MALAWI JOURNAL OF SCIENCE AND TECHNOLOGY VOLUME 13: ISSUE 1 | 2021

Investigation of Key Factors Affecting Quality of Patient Data from National Antiretroviral Therapy Electronic Medical Record System in Malawi

Deliwe B. Nkhoma [1*] & Kondwani G. Munthali [2]

[1]Baobab Health Trust, P. O. Box 31797, Lilongwe 3, Malawi.
[2] University of Malawi, Chancellor College, Computer Science Department, P.O. Box 280, Zomba, Malawi.

***Corresponding Author Email:** dbnkhoma@gmail.com

Abstract

The Ministry of Health in Malawi implemented a National Antiretroviral Therapy Electronic Medical Record system currently deployed in over 150 health facilities. It thus expected quality and timely quarterly cohort reports. However, the raw electronic reports are rarely complete, accurate and consistent requiring cleaning hence being delayed. Such reports are now very critical under the COVID-19 pandemic. Adopting a mixed-method approach, this study assessed the key factors that affect quality of data entered in the electronic medical records system and the reports produced by the system. The study interviewed 134 health-care workers in 17 sites and 10 Baobab Health Trust officers. Observations were conducted and secondary data analysed. The analysis shows that the EMRs lacks proper documentation and validation rules, making it hard to maintain and increasing chances of duplicate entry, respectively. Coupled with lack of trained personnel, it was revealed that one set of login credentials is used by multiple users and vital data elements being null compromising security and completeness, respectively. The electronic medical records system was not used at 40% of the sites as a point of care system hence being used as a back-data entry tool. Thus, there is need to revise the system to include necessary validations, security features, back data-entry form and data quality dashboards.

Keywords: *Electronic Medical Records system, Data Quality, System Quality, Information Quality.*

1 INTRODUCTION

The impact of human immunodeficiency virus infection / acquired immunodeficiency syndrome (HIV/AIDS) pandemic in African countries has been very huge. It has not only caused considerable strain on various national health-care system, but also increased the number of orphans, reduced productive human capital and productivity, eroded knowledge and skills. In addition, it has put pressure on national budgets, increased the poverty-stricken populace and reduced the quality of life, health and wellness (Fraser et al., 2005). A milestone in the history of HIV/AIDS has been the availability of new classes of drugs that have allowed the introduction of combination antiretroviral therapy (ART) and the gradual evolution of HIV infection into a chronical condition with lower mortality. This has seen the number of people receiving ART increase considerably reaching approximately 5.25 million by the end of 2009 about 3.9 million of which were living in sub-Saharan Africa. Despite the growth, coverage of the population in need of ART, calculated on the basis of the 2010 WHO guidelines, was 37% and 28%, and in eastern and southern Africa was 41% and 32% for adults and children respectively (Schouten et al.,2011). Again, mortality of patients starting ART has been substantially higher in these parts than in industrialized countries, particularly in the first few months of treatment (Keiser et al., 2008). Consequently, sub-Saharan governments begun to expand access to HIV care rapidly and ensure that all the health institutions are ready to receive and assist patients. For instance, in South Africa, approximately 900 000 people started treatment and another 1.2 million were targeted in the first 6 years and subsequent 2 – 3 years of starting the ART programme, respectively (Colvin et al., 2010). Similarly, the Malawi ART programme focused on the roll out of one first-line ART regimen using a fixed-dose generic combination of stavudine, lamivudine and nevirapine across the country in 2004. Two alternative first-line ARV regimens and a second-line ARV regimen were available in only four ART clinics in the country (Schouten, 2011). At the end of every quarter, cumulative treatment outcomes of all patients started on ART at each site and the number of patients retained alive and on treatment were recorded (Schouten et al., 2011).

As such, good health is not only important for individuals, but also for governments as it plays a central role in achieving sustainable economic development as well as effective use of resources (Coleman, 2013; Fraser et al., 2005). Notwithstanding, health care systems in developing countries are riddled with poor access to patient information, collaboration among care providers, decision support at the point of care, and patient information management. Coleman et al. (2015) argued that electronic medical record systems (EMRs) have become increasingly of great potential and use in alleviating these problems. The benefits of EMRs range from reduction of medical errors, easy access to previous records, incorporation of tools that send alerts in case of problems, to flexibility of patients moving from one facility to another (Williams and Boren, 2008; Coleman, 2013; Fraser et al., 2005; and Ohuabunwa et al., 2016). They also provide clinical, organisational, and societal benefits that include adequate record keeping for effective follow-up, multiple user access and efficient health-care workflows (Williams and Boren, 2008; Fraser et al., 2005; and Ohuabunwa et al., 2016). Beyond these advantages, EMRs have been a rich source of data for secondary purposes including research, public health surveillance, chronic disease management, and quality improvement.

It is thus not surprising that the Ministry of Health (MoH) in Malawi through the Department of HIV/AIDS (DHA) expects quarterly cohort analysis reports from all health facilities in Malawi. These reports are then reconciled at DHA into a national quarterly cohort report for ART drug and HIV management. To achieve this, in most of the high-burden clinics, health-care workers can take up to 10 or more days to prepare the quarterly cohort reports manually. Sometimes clinics are closed to complete the preparation of the reports. To reduce the burden on the health-care workers and improve data quality in generating the cohort reports, Baobab Health Trust (BHT), designed and built a National Antiretroviral Therapy (NART) EMRs in 2006. The NART EMRs was based on the specifications from MoH through DHA and is currently deployed in over 150 health facilities across the country. Since then, NART EMRs have evolved significantly along with the number of partners and participants (Giles, 2014). Through such partnerships Malawi has developed national ART Treatment Guidelines, which emphasize a structured and standardized approach for all aspects of ART delivery, including monitoring and evaluation. Using the successful Directly Observed Treatment Short-Course (DOTS) model adapted by National TB Control Programmes throughout the world, Malawi has developed a system of quarterly and cumulative ART cohort analysis report (Harries et al., 2007). Giles (2014) narrates that walking into one of the busiest HIV/AIDS clinics in Malawi with an annual facility volume of approximately 13,957 patients enrolled and receiving ART, clients seamlessly transition from point to point as clerks and clinicians use bar-code scanners and touch screen computers. At every step, a healthcare worker queries the EMRs to retrieve information from previous appointments and instantly record information from the current visit. At the end of the visit, a label with printed information from the visit is affixed to the patient's health passport, which acts as a vehicle for sharing information within or between health care facilities.

The introduction of the NART EMRs saw the quality of the quarterly case-finding data improve. However, the completeness and accuracy of the primary outcome data are often compromised. To ensure complete, consistent and credible quarterly reports, health-care workers and BHT staff are then engaged in data cleaning exercises. While the cleaning exercise does eliminate errors, duplicates and inconsistencies in the source data to achieve high-quality data, a prerequisite for credible reports for decision making (Galhardas, 2014; Rahm and Do Hai, 2016), it is a time-consuming exercise. Consequently, as observed by Forster et al. (2008), secondary outcome analyses of the primary cohort data are no longer feasible due to time constraints and are often delayed if done. Either way, management of the ARVs drug and overall HIV is compromised.

As shown, despite the improved efficiency in patient handling at point of care (POC), implementation of EMRs is replete with many challenges in the developing world. These range from lack of financial resources to monitor and support implemented POC EMRs on one hand to lack of hardware and software standardization making it hard to automatically generate alerts for decision making on the other (Douglas, 2009; Williams and Boren, 2008). It has also been argued that recording clinical data on paper and retrospective data entry is less invasive with respect to existing workflow and therefore less likely to negatively impact the delivery of clinical care (Douglas, 2009). Other challenges include differences in training, computer literacy, negative attitudes, and lack of interest by system users, data entry paradigm, work environment and culture and scepticism on the feasibility of POC EMRs (Douglas, 2009; WHO, 2012). Thus, the performance of an EMR as a Health Information System (HIS) should therefore be measured not only on the quality of data produced, but on evidence of the continued use of data to improve health system performance, to respond to emergent threats, and improve the health sector in general (Forster et al., 2008). It is imperative that these possible determinants of performance are mitigated to improve HIS in terms of data availability, quality and use.

The study, therefore, sets out to evaluate the quality of data entered in the NART EMRs and the reports produced by the system. Specifically, the NART EMRs software, hardware and networks are assessed to ascertain their influence on the quality of generated data on one hand. On the other, the study analyses how the current NART EMRs training program addresses the quality of data. The evaluation is also important as the world grapples with the COVID-19 pandemic.

1.1 EMRs data quality

Quality, accuracy and completeness of the information in medical records is fundamental to good patient care and the transition from paper-based records to EMRs is progressive (Greiver et al.,, 2012). More so when the data is extracted and compiled into databases for secondary research and surveillance health use (Coleman et al., 2015; Van Velthoven et al. 2016). Thus, the quality of the data entered in the EMRs needs to be assessed against inconsistent or missing diagnostic coding and risk factor designation, “dirty data” (misspelled words, inconsistent word strings, free text strings instead of structured data), missing “metadata” or description of data content and data entered in inconsistent or incorrect database fields (Greiver et al., 2012). These challenges are grouped into completeness (is all the data present?), reliability (is the data recorded in the same way across practices and over time?) and validity (is the data correct?). There are no agreed reference standards for reporting data quality in primary care and this limits measurement of data quality in electronic patient records (Coleman et al., 2015; Greiver et al., 2012). While there are many possible ways to measure EMR data quality and many areas that can be measured, systematic reviews of data quality assessment have noted a focus on diagnostic data, laboratory testing, risk factors and demographic information, with limited information on data quality regarding preventive services (Greiver et al., 2012).

Comparing the data reported in 3 sub-Saharan African counties (Mozambique, Rwanda, Zambia) by clinics with the data counted from source documents revealed a lot of inaccuracies that led to the conclusion that clinic data lacked validity (Gimbel et al. 2017; Wright and Odama, 2012). It was noted that not a single facility possessed data element definitions to guide data collection methods. As a result, the reasons used to calculate data elements changed from month to month depending on which clinician was doing the report write up. Poor quality routine data contributes to poor decision making, inefficient resource allocation, and loss of confidence in the health system, and may threaten the validity of impact evaluations (Gimbel et al., 2017; Wright and Odama, 2012).

1.2 System Quality and User Satisfaction

EMR applications have the potential to improve the quality of healthcare. Immediate advantages include decreased time spent on paperwork, reduction of the cost of health care and increased patient satisfaction (Joos et al., 2006) . Despite the advantages of EMR applications, implementation of the same may be resisted if users are not satisfied with the system. To fully satisfy the users with the EMR benefits the quality of the system must ensure that the physicians, clinicians and/or nurses, as decision makers and problem solvers, routinely use these applications (Joos et al., 2006; Maria, n.d.; Sittig et al., 1999). It is thus imperative that EMR application is optimal in its response time, logic and efficient flow of tasks, ability to complete desired tasks, ease of correcting mistakes, effects on an individual’s time, and proper training on the system (ibid). This helps to overcome already preconceived concerns about security and confidentiality, time incurred by EMR use, or perceived impacts on the quality of patient care by physicians and nurses.

1.2.1 Information Quality and Software Quality

Quality is defined as ‘fitness for use’. In the context of data quality, this means that data are of sufficient quality when they serve the needs of a given user pursuing specific goals be it medical care, status, and outcomes of a diverse population that is representative of actual patients. Data quality is critical to ensuring patient safety, communicating delivery of health services, coordinating care, and health-care reporting (Weiskopf and Weng, 2013). The quality of health-care across the continuum depends on the integrity, reliability, and accuracy of health information. The emergence of EMR-related errors results in data being lost or incorrectly entered, displayed, or transmitted, leading to loss of information integrity

(American Health Information Management Association (AHIMA), 2013; Bowman Sue, 2013; Weiskopf and Weng, 2013).

1.2.2 Conceptual Framework

The Updated DeLone & McLean Information System (Updated D&M IS) success model has six dimensions as follows: information quality, system quality, service quality, user satisfaction and net benefits (DeLone and Mclean, 2003). The Data Quality Assessment Framework (DQAF) also assesses data quality (Information quality) which is just one component of the updated D&M IS success model. The attributes assessed under this component in these two models slightly differ. DQAF assesses consistency, validity, timeliness, completeness and uniqueness while updated D&M IS success model assesses accuracy, timeliness, completeness, relevance and consistency. This study did secondary data analysis concentrating on the data elements that are reported on the ART cohort analysis report. To ascertain quality of data, a combination of some data elements are not to be recorded more than once, meaning no duplicates. For example: the combination of patient's name, date of birth, gender and some other elements should be unique to that particular patient. Similarly, these data elements should not contain any null/empty values, they should be complete, and in addition these data elements formats should be the same when comparing one dataset to the other or indeed the dataset and the manual register.

System quality dimension evaluates the overall quality of the system and it evaluates the system in terms of user friendliness, stability, security, availability, and reliability. Service quality evaluates the system according to the quality of services the system is to offer. It not only checks if the IS has up-to-date hardware and software, but also whether it is responsive for example, if employees are able to give prompt service to system users. It also evaluates the ability of the IS to: provide assurance to the employees that they will be able to do their job well given their knowledge level; guarantee that it has users’ best interests at heart (empathy) and nature of use (DeLone and Mclean, 2003; Halonen, 2009). This study used both the Modified Updated D&M IS success model and DQAF framework. The study used the latter for analysing Information Quality and former when looking at System and Service Quality respectively. However, this study also analysed secondary data to validate the Information Quality especially on the data elements that are reported on ART quarterly cohort report. On this, the study concentrated only on Uniqueness, Completeness, Consistency and Timeliness dimensions only of the DQAF. DQAF provides a structure for assessing data quality dimensions that are important because they enable one to understand why data is being measured (Sebastian-Coleman, 2014).

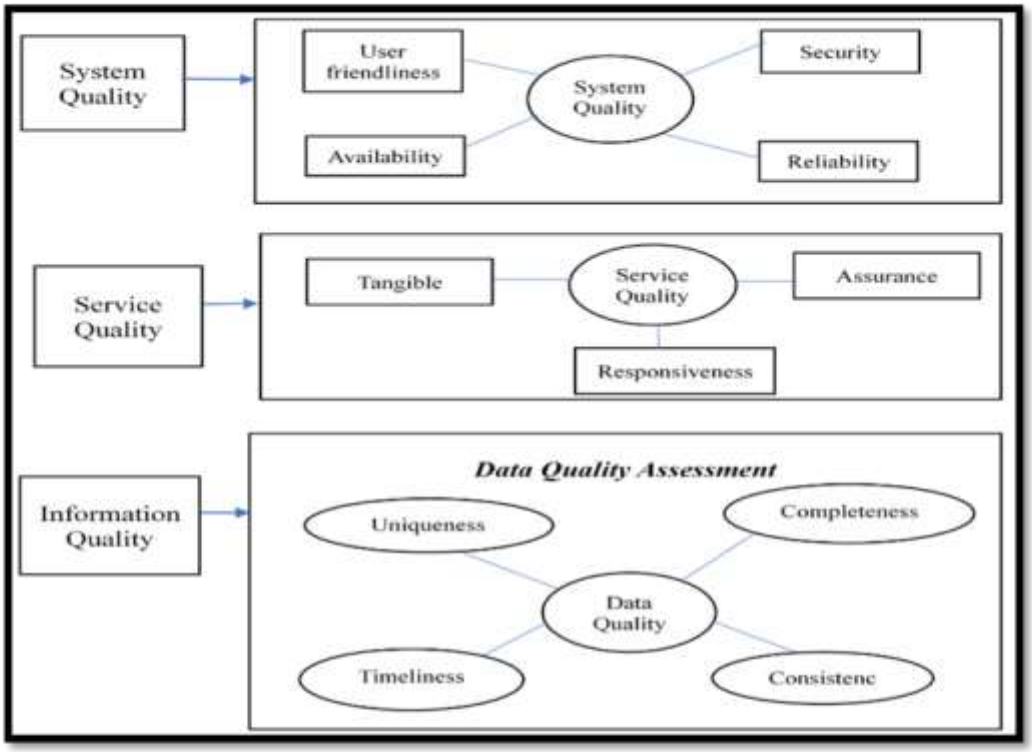


Figure 1: Conceptual Framework (adapted from DeLone & McLean, (2003) and DAMA UK Working Group on Data Quality Dimensions (2013).

Source: Authors

2 MATERIALS AND METHODS

At the time of this study, out of the 112 sites managed by BHT, 37 are old sites which initially were running on OpenMRS 1.1 database platform and they were migrated to BHT database platform based on OpenMRS 1.7 (category 1). 12 sites migrated from Fuchia software to BHT software (category 2), whereas 65 sites were deployed after the software and database platform changed to BHT database platform based on OpenMRS 1.7 hence they required back-data entry before going live (category 3). These provided natural categories for sampling and Table 1 provides the details of the sampled sites. The study has more sites and respondents in the first category not only because they are busy in that they handle more patients, but also because they have more experienced team of health-workers. In all sites the respondents were NART EMRs users that included: doctors, clinicians, nurses, clerks, monitoring and evaluation officers, other health-care workers and BHT officers altogether 144 (134 NART EMRs users and 10 BHT officers).

Table 1**: *Sample size as per cluster category***

Category	**Sample sites**	**Number of respondents**
1	• Tertiary sites: Queen Elizabeth (QECH) and Zomba Central Hospitals • Secondary sites: Nkhota-kota, Salima and Machinga District Hospitals. • Primary sites: St. Gabriels Mission Hospital, Area 25 Health Centre, Ndirande Health Centre	• Tertiary: 28\ • Secondary: 20 • Primary: 23
2	• Secondary Sites: Chiradzulo District Hospital • Primary sites: Namadzi, Nkalo and Mbulumbuzi Health Centres	• Secondary: 10 • Primary: 17
3	• Secondary sites: Ntchisi District Hospital • Primary Sites: Macro Blantyre and Lilongwe ART Clinics, Area 18 and Matiki Health Centres	• Secondary: 5 • Primary: 31

Source: Authors

The study used mixed methods approach and two-staged cluster sampling was adopted. Empirical data was collected through semi-structured interviews and participant observations. The former was used to collect data on NART EMRs software, hardware and network status, while the latter was adopted to assess the extent to which the current training program is addressing the quality of data entered and the reporting functionality of the NART EMRs. In so doing the researcher observed how user trainings are conducted and also how the application is used.

Existing documents reviewed included: NART EMRs user training manual, BHT help-desk Redmine system reports, and MoH ART guidelines. The specific secondary data elements of interest to the study are as shown in Table 2 with their corresponding priority. A My Structured Query Language (MySQL) script was developed to extract these data elements from the main data-set to a Comma Separate Values (CSV) file and R statistical software was used for further analysis.

Table 2: ***Secondary data elements analysed***

Data Element	Priority
Gender, Date of Birth, HIV enrolment date, Reason for starting ARVs, ARV regimen category, TB status, Side effects, ARV adherence Rate, HIV primary Outcomes	Mandatory
Pregnancy Status, Breastfeeding Status and Family Planning Method taken	Mandatory (only for female patients within the child bearing age of 9 to 55)

Source: Authors

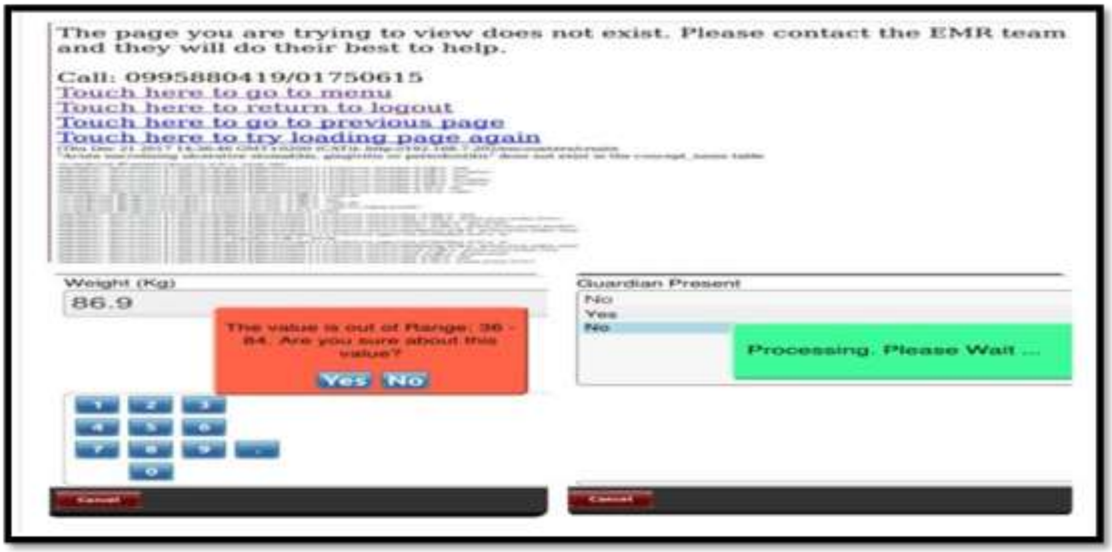
The conceptual framework of the study fused the Updated DeLone & McLean (D&M) Information System (IS) Success Model and the Data Quality Assessment Framework (DQAF). The latter was used to analyse the information quality while the former to analyse the system and service quality. As such, while the Updated D&M IS success model has 6 dimensions, this study only used two that evaluate: 1) how the system is able to deliver benefits – system quality, and 2) the system according to the quality of the services the system is to offer – service quality. In this case, the system (software, hardware, network infrastructure and system user training) quality of NART EMRs was evaluated along the dimensions of user friendliness, availability, security and reliability. On the other hand, service quality of the BHT Redmine help-desk system, a repository of defects (Error, Failure and Bugs) and all requests of new features of all EMRs by BHT including NART, was evaluated along the dimensions of tangibility, assurance and responsiveness. The condition of the data elements in relation to expectations of the cohort quarterly reports was assessed along the DQAF dimensions of information quality that included: uniqueness, completeness, timeliness and consistence.

Thematic Analysis (TA) was used to interpret qualitative data. The themes were derived from the conceptual framework classifications namely; System Quality, Service Quality and Information Quality. The semi-structured interviews and observations guidelines were structured in a way that the researcher collected diverse aspects of data. By using, thematic analysis the researcher was able to link the various concepts and opinions of the respondents and compare these with the data that has been gathered in different situation at different times during the research. Quantitative data was extracted using SQL script which extracted only the data variables that were of interest and then exported into CSV files. This nominal data was analysed mainly using the grouping method using the R statistical package and Libre Calc respectively.

3 RESULTS AND DISCUSSIONS

3.1 NART EMR System Quality

NART EMR software has evolved since it was first deployed at QECH in Blantyre in 2006 with biennial major software upgrades following revisions of the Malawi ART Guidelines by MoH through DHA. BHT, working closely with DHA, defines the features and functionality to be changed and/or added, and then implements and tests the features. Despite having no proper documentation of these software upgrades, the overall system is fairly user friendly. The findings show that the system User Interface (UI) has well defined validation rules on dates, numbers and other fields and provides feedback through meaningful error messages and effective error handling whenever possible as described in Figure 2. The system display pages are presented consistently in terms of both look and feel with the questions or items on the pages presented in English and, whenever possible, phrased in Chichewa for the users to capture the required data.



The page you are trying to view does not exist. Please contact the EMR team and they will do their best to help.

Call: 0995880419/01750615

Weight (Kg): 86.9

The value is out of Range: 36 - 84. Are you sure about this value?

Processing. Please Wait ...

Figure 2: NART EMR feedback.

Source: Authors

However, the NART EMRs has functionality gaps especially in the calculation of ARV adherence rate and patient's nutritional status, specifically body mass index (BMI), in pregnant women and under five children which do not use Mid-Upper Arm Circumference (MUAC). Again, the system has some missing data. For instance, ‘Skin rush’, as a vital sign, is missing on the list of side effects. These gaps affect the reliability of the system on these elements. In terms of security, the findings indicate that in 4 out of the 17 sites, users log into the NART EMRs using one set of credentials. This has the potential to compromise the security of the system as it does not only affect system user accountability but also the system has no way of controlling the user roles and permissions considering that the NART EMRs is accessed by a variety of users at different levels.

Of the 17 sites visited, only 7 use the NART EMRs as a point of care (PoC) system. The reasons for the unavailability of the NART EMRs at the PoC, according to the BHT Redmine help-desk system, range from prolonged power black outs which results in draining the power backup batteries faster than expected; running out of electricity prepaid units; software problems and sometimes faulty hardware. The availability of the system at PoC is further hampered by variations of workflows at the sites due to: (1) infrastructural challenges at the site, (2) the number of health-care workers present against the number of patients per visit, and (3) the number of champion users at a site.

3.2 NART EMRs Service Quality

Both NART EMRs hardware and software has evolved overtime. Software is revised biennially by MoH through DHA with the revision of the ART guidelines. Hardware, on the other hand, is revised due to: (1) growth in data hence requiring more space on the servers; (2) need to lower power consumption to cut budget lines; and (3) need to improve the speed, durability and performance of the hardware. Though the hardware and software are regularly revised and updated, the user training manual is not and therefore it lacks vital additional features. Aside from that, the user manual also lacks basic technical troubleshooting guidelines for some of the known problems and reporting for immediate response. For instance, NART EMRs enable users to void a record if they make any mistake or indeed update patient records, run reports and perform some site configurations. However, it was observed that 6 out of 10 users are not aware of this and once they have made a mistake, they either re-enter the whole record if the system permits or just record the actual results in the manual registers. Inadvertently, the former leads to duplicates while the latter more data inconsistencies. Therefore, BHT spends more man hours and money in supporting basic things that could have been tackled by the end users if they were equipped with the knowledge. This gets compounded by high staff turnover. Because of this, it was found that more than 50% of the users get on job training from champion users. While this does work, in the absence of a user manual, chances of the new users' carrying on with erroneous procedures are high leading to more duplicates or inconsistencies. Secondly, when it is the champion user that has moved chances of continued use of the system at the PoC reduces. The former compromises the quality of the data, and the latter the quality of the service.

The study also reveals that not all users report problems, recommendations and/or clarifications through BHT help-desk Redmine system. The users resort to either emails, phone calls or indeed through social media to contact the Support and Deployment Officer responsible for that site or any other officer in the software development department. Either way, the users fail to provide sufficient information when reporting a problem. With lack of documentation of the BHT support and help-desk system, the support officers must ask as many questions as possible to determine the nature of the problem. This affects the responsiveness in addressing the problem. The responsiveness is further compounded by the fact that the officers have no way of monitoring the system in real time over the network, due to intermittent network, forcing them to physically visit the site to tackle the problems. Consequently, technical knowledge becomes inevitable in support and maintenance of NART EMRs. Overall, these make maintenance of NART EMRs software or hardware very hard and expensive. Figure 3 shows Redmine category status and Figure 4 shows Redmine Item subject category.

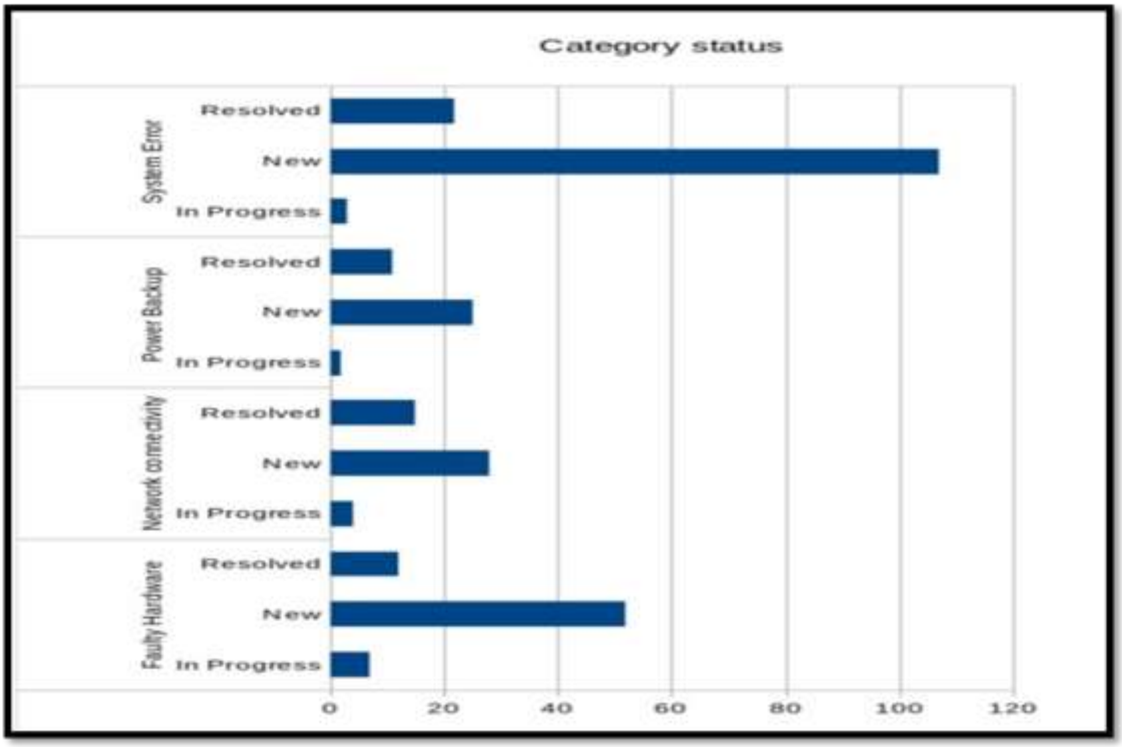


Figure 3: BHT Redmine category status.

Source: Authors

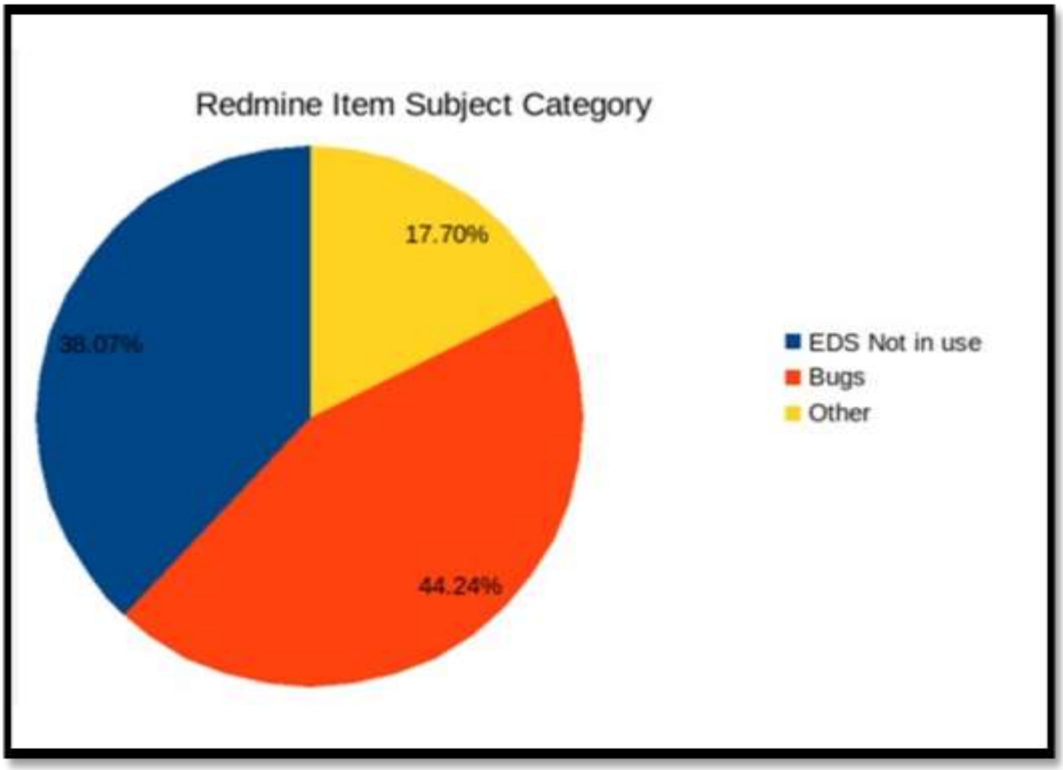


Figure 4: BHT redMine item subject category.

Source: Authors

3.3 NART EMRs Information Quality

By definition, every patient enrolled on an HIV program should be unique in which case no two patients should have exactly the same date of birth, gender, HIV enrolment date at the current site, date initiated on ART at the current site or another site, and reason for starting ART. However, the study has found that in all the 17 sites visited patients do not have unique data records where for instance some have multiple reasons for starting ART. The results show that at Macro Lilongwe ART clinic 8.2% of the patient data is not unique, followed by QECH and ZCH at 7.2% and 1% respectively (see figure 5).

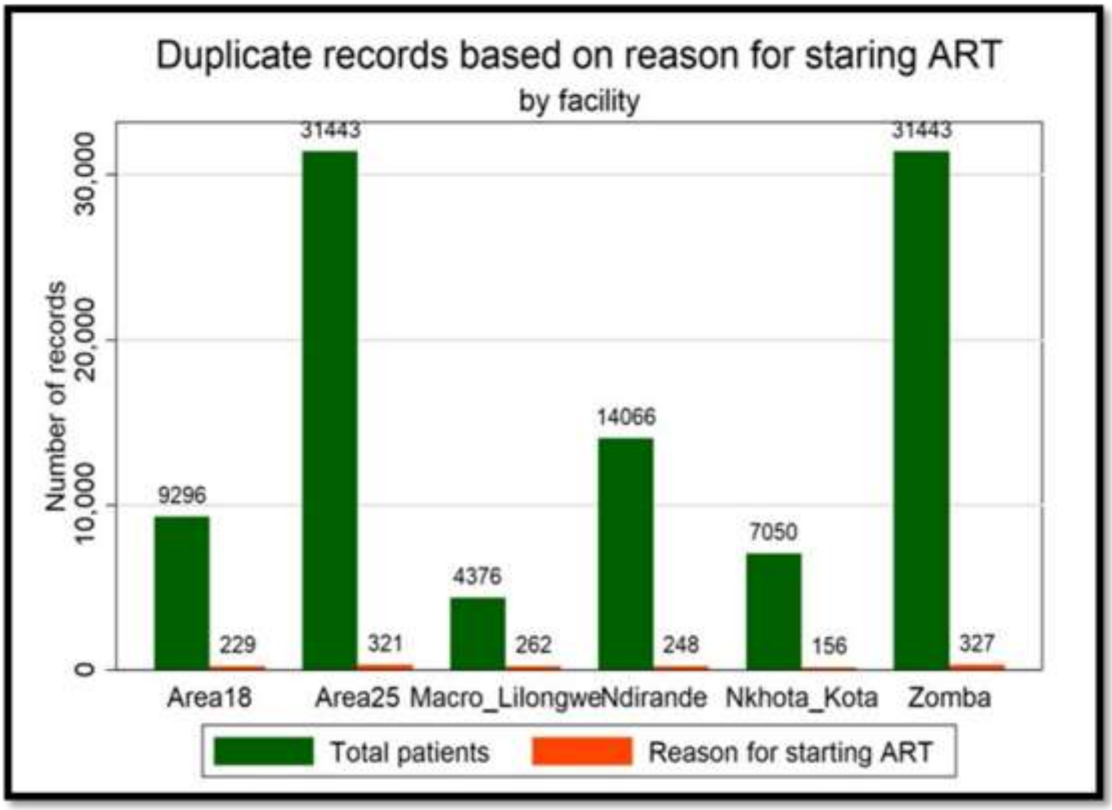


Figure 5: Duplicate ART start reasons records.

Source: Authors

In terms of completeness of the data elements analysed in this study (see Table 2), only regimen category and breastfeeding observations have 'null' values (see **Table 3**). Out of the 11 data elements that were checked for 'null' values only two have 'null' values in 7 out of 17 sites.

Table 3: ***Regimen category and breastfeeding missing records***

Facility	Number of records	Records Missing	% missing
Regimen category			
Area 25 Health Centre	117 576	331	0.28%
Queen Elizabeth Central Hospital	4660	743	15.94%
Salima District Hospital	20 564	92	0.45%
Zomba Central Hospital	5288	4	0.08%
Breastfeeding			
Namadzi Health Centre	6620	1	0.02%
Ndirande Urban Health Centre	85 430	3616	4.23%
Ntchisi District Hospital	16 787	1	0.01%

Source: Authors

The findings show that NART EMRs data had differences when comparing two or more representations of a thing against a definition. Some data elements with data type as number are saved as strings, or those with data type date are saved as strings. However, this study also looked at the consistency of data from two dimensions; (1) observations that are supposed to be done by women only and (2) if these female specific observations are asked to only women that are within the child bearing age (9 to 55 years old). The findings revealed that some men and women outside the childbearing age have observations that are supposed to be for women specifically those in the childbearing like pregnancy status, breasting feeding status, family planning method**.** The results show that Matiki Health Centre has 58.33% of the patient records having pregnancy status observations for males and female patients outside the childbearing age, while Macro Lilongwe ART clinic and Zomba Central Hospital had 55.07% and 18.68% respectively (see figure 6).

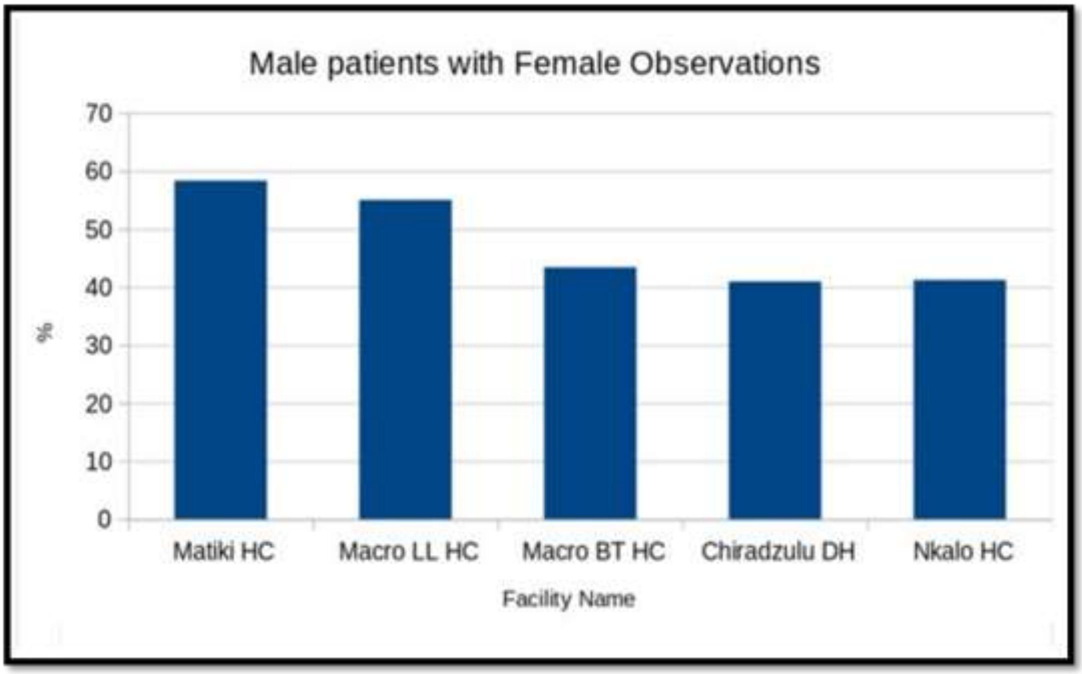


Figure 6: Male patients with female observations.

Source: Authors

It was established that 80% of these data inconsistencies were introduced by data migrations from OpenMRS1.1 and an MSF EMR running on Microsoft Access database platform to BHT database platform based on OpenMRS1.7 in 2012 and 2016 respectively. Nevertheless, the study also exposed gaps in the data, for example, patients having multiple encounters during the same visit; patients with missing regimen category; patients without demographics like names, gender and birthdate; patients without outcomes; patients with “Patient died” outcome but without death date; and patients without vitals. Despite efforts that saw 60% of the data anomalies being fixed through MySQL scripts more than half electronic quarterly reports are inconsistent with the manual ones. As a remedy, BHT together with health-care workers manually verifies patients' mastercards before updating the records until the report is both internally and externally consistent.

4 CONCLUSIONS AND RECOMMENDATIONS

Timeliness, availability and accessibility of data in making business decisions, is one of the most important aspects of database management. It also describes whether the data was submitted in due time, respecting the data gathering deadline (Mihăiloaie, 2015). NART EMR reports take time to be produced due to data cleaning processes and also back data entry exercise. This entails that data is not always available and reliable thereby delaying submission of reports to DHA and other stakeholders for decision making.

This is largely because NART EMRs do not have most of the important documentation ranging from user requirements, system design, test plan, test cases, and test results among others. The system's UI lack some validations which allows users 70% chance of entering duplicate tasks/encounters per patient visit. Similarly, security is compromised with multiple using single set of login details per site. BHT officers spend more man hours in addressing the problems reported from the sites due to (1) underutilization of BHT helpdesk, Redmine, lack of proper documentation of the Frequently asked Questions and Answers (FQA); (2) poor network connectivity; and (3) the user training manual and user manual not incorporating all crucial areas that the minimum end user should be trained in. Lastly, the study reveals that back data-entry is inevitable due to; (1) infrastructural challenges at the sites; (2) lack of trained personnel per site (3) faulty hardware; (4) network slowness problems; (5) software problems; and (6) running out of prepaid ESCOM units and also drainage of power-backup because of prolonged black-outs.

The study recommends building a proper NART EMRs documentation (training and user manuals) and a revision of the NART EMRs to include all the necessary validations and build data quality checks dashboards. It is also recommended that users be encouraged to use own credentials to improve system security. There is also need to: improve network infrastructure; build back-data entry form which will contain all the necessary data components; redesign the power-backup system so that it suits the current situation in Malawi; raise awareness on the importance of using helpdesk; and build helpdesk Frequently Questions and Answers documentation.

REFERENCES

American Health Information Management Association (AHIMA) (2013). Data Standards, Data Quality, and Interoperability. *Journal of AHIMA* **84(11|):**64-69. http://library.ahima.org/doc?oid=107104#.V6P5JPkrJD-

Bowman, S. (2013). Impact of Electronic Health Record Systems on Information Integrity Quality and Safety Implications Perspectives. *Online Journal Perspectives in Health Information Management.* http://perspectives.ahima.org/impact-of-electronic-health-record-systems-on-information-integrity-quality-and-safety-implications/

Coleman, A. (2013). *An Integrated Model to Share Patient Health Records in Public and Private Hospitals in South Africa* ***7***(2):87–93.

Coleman, N., Halas, G., Peeler, W., Casaclang, N., Williamson, T. & Katz, A. (2015). From Patient Care To Research: A Validation Study Examining the Factors Contributing to Data Quality in a Primary Care Electronic Medical Record Database. *BMC Family Practice* ***16***(1):1–8. https://doi.org/10.1186/s12875-015-0223-z

Colvin, C. J., Fairall, L., Lewin, S., Georgeu, D., Zwarenstein, M., Bachmann, M., Uebel, K. E. & Bateman, E. D. (2010). *Expanding access to ART in South Africa: The Role Of Nurse-Initiated Treatment Lessons From The STRETCH Trial In The Free State* ***100***(4): 210. https://doi.org/10.1186/1478-4491-7-75

DAMA UK Working Group on Data Quality Dimensions. (2013). *The Six Primary Dimensions For Data Quality Assessment: Defining Data Quality Dimensions*. https://www.dqglobal.com/wp-content/uploads/2013/11/DAMA-UK-DQ-Dimensions-White-Paper-R37.pdf

DeLone, W. H. & Mclean, E. R. (2003). The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of Management Information Systems / Spring* ***19***(4): 9–30. https://doi.org/10.1073/pnas.0914199107

Douglas, G. (2009). *Engineering an EMR System in the Developing World Necessity is the Mother of Invention* 1–154. papers2://publication/uuid/1F5C1748-4778-4231-BA34-B65F80CF8EBE

Forster, M., Bailey, C., Brinkhof, M. W. G., Graber, C., Boulle, A., Spohr, M. & Eggera, M.

(2008). Electronic Medical Record Systems, Data Quality and Loss To Follow-Up: Survey Of Antiretroviral Therapy Programmes In Resource-Limited Settings. *Bulletin of the World Health Organization* ***86***(12):939–947. Retrieved from https://doi.org/10.2471/BLT.07.049908

Fraser, H. S. F., Biondich, P., Moodley, D., Choi, S., Mamlin, B. W. & Szolovits, P. (2005). Refereed papers Implementing electronic medical record systems in developing countries. *Informatics in Primary Care* ***13***:83–95. https://doi.org/15992493

Galhardas, H. (2014). *Introduction to Data Cleaning* in Carlo Batini, M.S. & Catarci, T. (Eds.); 2009:1–36. https://docplayer.net/19760199-Data-cleaning-helena-galhardas-dei-ist.html

Giles, D. (2014). Public Health Informatics in Action in Malawi_ Making Life Easier For Healthcare Workers And Patients While Improving Quality Through An Innovative National Electronic Medical Record System_Blogs_CDC. Retrieved from https://blogs.cdc.gov/global/2014/05/21/public-health-informatics-in-action-in-malawi-making-life-easier-for-healthcare-workers-and-patients-while-improving-quality-through-an-innovative-national-electronic-medical-record-system/

Gimbel, S., Mwanza, M., Nisingizwe, M. P., et al. (2017). Improving data quality across 3 sub-Saharan African countries using the Consolidated Framework for Implementation Research (CFIR): Results from the African Health Initiative. *BMC Health Services Research*, ***17***(Suppl 3). https://doi.org/10.1186/s12913-017-2660-y

Greiver, M., Barnsley, J., Glazier, R. H., Harvey, B. J. & Moineddin, R. (2012). Measuring Data Reliability For Preventive Services in Electronic Medical Records. *BMC Health Services Research* ***12***(1). https://doi.org/10.1186/1472-6963-12-116

Harries, A., Schouten, E.J., Makombe, S.D., Libamba, E., Neufville, H.N., Some, E., Kadewere, G. & Lungu, D. (2007). Ensuring uninterrupted supplies of antiretroviral drugs in resource-poor settings: An Example from Malawi. *Bulletin of the World Health Organization* **85**(2):152–155. doi:10.2471/blt.06.032060.

Halonen, R., Acton, T., Golden, W. & Conboy, K. (2009). DeLone & McLean Success Model As a Descriptive Tool in Evaluating a Virtual Learning Environment. *International Conference on Organizational Learning, Knowledge and Capabilities (OLKC 2009)* ***16.*** https://doi.org/10.4018/jissc.2010040103

Joos, D., Chen, Q., Jirjis, J. & Johnson, K. B. (2006). An Electronic Medical Record In Primary Care: Impact On Satisfaction, Work Efficiency And Clinic Processes. *Annual Symposium Proceedings / AMIA Symposium. AMIA Symposium*, 394–398. https://doi.org/86267 [pii]

Keiser, O., Anastos, K., Schechter, M., et al. (2008). Antiretroviral therapy in resource-limited settings 1996 to 2006: Patient Characteristics, Treatment Regimens and Monitoring In Sub-Saharan Africa, Asia And Latin America. *Tropical Medicine and International Health,* ***13***(7):870–879. https://doi.org/10.1111/j.1365-3156.2008.02078.x

Maria, J. (n.d.). *Evaluation Of Electronic Healthcare Record EHR Systems.* http://www.maria-johnsen.com/electronic-healthcare-record/

Mihăiloaie, C. (2015). *Data Quality Dimensions–From Accuracy To Uniqueness.* https://www.performancemagazine.org/data-quality-dimensions-from-accuracy-to-uniqueness.

Ohuabunwa, E. C., Sun, J., Jean Jubanyik, K. & Wallis, L. A. (2016). Electronic

Medical Records in Low To Middle Income Countries: The case of Khayelitsha Hospital, South Africa. *African Journal of Emergency Medicine* ***6***(1):38–43. https://doi.org/10.1016/j.afjem.2015.06.003

Rahm, E., and Do Hai, H. (2016). Data Cleaning: Problems and Current Approaches. *University of Leipzig, Germany* ***25***(4):11–20. https://doi.org/10.1145/1317331.1317341

Schouten, E. J., Jahn, A., Ben-Smith, A., Makombe, S. D., Harries, A. D., Aboagye-Nyame, F., & Chimbwandira, F. (2011). Antiretroviral drug supply challenges in the era of scaling up ART in Malawi. *Journal of the International AIDS Society* ***14***(SUPPL. 1):1–8. https://doi.org/10.1186/1758-2652-14-S1-S4

Sebastian-Coleman, L. (2014). *Dissecting Data Measurement_ Key Metrics For Assessing Data Quality.* http://searchdatamanagement.techtarget.com/feature/Dissecting-data-measurement-Key-metrics-for-assessing-data-quality

Sittig, D., Kuperman, G. & Fiskio, J. (1999). Evaluating Physician Satisfaction Regarding User Interactions with an Electronic Medical Record System. *Proceedings / AMIA ... Annual Symposium. AMIA Symposium* 400–404. https://doi.org/D005736 [pii]

Van Velthoven, M. H., Mastellos, N., Majeed, A., O'Donoghue, J. & Car, J. (2016). Feasibility Of Extracting Data From Electronic Medical Records For Research: An International Comparative Study 90. *BMC Medical Informatics and Decision Making* ***16***(1):1–10. https://doi.org/10.1186/s12911-016-0332-1

Weiskopf, N. G. & Weng, C. (2013). Methods and Dimensions of Electronic Health Record Data Quality Assessment: Enabling Reuse for Clinical Research. *Journal of the American Medical Informatics Association* ***20***(1):144–151. https://doi.org/10.1136/amiajnl-2011-000681

WHO (2012). Framework and Standards for Country Health Information Systems. *World Health, 2nd Editio*(January), 72. https://doi.org/10.4018/978-1-60566-988-5

Williams, F. & Boren, S. A. (2008). The Role of the Electronic Medical Record (EMR) in Care Delivery Development in Developing Countries: A Systematic Review. *Informatics in Primary Care* ***16***:39–145. https://doi.org/10.1016/j.ijinfomgt.2008.01.016

Wright, G. & Odama, A. (2012). Health Data Ownership and Data Quality: Clinics in the Nyandeni District, Eastern Cape, South Africa. *Engineering Management Research* ***1***(2):146–152. https://doi.org/10.5539/emr.v1n2p146