in children and adolescents. With regard to limitations, the observational design renders this evaluation vulnerable to confounding and causation cannot be inferred. Additionally, the proportion with complete documentation of the first VL is likely to be an underestimate because patients eligible for first VL who did not get the test will erroneously be classified as having incomplete data. Further VL specific analysis using mixed methods designs would be instrumental in understanding the VL test cascade and identify where the gaps occur between VL ordering to VL result documentation.

CONCLUSION

Overall, children had lower CCS scores compared to adolescents and adults, at both enrollment and return visits. Children were most likely to have poorly documented WHO staging, cotrimoxazole prophylaxis, and first viral load result documentation, thus potentially least likely to have optimal HIV management. As digital health interventions like EMR expand, the presence of digital health inequities should be anticipated and measured to ensure that no one left behind. Further evaluation and health provider support are needed to close the documentation gap and ensure complete data for comprehensive HIV management, especially among children, adolescents, and new patients in expanding HIV programs.

Chapter 5: Conclusions

SUMMARY OF FINDINGS

In this chapter 2 of this evaluation, we demonstrated that the DHIS2 is a convenient and readily available source of nationwide longitudinal HIV data for DQ assessment in Kenya. Data driven approaches to DQ assessment can be efficient and time-saving, thus potentially feasible to implement frequently. We successfully developed a data-driven approach for examining large volumes of data by computing three data-driven composite scores reflecting completeness, consistency, and discrepancy (plausibility).

These findings add to the base of evidence on national data quality in Kenya given they are uniquely based on a large, longitudinal sample and thus offers a broader snapshot of the status of data completeness in Kenya. Our evaluation builds on findings from studies in different settings which were disadvantaged by shorter follow-up periods and/or small sample sizes. We demonstrated that overall data completeness in Kenya was strikingly low with 40% missing data on average. The most problematic indicators were related to ANC- and L&D-related data which were largely incomplete – these contrasts several smaller scale studies which report higher levels of data completeness in SSA.

In contrast to data completeness scores, facilities had better accuracy, as captured by consistency and discrepancy scores – a finding that was consistent with findings from other studies. After stratifying the data, we found that general HIV care data had higher levels of discrepancy compared to ANC- and L&D-related data. In-depth exploration of the disparate performance across service departments and specific indicators is needed to understand context and barriers to tailor appropriate interventions. The DHIS2 can regularly be assessed using these efficient approaches to determine the quality of its data and ensure it fit for use by national policy makers.

In chapter 3, we assess the impact of EMR on DHIS2 data quality. There is existing research describing data quality within electronic data systems, but less research examining the evidence for causal effects of EMR implementation on quality of aggregate data reported from health facilities to the national HMIS. Our interrupted time series analysis showed no evidence that transitioning to EMR implementation was associated with DHIS2 data quality as captured by the discrepancy score. This could be attributed to the low degree of discrepancy at baseline (i.e., good or favorable). These null findings suggest that the presence of EMRs alone is insufficient to achieve strong aggregate data quality in the national HMIS. This could be an indication that further optimization of EMR usage is needed to improve data quality within the EMR (which we did not directly study), or that automated data exchange between the EMR and DHIS2 is necessary to realize better HMIS data quality, or both. Studies have shown several barriers to EMR implementation which may be associated with poor data quality including lack of provider training and failure to standardized EMR use such as dual or parallel documentation, EMR features, frequency of data quality assessments, and resource constraints. These factors need systematic exploration and design of remedial interventions to address context-specific challenges.

Our cross-sectional facility survey revealed facility and EMR implementation factors that were associated with data quality. Health centers and higher facility volumes were associated with higher discrepancy scores (thus worse data quality) compared to dispensaries. The type of EMR software used was associated with data quality. We postulate that potential disparities in usability of the software along with variations in the extent of technical support and funding provided by implementing partners may have contributed to this association. We found that there was a weak association between the data quality and the frequency of RDQAs; however, the mode of EMR implementation (i.e., whether point-of-care, retrospective or hybrid data entry) was not associated with data quality.

More work is needed to understand systemic deficits that hinder health provider acceptance, uptake, and meaningful use of EMRs. This work should be coupled with continuous quality improvement efforts and design of contextually appropriate implementation strategies to strengthen the uptake of EMRs. Ministries of Health need to further invest in standardized processes for supporting, monitoring, and evaluating EMRs regardless of the type of EMR software in use. Further consideration is needed to synergize EMR platforms to ensure interoperability between different software and minimize gaps in their performance. Strategies such as training health providers, infrastructure support, modifying electronic tools, designing usable and interoperable systems, data audits, and sustainable technical support from ministries of health are critical for optimal implementation of EMRs. Finally, evaluation of EMRs takes time, thus researchers, implementers and policy makers need to factor in longer timeframes for follow-up and concurrently focus on EMR usage, data quality, and identification of associated facility or EMR factors.

Chapter 4 describes our retrospective evaluation on data completeness in HIV clinics in Haiti. Our evaluation demonstrated that completeness of documentation of clinical variables and care processes was less than 50% at enrollment and return visits. Documentation of care processes and clinical variables like WHO staging, adherence assessment, cotrimoxazole prophylaxis, and first viral load results was assessed across age groups i.e., children, adolescents, and adults. To our knowledge, this is the first large-scale study to assess EMR data quality and compare documentation completeness across age groups in a low-resource national HIV program.

Children had significantly lower data completeness in the EMR compared to adolescents and adults. Further, newer enrollments had less complete data, and overall data completeness declined with time; however, timely first viral load documentation completeness increased over time and with longer duration of EMR use. Documentation completeness was lowest for the first VL and cotrimoxazole prophylaxis. Factors that were significantly associated with overall documentation completeness were age, sex, and year of enrollment in HIV care. Age, year of enrollment, and duration of EMR use were significantly associated with timely first VL documentation completeness. While facility ownership was the only facility characteristic that was associated with completeness, the association was weak.

Our evaluation in Haiti demonstrates that documentation continues to be a challenge within routine HMIS in resource-constrained settings; particularly, less complete documentation for children. The declining trend in completeness is concerning as it has potential ramifications on the care of newly enrolled patients who are potentially more vulnerable to loss of follow-up and poor retention. Poor documentation from the outset puts patients at an initial disadvantage as the pattern is likely to be carried forward and may delay timely clinical intervention. It is expected that digital health interventions would increase access to healthcare and improve health outcomes for all. However, the use of EMRs is still associated with inequities in data quality – an issue that needs further exploration and research.

Overall, missing data is one of the most prevalent challenges of routine HMIS in LMICs. The impact of incomplete data goes beyond compromising patient management - it undermines: 1) utility of data for research and evaluation of policy implementation; 2) upstream organizational and national operations required for forecasting programmatic needs; and 3) informed planning of the national – and global – HIV response. Various strategies to consider include the use of more robust data quality evaluation designs, data quality audits, staff training and feedback, supportive supervision, and technical support including use of unique patient identifiers to allow for data linkages across routine data sets. The findings from our evaluation can inspire more attention and investment in data quality and may light the path to effective expansion of EMR to comprehensive care – a crucial step to realizing universal health coverage.

IMPLICATIONS FOR RESEARCH

Further understanding of the mechanistic link between EMRs and data quality in HMIS will be necessary to integrate EMRs fully and meaningfully into routine healthcare settings in LMICs. More robust experimental designs that incorporate qualitative methodology are needed to explore reasons behind poor data quality and inequities. There is also an important role for implementation science research, to provide a framework for understanding barriers and facilitators of EMR use, the core components of EMR implementation, and how these can be adapted to suit different contexts in which EMRs are introduced, e.g., rural vs urban, private vs public health facilities. Further, mixed methods research is needed to characterize population groups that are more likely to have suboptimal data quality, the reasons for disparities in data quality, and investigation of strategies to address these disparities.

Acronyms

AIDS	Acquired immunodeficiency syndrome
ACF	Autocorrelation function
ANC	Antenatal care
AOR	Adjusted Odds Ratio
ART	Antiretroviral therapy
CCS	Composite completeness score
CI	Confidence interval
CDC	Center for Disease Control and Prevention
CLHIV	Children living with HIV
DHIS2	District Health Information System version 2
DQ	Data quality
DQA	Data quality assessment
EMR	Electronic Medical Record
ESRC	Ethical Scientific Review Committee
GLM	Generalized Linear Models
GLMM	Generalized Linear Mixed Models
HAART	Highly Active Antiretroviral therapy
HMIS	Health Management information system
HIV	Human immunodeficiency virus
ICC	Intracluster correlation
IQR	Interquartile range
L&D	Labor and Delivery care
LMIC	Low- and middle-income countries
MOH	Ministry of Health
MSPP	Haitian Ministry of Health
OR	Odds Ratio
PACF	Partial autocorrelation function
PEPFAR	President's Emergency Plan for AIDS Relief
PLHIV	people living with HIV
PMTCT	prevent mother-to-child transmission of HIV

PRISM	Performance of Routine Information System Management framework
RDQA	Routine data quality assessment
SD	Standard Deviation
SSA	sub-Saharan Africa
TB	Tuberculosis
UNAIDS	Joint United Nations Programme on HIV/AIDS
VL	Viral Load
WHO	World Health Organization

References

1. UNAIDS Fast-track: Ending the AIDS epidemic by 2030. https://www.pedaids.org/2014/11/20/unaids-issues-new-fast-track-strategy-to-end-aids-by-2030/.

2. WHO. Electronic health records: manual for developing countries. Manila: WHO Regional Office for the Western Pacific; 2006.

3. WHO. Global diffusion of eHealth: making universal health coverage achievable: report of the third global survey on eHealth. World Health Organization; 2017.

4. WHO. WHO guideline: recommendations on digital interventions for health system strengthening: web supplement 2: summary of findings and GRADE tables. World Health Organization; 2019.

5. WHO. WHO guideline: recommendations on digital interventions for health system strengthening: web supplement 2: summary of findings and GRADE tables. World Health Organization; 2019.

6. Kumar M, Mostafa J. Electronic health records for better health in the lower-and middle-income countries: a landscape study. Libr Hi Tech. 2020.

7. Muinga NA-O http orcid org, Magare SA-O http orcid org, Monda JA-O http orcid org, Kamau OA-O http orcid org, Houston SA-O http orcid org X, Fraser HA-O http orcid org, et al. Implementing an Open Source Electronic Health Record System in Kenyan Health Care Facilities: Case Study.

8. Suthar AB, Khalifa A, Joos O, Manders E-J, Abdul-Quader A, Amoyaw F, et al. National health information systems for achieving the Sustainable Development Goals. BMJ Open. 2019;9: e027689.

9. OpenMRS. Accessed May 16, 2020. Available: https://openmrs.org/about/

10. Amoroso CL, Akimana B, Wise B, Fraser HS. Using electronic medical records for HIV care in rural Rwanda. 2010. pp. 337–341.

11. Castelnuovo B, Kiragga A, Afayo V, Ncube M, Orama R, Magero S, et al. Implementation of provider-based electronic medical records and improvement of the quality of data in a large HIV program in Sub-Saharan Africa. Plos One. 2012;7: e51631. doi:10.1371/journal.pone.0051631

12. Chaplin B, Meloni S, Eisen G, Jolayemi T, Banigbe B, Adeola J, et al. Scale-up of networked HIV treatment in Nigeria: creation of an integrated electronic medical records system. Int J Med Inf. 2015;84: 58–68. doi:10.1016/j.ijmedinf.2014.09.006

13. Tweya H, Feldacker C, Gadabu OJ, Ng'ambi W, Mumba SL, Phiri D, et al. Developing a point-of-care electronic medical record system for TB/HIV co-infected patients: experiences from Lighthouse Trust, Lilongwe, Malawi. BMC Res Notes. 2016;9: 146. doi:10.1186/s13104-016-1943-4

14. Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy Plan. 2009;24: 217–28. doi:10.1093/heapol/czp010

15. RHINO. 2003. Enhancing quality and use of routine health information at district level. Second international workshop, September–October, Eastern Cape, South Africa. MEASURE Evaluation. 2003.

16. Sood SP, Nwabueze SN, Mbarika VW, Prakash N, Chatterjee S, Ray P, et al. Electronic medical records: A review comparing the challenges in developed and developing countries. IEEE; 2008. pp. 248–248.

17. Silvestre, E. and Wood, F. (2019), "Health information systems: analysis of country-level strategies,indicators and resources." Assessed May 16,2020. Available: https://www.measureevaluation.org/resources/tools/health-information-systems

18. Fritz F, Tilahun B, Dugas M. Success criteria for electronic medical record implementations in low-resource settings: a systematic review. J Am Med Inform Assoc. 2015;22: 479–488.

19. Kumar M, Mostafa J. Research evidence on strategies enabling integration of electronic health records in the health care systems of low-and middle-income countries: A literature review. Int J Health Plann Manage. 2019;34: e1016–e1025.

20. Odekunle FF, Odekunle RO, Shankar S. Why sub-Saharan Africa lags in electronic health record adoption and possible strategies to increase its adoption in this region.

21. Shrivastava SR, Shrivastava PS. Employing digital technology to strengthen public health services in African region. Niger J Gen Pract. 2018;16: 36.

22. Kang'a SG, Muthee VM, Liku N, Too D, Puttkammer N. People, Process and Technology: Strategies for Assuring Sustainable Implementation of EMRs at Public-Sector Health Facilities in Kenya.

23. Lober WB, Quiles C, Wagner S, Cassagnol R, Lamothes R, Alexis DRP, et al. Three years experience with the implementation of a networked electronic medical record in Haiti. AMIA Annu Symp Proc AMIA Symp. 2008;2008: 434–438.

24. deRiel E, Puttkammer N, Hyppolite N, Diallo J, Wagner S, Honore JG, et al. Success factors for implementing and sustaining a mature electronic medical record in a low-resource setting: a case study of iSante in Haiti. Health Policy Plan. 2018;33: 237–246. doi:10.1093/heapol/czx171

25. Wagenaar BH, Sherr K, Fernandes Q, Wagenaar AC. Using routine health information systems for well-designed health evaluations in low-and middle-income countries. Health Policy Plan. 2016;31: 129–135.

26. WHO. Data quality review: a toolkit for facility data quality assessment. Module 1. Framework and metrics. Geneva: World Health Organization; 2017. Licence: CC BY-NC-SA 3.0 IGO. 2017. Available: https://www.who.int/healthinfo/tools_data_analysis/dqr_modules/en/ [Accessed March 2019]

27. Wagenaar BH, Gimbel S, Hoek R, Pfeiffer J, Michel C, Manuel JL, et al. Effects of a health information system data quality intervention on concordance in Mozambique: time-series analyses from 2009–2012. Popul Health Metr. 2015;13: 9.

28. Muthee V, Bochner AF, Osterman A, Liku N, Akhwale W, Kwach J, et al. The impact of routine data quality assessments on electronic medical record data quality in Kenya. Plos One. 2018;13: e0195362. doi:10.1371/journal.pone.0195362

29. Lee K, Weiskopf N, Pathak J. A framework for data quality assessment in clinical research datasets. American Medical Informatics Association; 2017. p. 1080.

30. 2025 AIDS targets -UNAIDS. https://www.unaids.org/sites/default/files/2025-AIDS-Targets_en.pdf. Accessed October 18, 2021.

31. Garrib A, Stoops N, McKenzie A, Dlamini L, Govender T, Rohde J, et al. An evaluation of the District Health Information System in rural South Africa. South Afr Med J Suid-Afr Tydskr Vir Geneeskd. 2008;98: 549–552.

32. Okello G, Gerrets R, Zakayo S, Molyneux S, Jones C. “Every day they keep adding new tools but they don’t take any away”: Producing indicators for intermittent preventive treatment for malaria in pregnancy (IPTp) from routine data in Kenya. Plos One. 2018;13: e0189699. doi:10.1371/journal.pone.0189699

33. Gloyd S, Wagenaar BH, Woelk GB, Kalibala S. Opportunities and challenges in conducting secondary analysis of HIV programmes using data from routine health information systems and personal health information. J Int AIDS Soc. 2016;19: 20847. doi:10.7448/IAS.19.5.20847

34. Ndira SP, Rosenberger KD, Wetter T. Assessment of data quality of and staff satisfaction with an electronic health record system in a developing country (Uganda): a qualitative and quantitative comparative study. Methods Inf Med. 2008;47: 489–498.

35. Numair T, Harrell DT, Huy NT, Nishimoto F, Muthiani Y, Nzou SM, et al. Barriers to the Digitization of Health Information: A Qualitative and Quantitative Study in Kenya and Lao PDR Using a Cloud-Based Maternal and Child Registration System. Int J Env Res Public Health. 2021;18. doi:10.3390/ijerph18126196

36. Health Information Systems Project: End of Project Evaluation Report for International Training and Evaluation Center for Health (I-TECH). 2016.

37. Abiy R, Gashu K, Asemaw T, Mitiku M, Fekadie B, Abebaw Z, et al. A Comparison of Electronic Medical Record Data to Paper Records in Antiretroviral Therapy Clinic in Ethiopia: What is affecting the Quality of the Data? Online J Public Health Inf. 2018;10: e212. doi:10.5210/ojphi.v10i2.8309

38. Kumar M, Mostafa J. Electronic health records for better health in the lower-and middle-income countries. Libr Hi Tech. 2020.

39. Suthar AB, Khalifa A, Joos O, Manders E-J, Abdul-Quader A, Amoyaw F, et al. National health information systems for achieving the Sustainable Development Goals. BMJ Open. 2019;9: e027689.

40. Nicol E, Hanmer LA. Routine Health Information Systems in South Africa--Opportunities for Improvement. Stud Health Technol Inform. 2015;216: 993.

41. WHO. Data quality review: a toolkit for facility data quality assessment: module 3: data verification and system assessment. 2017.

42. Kahn MG, Raebel MA, Glanz JM, Riedlinger K, Steiner JF. A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research. Med Care. 2012;50 Suppl: S21-29. doi:10.1097/MLR.0b013e318257dd67

43. Mphatswe W, Mate KS, Bennett B, Ngidi H, Reddy J, Barker PM, et al. Improving public health information: a data quality intervention in KwaZulu-Natal, South Africa. Bull World Health Organ. 2012;90: 176–182.

44. Ministry of Health, Zambia. Data Quality Audit (DQA) Guidelines (Health Management Information System). 2014. Available: https://pdf.usaid.gov/pdf_docs/PA00K5N6.pdf

45. Braa J, Heywood A, Sahay S. Improving quality and use of data through data-use workshops: Zanzibar, United Republic of Tanzania. Bull World Health Organ. 2012;90: 379–384.

46. Nisingizwe MP, Iyer HS, Gashayija M, Hirschhorn LR, Amoroso C, Wilson R, et al. Toward utilization of data for program management and evaluation: quality assessment of five years of health management information system data in Rwanda. Glob Health Action. 2014;7: 25829.

47. Nagbe T, Yealue K, Yeabah T, Rude JM, Fallah M, Skrip L, et al. Integrated disease surveillance and response implementation in Liberia, findings from a data quality audit, 2017. Pan Afr Med J. 2019;33: 10. doi:10.11604/pamj.supp.2019.33.2.17608

48. Marchant T, Schellenberg J, Peterson S, Manzi F, Waiswa P, Hanson C, et al. The use of continuous surveys to generate and continuously report high quality timely maternal and newborn health data at the district level in Tanzania and Uganda. Implement Sci. 2014;9: 112.

49. Rowe AK. Potential of integrated continuous surveys and quality management to support monitoring, evaluation, and the scale-up of health interventions in developing countries. Am J Trop Med Hyg. 2009;80: 971–979.

50. Chen H, Hailey D, Wang N, Yu P. A review of data quality assessment methods for public health information systems. Int J Environ Res Public Health. 2014;11: 5170–5207.

51. Endriyas M, Alano A, Mekonnen E, Ayele S, Kelaye T, Shiferaw M, et al. Understanding performance data: health management information system data accuracy in Southern Nations Nationalities and People's Region, Ethiopia. BMC Health Serv Res. 2019;19: 175. doi:10.1186/s12913-019-3991-7

52. WHO. WHO Data Quality Tool for DHIS2. Available: https://dhis2.org/who-dq/. Assessed February 5, 2022

53. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017;390: 1151–1210. doi:10.1016/s0140-6736(17)32152-9

54. O'Hagan R, Marx MA, Finnegan KE, Naphini P, Ng'ambi K, Laija K, et al. National assessment of data quality and associated systems-level factors in Malawi. Glob Health Sci Pract. 2017;5: 367–381.

55. Cohen JF, Coleman E, Kangethe MJ. An importance-performance analysis of hospital information system attributes: A nurses' perspective. Int J Med Inf. 2016;86: 82–90. doi:10.1016/j.ijmedinf.2015.10.010

56. Kenya AIDS response progress report. Ministry of Health, 2018. National AIDS Control Council. 2018.

57. Dehnavieh R, Haghdoost A, Khosravi A, Hoseinabadi F, Rahimi H, Poursheikhali A, et al. The District Health Information System (DHIS2): A literature review and meta-synthesis of its strengths and operational challenges based on the experiences of 11 countries. Health Inf Manag. 2018; 1833358318777713. doi:10.1177/1833358318777713

58. Braa J, Sahay S. The DHIS2 open source software platform: evolution over time and space. LF Celi Glob Health Inform. 2017;451.

59. Spiegelhalter D, Sherlaw-Johnson C, Bardsley M, Blunt I, Wood C, Grigg O. Statistical methods for healthcare regulation: rating, screening and surveillance. J R Stat Soc Ser A Stat Soc. 2012;175: 1–47.

60. Nicol E, Dudley L, Bradshaw D. Assessing the quality of routine data for the prevention of mother-to-child transmission of HIV: An analytical observational study in two health districts with high HIV prevalence in South Africa. Int J Med Inf. 2016;95: 60–70. doi:10.1016/j.ijmedinf.2016.09.006

61. Jamieson L, Evans D, Berhanu R, Ismail N, Aucock S, Wallengren K, et al. Data quality of drug-resistant tuberculosis and antiretroviral therapy electronic registers in South Africa. BMC Public Health. 2019;19: 1638. doi:10.1186/s12889-019-7965-9

62. Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy Plan. 2009;24: 217–28. doi:10.1093/heapol/czp010

63. Wagenaar BH, Sherr K, Fernandes Q, Wagenaar AC. Using routine health information systems for well-designed health evaluations in low-and middle-income countries. Health Policy Plan. 2016;31: 129–135.

64. Leon N, Balakrishna Y, Hohlfeld A, Odendaal WA, Schmidt B-M, Zweigenthal V, et al. Routine Health Information System (RHIS) improvements for strengthened health system management. Cochrane Database Syst Rev. 2020;8: CD012012. doi:10.1002/14651858.CD012012.pub2

65. Jacob N, Rice B, Kalk E, Heekes A, Morgan J, Brinkmann S, et al. Consolidating strategic information to monitor progress against the UNAIDS 90-90-90 targets: evaluating the operational feasibility of an electronic HIV testing register in Cape Town, South Africa. BMC Health Serv Res. 2020;20: 720. doi:10.1186/s12913-020-05517-7

66. Ley B, Rijal KR, Marfurt J, Adhikari NR, Banjara MR, Shrestha UT, et al. Analysis of erroneous data entries in paper based and electronic data collection. BMC Res Notes. 2019;12: 537. doi:10.1186/s13104-019-4574-8

67. McLean E, Dube A, Saul J, Branson K, Luhanga M, Mwiba O, et al. Implementing electronic data capture at a well-established health and demographic surveillance site in rural northern Malawi. Glob Health Action. 2017;10: 1367162. doi:10.1080/16549716.2017.1367162

68. Mukasa O, Mushi HP, Maire N, Ross A, de Savigny D. Do surveys with paper and electronic devices differ in quality and cost? Experience from the Rufiji Health and demographic surveillance system in Tanzania. Glob Health Action. 2017;10: 1387984. doi:10.1080/16549716.2017.1387984

69. Flaxman AD, Stewart A, Joseph JC, Alam N, Alam SS, Chowdhury H, et al. Collecting verbal autopsies: improving and streamlining data collection processes using electronic tablets. Popul Health Metr. 2018;16: 3. doi:10.1186/s12963-018-0161-9

70. Singh Y, Jackson D, Bhardwaj S, Titus N, Goga A. National surveillance using mobile systems for health monitoring: complexity, functionality and feasibility. BMC Infect Dis. 2019;19: 786. doi:10.1186/s12879-019-4338-z

71. WHO. eHealth and innovation in women's and children's health: a baseline review: based on the findings of the 2013 survey of CoIA countries by the WHO Global Observatory for eHealth, March 2014. 2014.

72. Oluoch TO. Electronic medical records and clinical Decision Support Systems in HIV care in resource-limited settings. 9789462596313; 2015.

73. Oluoch T, Kwaro D, Ssempijja V, Katana A, Langat P, Okeyo N, et al. Better adherence to pre-antiretroviral therapy guidelines after implementing an electronic medical record system in rural Kenyan HIV clinics: a multicenter pre-post study. doi:D - NLM: HHSPA796290 OTO - NOTNLM

74. Oluoch T, Santas X Fau - Kwaro D, Kwaro D Fau - Were M, Were M Fau - Biondich P, Biondich P Fau - Bailey C, Bailey C Fau - Abu-Hanna A, et al. The effect of electronic medical record-based clinical decision support on HIV care in resource-constrained settings: a systematic review.

75. Chan KS, Fowles JB, Weiner JP. Review: Electronic Health Records and the Reliability and Validity of Quality Measures: A Review of the Literature. Med Care Res Rev. 2010;67: 503–527. doi:10.1177/1077558709359007

76. Bian J, Lyu T, Loiacono A, Viramontes TM, Lipori G, Guo Y, et al. Assessing the practice of data quality evaluation in a national clinical data research network through a systematic scoping review in the era of real-world data. J Am Med Inform Assoc JAMIA. 2020;27: 1999–2010. doi:10.1093/jamia/ocaa245

77. District Health Information Software 2 [Internet]. Oslo: Health Information Systems Programme; 2012. [Accessed April 9, March 2022]. Available: http://www.hisp.uio.no/

78. Amoroso CL, Akimana B, Wise B, Fraser HS. Using electronic medical records for HIV care in rural Rwanda. 2010. pp. 337–341.

79. Tsai CH, Eghdam A, Davoody N, Wright G, Flowerday S, Koch S. Effects of Electronic Health Record Implementation and Barriers to Adoption and Use: A Scoping Review and Qualitative Analysis of the Content. Life Basel Switz. 2020;10. doi:10.3390/life10120327

80. Odeny B, Hughes JP, Gloyd S, Njoroge A, Wagenaar BH, Puttkammer N. Development of novel composite data quality scores to evaluate facility-level data quality in the District Health Information System (DHIS2) in Kenya: a nationwide retrospective cohort study. doi:Under review

81. Xiao H, Augusto O, Wagenaar BH. Reflection on modern methods: a common error in the segmented regression parameterization of interrupted time-series analyses. Int J Epidemiol. 2021;50: 1011–1015.

82. Noureldin M, Mosallam R, Hassan SZ. Quality of documentation of electronic medical information systems at primary health care units in Alexandria, Egypt. East Mediterr Health J Rev Sante Mediterr Orient Al-Majallah Al-Sihhiyah Li-Sharq Al-Mutawassit. 2014;20: 105–111.

83. Ahmed R, Robinson R, Elsony A, Thomson R, Squire SB, Malmborg R, et al. A comparison of smartphone and paper data-collection tools in the Burden of Obstructive Lung Disease (BOLD) study in Gezira state, Sudan. PloS One. 2018;13: e0193917. doi:10.1371/journal.pone.0193917

84. Mohammadi Firouzeh M, Jafarjalal E, Emamzadeh Ghasemi HS, Bahrani N, Sardashti S. Evaluation of vocal-electronic nursing documentation: A comparison study in Iran. Inform Health Soc Care. 2017;42: 250–260. doi:10.1080/17538157.2016.1178119

85. Walther B, Hossin S, Townend J, Abernethy N, Parker D, Jeffries D. Comparison of electronic data capture (EDC) with the standard data capture method for clinical trial data. PloS One. 2011;6: e25348. doi:10.1371/journal.pone.0025348

86. Zeleke AA, Worku AG, Demissie A, Otto-Sobotka F, Wilken M, Lipprandt M, et al. Evaluation of Electronic and Paper-Pen Data Capturing Tools for Data Quality in a Public Health Survey in a Health and Demographic Surveillance Site, Ethiopia: Randomized Controlled Crossover Health Care Information Technology Evaluation. JMIR MHealth UHealth. 2019;7: e10995. doi:10.2196/10995

87. Ngugi P, Babic A, Were MC. A multivariate statistical evaluation of actual use of electronic health record systems implementations in Kenya. PloS One. 2021;16: e0256799. doi:10.1371/journal.pone.0256799

88. Bisrat A, Minda D, Assamnew B, Abebe B, Abegaz T. Implementation challenges and perception of care providers on Electronic Medical Records at St. Paul's and Ayder Hospitals, Ethiopia. BMC Med Inform Decis Mak. 2021;21: 306. doi:10.1186/s12911-021-01670-z

89. Douglas G, McKay M, Bwanali M, Gadabu O, Mumba S, Gondwe A. A comparison of data captured at point of care versus data captured retrospectively. AMIA Annu Symp Proc AMIA Symp. 2008; 934.

90. Tweya H, Feldacker C, Ben-Smith A, Harries AD, Komatsu R, Jahn A, et al. Simplifying ART cohort monitoring: can pharmacy stocks provide accurate estimates of patients retained on antiretroviral therapy in Malawi? BMC Health Serv Res. 2012;12: 210. doi:10.1186/1472-6963-12-210

91. Myburgh H, Murphy JP, van Huyssteen M, Foster N, Grobbelaar CJ, Struthers HE, et al. Implementation of an electronic monitoring and evaluation system for the antiretroviral treatment programme in the Cape Winelands district, South Africa: a qualitative evaluation. PloS One. 2015;10: e0127223. doi:10.1371/journal.pone.0127223

92. Aminpour F, Sadoughi F, Ahamdi M. Utilization of open source electronic health record around the world: A systematic review. J Res Med Sci. 2014;19: 57–64.

93. Hillestad R, Bigelow J, Bower A, Girosi F, Meili R, Scoville R, et al. Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. Health Aff (Millwood). 2005;24: 1103–1117.

94. Bhavaraju SR. From subconscious to conscious to artificial intelligence: A focus on electronic health records. Neurol India. 2018;66: 1270–1275. doi:10.4103/0028-3886.241377

95. Akanbi MO, Ocheke AN, Agaba PA, Daniyam CA, Agaba EI, Okeke EN, et al. Use of Electronic Health Records in sub-Saharan Africa: Progress and challenges. 2012.

96. Muthee V, Bochner AF, Kang'a S, Owiso G, Akhwale W, Wanyee S, et al. Site readiness assessment preceding the implementation of a HIV care and treatment electronic medical record system in Kenya.

97. King JD, Buolamwini J, Cromwell EA, Panfel A, Teferi T, Zerihun M, et al. A novel electronic data collection system for large-scale surveys of neglected tropical diseases. PloS One. 2013;8: e74570. doi:10.1371/journal.pone.0074570

98. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol. 2017;46: 348–355.

99. IN DANGER: UNAIDS Global AIDS Update 2022. Geneva: Joint United Nations Programme on HIV/ AIDS; 2022. Available: https://www.unaids.org/sites/default/files/media_asset/2022-global-aids-update_en.pdf [Accessed Oct 11, 2022].

100. Kim MH, Ahmed S, Abrams EJ. Pediatric HIV: progress on prevention, treatment, and cure. Curr Pediatr Rep. 2015;3: 219–229.

101. Forster M, Bailey C, Brinkhof MW, Graber C, Boulle A, Spohr M, et al. Electronic medical record systems, data quality and loss to follow-up: survey of antiretroviral therapy programmes in resource-limited settings. Bull World Health Organ. 2008;86: 939–947.

102. UNAIDS. Country factsheets. HIV and AIDS estimates. Haiti 2021. Available: https://www.unaids.org/en/regionscountries/countries/haiti [accessed Oct 11, 2022].

103. PEPFAR. Haiti Country Operational Plan (COP/ROP) 2020 Strategic Direction Summary. Available: https://www.state.gov/wp-content/uploads/2020/07/COP-2020-Haiti-SDS-FINAL-with-Appendix-C.pdf [Accessed Oct 11, 2022].

104. Auld AF, Pelletier V, Robin EG, Shiraishi RW, Dee J, Antoine M, et al. Retention throughout the HIV care and treatment cascade: from diagnosis to antiretroviral treatment of adults and children living with HIV—Haiti, 1985–2015. Am J Trop Med Hyg. 2017;97: 57.

105. PEPFAR. HAITI Country Operational Plan (COP/ROP) 2016 Strategic Direction Summary. 2016. Available: https://www.state.gov/wp-content/uploads/2019/08/Haiti-15.pdf [Accessed Oct 11, 2022].

106. deRiel E, Puttkammer N, Hyppolite N, Diallo J, Wagner S, Honore JG, et al. Success factors for implementing and sustaining a mature electronic medical record in a low-resource setting: a case study of iSante in Haiti. Health Policy Plan. 2018;33: 237–246. doi:10.1093/heapol/czx171

107. Puttkammer N, Baseman JG, Devine EB, Valles JS, Hyppolite N, Garilus F, et al. An assessment of data quality in a multi-site electronic medical record system in Haiti. Int J Med Inf. 2016;86: 104–116.

108. Mate KS, Bennett B, Mphatswe W, Barker P, Rollins N. Challenges for routine health system data management in a large public programme to prevent mother-to-child HIV transmission in South Africa. PloS One. 2009;4: e5483.

109. Podewils LJ, Bantubani N, Bristow C, Bronner LE, Peters A, Pym A, et al. Completeness and Reliability of the Republic of South Africa National Tuberculosis (TB) Surveillance System. BMC Public Health. 2015;15: 765. doi:10.1186/s12889-015-2117-3

110. Carlucci JG, Liu Y, Clouse K, Vermund SH. Attrition of HIV-positive children from HIV services in low-and middle-income countries: a systematic review and meta-analysis. AIDS Lond Engl. 2019;33: 2375.

111. Clouse K, Phillips T, Myer L. Understanding data sources to measure patient retention in HIV care in sub-Saharan Africa. Int Health. 2017;9: 203–205.

112. Amouzou A, Maïga A, Faye CM, Chakwera S, Melesse DY, Mutua MK, et al. Health service utilisation during the COVID-19 pandemic in sub-Saharan Africa in 2020: a multicountry empirical assessment with a focus on maternal, newborn and child health services. BMJ Glob Health. 2022;7: e008069.

113. Aliyu MH, Blevins M, Arinze F, Megazzini KM, Bussell S, Dunlap J, et al. Enrolment trends in a comprehensive HIV programme in rural north-central Nigeria: improved care indices, but declining quality of clinical data over time. Pathog Glob Health. 2015;109: 75–83.

114. Gumede-Moyo S, Todd J, Bond V, Mee P, Filteau S. A qualitative inquiry into implementing an electronic health record system (SmartCare) for prevention of mother-to-child transmission data in Zambia: a retrospective study. BMJ Open. 2019;9: e030428.

115. Verani AR, Lane J, Lim T, Kaliel D, Katz A, Palen J, et al. HIV Policy Advancements in PEPFAR Partner countries: a review of data from 2010–2016. Glob Public Health. 2021;16: 390–400.

116. Ehrenkranz P, Grimsrud A, Rabkin M. Differentiated service delivery: navigating the path to scale. Curr Opin HIV AIDS. 2019;14: 60–65.

117. Hoxha K, Hung YW, Irwin BR, Grepin KA. Understanding the challenges associated with the use of data from routine health information systems in low-and middle-income countries: A systematic review. Health Inf Manag J. 2022;51: 135–148.

118. Feng S, Hategeka C, Grépin KA. Addressing missing values in routine health information system data: an evaluation of imputation methods using data from the Democratic Republic of the Congo during the COVID-19 pandemic. Popul Health Metr. 2021;19: 1–14.

119. Sohn AH, Judd A, Mofenson L, Vicari M, Jerene D, Leroy V, et al. Using observational data to inform HIV policy change for children and youth. J Acquir Immune Defic Syndr 1999. 2018;78: S22.

120. Hong TT, Phuong Hoa N, Walker SM, Hill PS, Rao C. Completeness and reliability of mortality data in Viet Nam: Implications for the national routine health management information system. PloS One. 2018;13: e0190755. doi:10.1371/journal.pone.0190755