Development of a systematic method for assessing HIV-indicator data reporting in Kenya

Milka Bochere Gesicho

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2021

UNIVERSITY OF BERGEN

Development of a systematic method for assessing HIV-indicator data reporting in Kenya

Milka Bochere Gesicho

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 24.08.2021

Year: 2021

Title: Development of a systematic method for assessing HIV-indicator data reporting in Kenya

Name: Milka Bochere Gesicho

Print: Skipnes Kommunikasjon / University of Bergen

To my husband Mr. Albert M. Boloji, and my parents

Mr. Robert O. Gesicho and Mrs. Esther K. Gesicho

Scientific environment

This research is a result of the collaboration between University of Bergen-Norway, Moi University-Kenya, and Makerere University- Uganda under the project called Health Informatics Training and Research in East Africa for Improved Health Care (HI-TRAIN). HI-TRAIN project was supported by the Norwegian Programme for Capacity Building Development in Higher Education and Research for Development (NORHED), which is under the Norwegian Agency for Development Cooperation (Norad) that funded this PhD scholarship. I have also benefited as a member in the Norwegian Research School of Global Health, which offered PhD courses that contributed to my course credits as well as travel grants for scientific conferences.

Acknowledgements

This has been an arduous yet fruitful journey, which I could not have achieved without the contribution in one way or the other of everyone acknowledged here.

I begin by expressing my sincere and deepest gratitude to my esteemed main supervisor Prof. Ankica Babic who doubles up as my mentor. I attribute my progressive growth in academic research throughout my PhD journey to Prof. Ankica's adept mentorship, patience, encouragement, and invaluable advice. She tactfully asked tough thought-provoking questions in a manner that made my efforts not seem in vain. Despite her busy schedule, she selflessly and tirelessly sacrificed hours upon hours of her precious time (including late nights) to meet with me, and meticulously scrutinized my analyses and manuscripts. Moreover, her constructive comments paired with her adroitness left me charged up after each session and bouncing with not only inspiration to think outside the box, but also provoked careful thinking in my work from multiple perspectives.

It is in some of these sessions that my knowledge for Design Science Research was birthed and my thirst to understand particular aspects in statistics and machine learning quenched. More so, I am also grateful to Prof. Ankica for enlightening me on the ropes to publishing, something I was clueless about. I also vividly remember the crippling fear that engulfed me as I embarked on this journey, which often begins with a storm as most PhDs can relate. It is the constant affirmations from Prof. Ankica such as "you are smart, believe me", as well as her sense of humor even in the tough moments, which gave me the motivation to push on.

A special thanks also goes to my second esteemed supervisor Prof. Martin Were for his invaluable support and advice on my academic research. Discussions with Prof. Martin were illuminating in a manner that often-invoked light bulb moments. He often engaged his eagle eye and intelligence, critically uncovering errors and bluntly pointed out questionable statements in my manuscripts to achieve the utmost excellence. In return, I received fashionably colorful arrays of red and green on my manuscripts awing to track changes, which spurred determination and thoroughness in my work. I am very grateful for the dedication, time and effort Prof. Martin Were put to ensure impeccable work.

A warm thanks to the HI-TRAIN team for their support in one in one way or another. Much thanks also to Prof. Khalid and Prof. Thorkild for their advice as well as assistance within the project. I would also like to appreciate Tor Bastiansen, and Margun Skarbø for their prompt assistance in financial aspects within the project. To my PhD colleagues Bambi-Alice Mugisha (graduated), Philomena Ngugi, Job Nyameino, Thomas Mwogi, and Moses Musiime, each of you played a different yet special role in making this journey enjoyable and for that, I am deeply grateful.

Special thanks to my husband and best friend Albert Boloji for his consistent encouragement, support and patience throughout my PhD study. He always fired me up when my flames died down whenever it got tough. My deepest appreciation to my loving parents Esther Gesicho and Robert O. Gesicho who always supported, encouraged, and inspired me to achieve the most in life especially in education. As my father would tell me "there are some doors to progress that can only be unlocked with the key of education". I am also grateful to my siblings (Gloria, Susan, Rose and Moses) and my best Jackline Njeri, for their immense emotional support each in their own special way. I also thank my close friends Mercy Njue, Olive Mumbi and John Maneno for those getaways and conversations, which provided an 'escape' when my head was buried in books and a sense of social life.

Finally, and most importantly, I truly thank God without whom, this would not have been possible.

Table of Contents

List of Tables

List of figures

Abbreviations

ART:	Antiretroviral therapy
BS:	Blood Safety
CPC:	Cumulative Percent Completion
CRT:	Care and Treatment
DHIS2:	District Health Information System Version 2
EMRS:	Electronic Medical Record System
HTC:	HIV Testing and Counselling
HIV:	Human Immunodeficiency Virus
HMISs:	Health Management Information Systems
HIS:	Health Information System
LMICs:	Low-and Middle-Income Countries
M&E:	Monitoring and Evaluation
MoH:	Ministry of Health
NGO:	Non-Governmental Organization
RHIS:	Routine Health Information Systems
RR:	Reporting Rate
RRT:	Reporting Rate on Time
UNAIDS:	Joint United Nations Program on HIV/AIDS
PEPFAR:	President's Emergency Plan for AIDS Relief
WHO:	World Health Organization
KDD:	Knowledge Discovery in Databases
DR:	Design Research
DSR:	Design Science Research
SUS:	System Usability Scale

Abstract

Background: In a bid to eradicate the HIV epidemic, Low-and Middle-Income Countries (LMICs) have taken strides in strengthening monitoring and evaluation through building capacity in data collection and data use. As such, the District Health Information System Software version 2 (DHIS2) has been adopted by numerous countries in LMICs for purposes of monitoring and evaluating the progress made towards eradication of the epidemic. Nonetheless, despite a longstanding requirement to report HIV-indicator data from facilities into DHIS2 for many LMICs, few rigorous evaluations exist to evaluate performance of facilities at meeting completeness and timeliness reporting requirements for HIV-indicator data to DHIS2. Hence, the aim of this dissertation was to develop and apply a systematic method that incorporates the use of both quantitative and qualitative research approaches in assessing facility reporting performance over time (2011 to 2018), using completeness and timeliness facility reporting requirements to DHIS2.

Methods: This dissertation was anchored on Design Science Research (DSR) methodology. A DSR process model proposed by Vaishnavi et al. was employed, and consisted of five steps (awareness of problem, suggestion, development, evaluation, and conclusion). The development step is key in design and development of the artifact and consisted of four sub-cycles in this dissertation with each applying different approaches to obtain the various expected outcomes.

Results: A systematic method of assessing facility reporting performance resulted from the combination of four sub-cycles within the development step. This entailed systematic process of data cleaning (sub-cycle 1); application of the resultant clean dataset in evaluation of facility reporting performance (sub-cycle 2); conducting qualitative case study based on facility reporting performance results in sub-cycle 2 (sub-cycle 3); and development of facility reporting performance dashboard comprising visualizations using data and results in sub-cycle 1 and 2 (sub-cycle 4). Results in each of the sub-cycles also varied based on expected outcomes.

Conclusions: The developed systematic method artifact in this dissertation can be of benefit to HIV monitoring and evaluation teams in ministries of health in LMICs as well as other relevant stakeholders.

List of Publications

This dissertation is based on the following papers, which were published in peer reviewed international conference proceedings as well as open access journals (CC BY and CC BY-NC licenses).

I. Gesicho, Milka Bochere, Were, M. C., & Babic, A. (2020). Data cleaning process for HIV-indicator data extracted from DHIS2 national reporting system: a case study of Kenya. BMC Medical Informatics and Decision Making, 20(1), 293.

II. Gesicho, M. B., Were, M. C., & Babic, A. (2021). Evaluating performance of health care facilities at meeting HIV-indicator reporting requirements in Kenya: an application of K-means clustering algorithm. *BMC Medical Informatics and Decision Making*, *21*(1), 6.

III. Gesicho, Milka B., Babic, A., & Were, M. C. (2020). K-means Clustering in Monitoring Facility Reporting of HIV Indicator data: Case of Kenya. In Studies in Health Technology and Informatics Vol. 272, pp. 143–146.

IV. Gesicho, M., Babic, A., & Were, M. (2021). A Retrospective Observational Study of Health Facility Ownership Type and Performance on HIV Indicator Data Reporting in Kenya. In IFMBE Proceedings Vol. 80, pp. 38–44.

V. Gesicho, M. B., & Babic, A. (2021). Identifying barriers and facilitators in HIV-indicator reporting for different health facility performances: A qualitative case study. PLOS ONE, 16(2), e0247525.

VI. Gesicho, M. B., & Babic, A. (2021). Facility Performance Dashboard Instance of HIV-Indicator reporting: Case example of Kenya. In press (Manuscript)

Related Publications

1. Gesicho, Milka B, Babic, A., & Were, M. C. (2017). Critical Issues in Evaluating National-Level Health Data Warehouses in LMICs: Kenya Case Study. Studies in Health Technology and Informatics, 238, 201–204.

2. Gesicho, M., & Babic, A. (2019). Task-based approach recommendations to enhance data visualization in the Kenya national health data warehouse. In IFMBE Proceedings Vol. 68, pp. 467–470.

The published papers are reprinted with permission from the publishers.

1. Chapter 1

1.1 Introduction

1.1.1 HIV Global Burden

The HIV epidemic remains a challenge globally with highest infected numbers of populations found in countries in eastern and southern Africa, which account to 20.6 million of the total number (estimated 37.9 million) of people living with HIV in the word as at 2018 [1]. There were estimated 800,000 new HIV infections and 310,000 AIDs –related deaths as at 2018 [1]. Among the countries that accounted for more than 50% of new infections include: Mozambique (150,000), Tanzania (72,000), Uganda (53,000), Zambia (48,000), Kenya (46,000), Malawi (38,000), and Zimbabwe (38,000) [2]. Nonetheless, new infections have generally declined in eastern and southern Africa by 28%, and deaths by 44% since 2010 [3].

This is due to the numerous strides that have been put in place to achieve ambitious targets such as the Joint United Nations Program on HIV/AIDS (UNAIDS) 90 90 90 targets, whose goal was that by 2020, "90% of all people living with HIV will know their HIV status; 90% of all people with diagnosed HIV infection will receive sustained antiretroviral therapy; and 90% of all people receiving antiretroviral therapy will have viral suppression" in order to end the epidemic by 2030 [4]. This target is yet to be realized despite incidences of HIV/AIDS and mortality decreasing by almost 50% since 2000.

As such low-and middle-income countries (LMICs) have received substantial support from donors and multilateral global organizations to scale-up HIV services such as antiretroviral therapy (ART) , prevention of mother-to-child transmission (PMTCT) of HIV and HIV testing and counselling [5]. The three major international donor organizations that have supported HIV interventions include the United States President's Emergency Plan for AIDS Relief (PEPFAR), the Global Fund to Fight AIDS, Tuberculosis, and Malaria, and the Joint United Nations Program on HIV/AIDS [6]. An estimated $332.00 million dollars was allocated in 1990 by the Development

Assistant for Health (DAH) for purposes of combating HIV/AIDS [6]. The largest source of financial assistance between 1990 and 2015 in descending order with cumulative estimates is as follows: the US government ($67.4 billion), the United Kingdom ($ 6.7 billion), the German government ($ 3.5 billion) and the Gates Foundation ($ 4.2 billion since 1999) [6]. These international donor organizations have also supported health system strengthening in LMICs. Development assistance has however stagnated since 2010 and decreased by 20% between 2012 and 2016 [7]. This has led to the need for LMICs to fill in the gap left by funding agencies, which constituted 85% of all HIV/AIDS expenditure [7].

Hence, there is increased emphasis on using more efficient and cost-effective approaches by LMICs in the continuous efforts to combat the epidemic. In addition, these approaches have the potential to further increase the importance of understanding and using HIV-data by respective countries in order to monitor trends and identify specific needs of geographic regions, hence enabling targeted responses. As such, an advantage that emerges amidst the financial challenges is the need for generation of HIV-data, which promotes ownership and accountability in HIV response and sustainability.

1.1.2 Strengthening of Health Information Systems

A well-functioning health system is essential to achieve better health outcomes [8]. As such strengthening of health systems has been considered salient in LMICs and also received substantial support from domestic budgets as well as international donor agencies [9]. Six building blocks are identified in the World Health Organization (WHO)'s framework for health systems strengthening and include: health service delivery; health workforce; health financing; health information; medical products, vaccines and technologies; and leadership and governance (stewardship) [9]. Each of these building blocks plays a salient role in improving health systems and ultimately health outcomes. Of the six building blocks, health information is considered an integral component of the overall system as it informs decision-making in the other five building blocks [8, 10]. Health Information Systems (HIS) therefore play a critical role in the management of information. Lippeveld et al. define HIS as "*a set of* *components and procedures organized with the objective of generating information which will improve health care management decisions at all levels of the health system*" [8]. Therefore, high quality and timely data generated from a HIS are essential for decision-making [10].

The substantive financial investments put in place by the various international donors with the aim of scaling up HIV services, comes with the need for data in order to provide information that can inform decisions and processes, such as evaluating patient and program progress, as well as guiding allocation of resources. However, data use and demand in LMICs have been considered weak in large part due to data quality issues, rendering some of the countries data rich but information poor [11–13]. As a result, LMICs have witnessed continuous efforts aimed at strengthening quality of collected data through HIS, which have resulted to transition from paper-based medical records to electronic medical records in numerous sites, as well as implementation of HIV national Health Management Information Systems (HMIS) for aggregate data collection [14, 15]. Transitioning to HIS and HMIS has led to collection of large amounts of routine and non-routine health data, which have potential for use in decision-making at facility, county, sub-county and national levels.

However, even with years of existence after implementation of HIS, cases of inadequate use of data are still being reported [16–19]. Data utilization in decision-making, which is also referred to as Data Informed Decision Making (DIDM) is essential in informing policy and advocacy, program design and improvement, program operations, and management. To date, emphasis has been placed on data collection [16, 20] with relatively less attention to DIDM. As a result, more often than not, the collected data is not used sufficiently in strategic planning, advocacy or program development and management [21]. Nonetheless, efforts have been put in place to improve DIDM [22].

1.1.3 HIV Monitoring and Evaluation

Monitoring and Evaluation (M&E) systems, which are regarded as the cornerstone of HIV services, have been established in LMICs to provide high quality strategic information for decision-making [12, 23]. Monitoring and Evaluation are two separate yet interconnected activities. Monitoring focuses on tracking the progress of a project or program through systematic collection and analysis of information using predefined indicators, which reveal their success or failure [24]. Evaluation on the other hand focuses on identifying whether the intended outcome(s) for a project or program was achieved, with the aim of informing areas such as policy formulation, interventions and so on. To inform decision-making and management, M&E systems convert raw data, such as aggregate patient data, to indicators [24]. Good indicators are a fundamental measure used by decision-makers hence providing information on a broad range of conditions [25]. Ministries of Health (MoH), as well as international donor organizations require facilities to report several aggregated indicators as part of M&E programs [26]. In many LMICs, aggregate HIV data reporting is done through the District Health Information Software 2 (DHIS2) [27].

While systems like DHIS2 have contributed to improved availability of routinely generated HIV data for reporting and M&E, significant gaps persist in completeness, timeliness and inaccuracy in these reporting data [11, 12]. Contributing factors are due to issues such as, lack of robust systems for data collection, and analysis, with interoperability as a main challenge [28]; inadequate training, and skills [29–31]; inadequate financing of M&E infrastructure [23]; irrelevant indicators; lack of proper reporting tools; lack of meaningful demand and utilization of data across various stakeholders in various levels and sectors [16, 21]; shortage of staff; and lack of feedback [30]. In addition, reports on countries that have made efforts to implement single national M&E systems reveal that monitoring efforts for programs have resulted to duplicative reporting processes [32]. Part of the challenge is attributed to lack of coordination between multiple donors and implementing partners that support HIV scale-up, with resultant creation of parallel M&E systems [33, 34]. This in part inhibits effective utilization of data for decision-making.

Several other challenges exist in reporting. In many LMIC settings, routine data are collected and recorded in paper based registers, and summary forms, and these have to be entered manually into reporting systems, with multiple potential areas of problems

[32],[35]. Furthermore, in many facilities where Electronic Medical Record Systems (EMRS) have been implemented, there is often lack of robust data exchange to aggregate data systems [28]. Ideally, facilities that have EMRS should be set up to generate aggregate reports that can then be transmitted automatically to DHIS2. However, this is often not the case, and reports from facilities are sometimes still manually entered or uploaded into the DHIS2 system by the facilities [28]. Nonetheless, efforts to ensure data exchange between EMRS and DHIS2 have the potential to improve data quality availability [36].

1.1.4 Approaches in Evaluating Data Quality and Facility Reporting Performance

To improve data quality and availability, various evaluations have been conducted based on the different dimensions of data quality [37–40] However, evaluating data quality is often a complex undertaking due to its multiple dimensions [41]. In addition, the definition of data quality varies based on aspects such as different perspectives, the evaluation approach selected, and whether the evaluation is conducted at national or subnational level [41, 42]. Among the most frequently assessed dimensions of data quality are completeness, accuracy and timeliness [41]. Data quality evaluations conducted within DHIS2 have leveraged various approaches, ranging from desk reviews, data verification to system assessments [37–40, 43]

Moreover, evaluations conducted have often focused on selected regions, periods and specific diseases and indicators within health care. Despite a longstanding requirement to report HIV-indicator data from facilities into DHIS2 for many LMICs, few rigorous evaluations exist that have evaluated completeness and timeliness of reports from these facilities. To our knowledge, there are even more limited studies and reports that provide comprehensive and systematic descriptions steps in data extraction for national HIV reporting, data cleaning process of the reporting data, analyses (assessment) of reporting data to inform performance, and presentation of HIV reporting performance for all facilities nationally.

In this dissertation, a systematic method for assessment was developed and applied. This method incorporated the use of both quantitative and qualitative research approaches for assessing facility reporting performance of HIV data over time, using reporting data from Kenya between (2011 to 2018).

1.2 Research contribution

This dissertation developed and applied a systematic 'method' artifact that can be replicated in settings and countries using DHIS2 as the national data aggregation, reporting and surveillance system, with findings also extensible to other HMIS. Below, key contributions in this dissertation are highlighted.

The first contribution of this dissertation is the development of a generic five step sequence for data cleaning as demonstrated in **Paper I -** Gesicho et al. [44]. Data cleaning is an important aspect when preparing data for analyses or decision-making. Comprehensive, systematic and transparent procedures for data cleaning were presented, that improve on existing processes [44], and which provide insights on the status of data quality in the DHIS2. Our data cleaning process improves the ultimate dataset on which reports are generated. The approach used in data cleaning as well as reporting can also be replicated by researchers and relevant stakeholders.

The second contribution involves implementing innovative approaches to derive new insights from HIV-indicator reporting data. HIV-indicator reporting data lies in HMISs such as DHIS2 with little to no exploration or use. New insights were derived on performance of facilities at meeting completeness and timeliness reporting requirements in DHIS2 over an eight year period using the obtained clean data-set [44]. This also facilitated better understanding of the evolution of reporting performance from the time this national reporting system was implemented. By leveraging on machine learning algorithms as demonstrated in **Paper II -** Gesicho et al. [45] and **Paper III -** Gesicho et al. [46], this dissertation provides a different approach to evaluating reporting performance from previous studies. In addition, this dissertation presents insights derived from statistical analyses on reporting performance based on facility ownership (private and public), as described in **Paper IV-** Gesicho et al. [47]. Therefore, the approaches used in this dissertation to derive insights from data can also be applied by researchers and HIV-monitoring and evaluation teams in ministries of health.

The third contribution entails presentation of findings from a qualitative case study on barriers and facilitators in HIV-indicator reporting based on the different facility performance categories as described in Gesicho et al. [45]. To our knowledge, no study exists that has used a multiple qualitative case study approach to understand barriers and facilitators in HIV-indicator reporting by facilities based on their reporting performance categories as described in **Paper V**- Gesicho et al. [48]. As such, this dissertation demonstrates that insights derived from reporting data can be used in conducting further qualitative inquiries, that further inform areas and approaches for improvement.

The fourth contribution in this dissertation entailed use of good design principles in development of a facility reporting performance dashboard using the results in Gesicho et al. [44, 45]. Data visualization is often advocated in representing health data in LMICs [17, 49], but oftentimes, attention to principles for good design are meagerly addressed in the literature. The visualizations in this dissertation not only provided in part a summary of the results based on the systematic method applied, but also aimed at promoting data exploration and development of insights by various stakeholders in the health sector.

1.3 Justification of the study

When there is a pandemic or epidemic, countries have no choice but to look for ways to manage it and most importantly, eradicate it. Although the terms pandemic or epidemic are often used interchangeably, their meanings vary. The Center for Disease and Control (CDC) defines an epidemic as “a sudden increase in the number of cases of a disease above what is normally expected in that population in that area ”, whereas a pandemic is defined as “an epidemic that has spread over several countries or continents, affecting a large number of people” [50]. While HIV is referred in some studies as a global pandemic [51], the WHO refers to HIV as a ‘global epidemic’[52].

Nonetheless, this goes to show the magnitude of HIV as it affects a large population and posing as a public health concern globally.

Given the prevalence and evolution of HIV, there is need for continuous efforts towards tracking the response to the disease and formulating actions aimed at prevention and treatment. This requires availability of high-quality HIV-data. As previously stated, health information plays a salient role among the six building blocks in health systems strengthening [9]. Systems such as DHIS2 ensure collection of data across health facilities. DHIS2 has contributed in promoting availability of routinely generated HIV-data from health facilities and is being used in over 70 countries [53]. In Kenya, DHIS2 has been in use since 2011 [15].

Good quality aggregate HIV-data from systems such as DHIS2 is necessary for decision-making by MoHs as well as other stakeholders, if the targets aimed at eradicating HIV are to be achieved [4]. Therefore, approaches that evaluate HIV-indicator reporting performance by facilities over time are of benefit as they inform the progress, as well as weaknesses in reporting. This in turn promote formulation of solutions and approaches for improvements as needed. Nonetheless, despite a longstanding requirement to report HIV-indicator data to DHIS2 in LMICs, few rigorous evaluations exist to evaluate performance of various care programs and facilities at meeting reporting requirements such as completeness, and timeliness over time.

This dissertation recognized the importance of decision-makers receiving timely and high-quality data, for purposes such as resource allocations and conducting timely interventions. Countries that have implemented DHIS2 need to evaluate the status of HIV reporting by all facilities in order to identify issues, thus contributing to improvement of M&E efforts of HIV. As such, the systematic method developed in this body of work is a step toward achieving this goal.

1.4 Dissertation summary

In **Chapter 1,** the background for the study is provided, with description of the burden of HIV **(Section 1.1.1)** and importance of strengthening HIS **(Section 1.1.2).** M&E is also described as salient in providing strategic information needed for decision-making **(Section 1.1.3).** A description of evaluation of data quality and reporting performance of facilities as important for ensuring data is used in decision-making is provided in **(Section 1.1.4).** Approaches used by various studies in evaluating data quality are also mentioned, prior to introducing the gap that this dissertation aims to fill **(Section 1.1.4),** as well as the contribution and justification of this dissertation **(Section 1.2 and 1.3).**

In **Chapter 2**, a broad perspective of HIV prevalence is provided, as well as responses to HIV epidemic, with details provided for the study country, Kenya **(Section 2.1.1 and 2.1.2).** In-depth descriptions of HIV programmatic areas assessed in this dissertation are also presented in **(Section 2.1.3)**. Existing limitations of data use, which are interrelated and affected by the quality of data, and mechanisms used in combating these limitations are described in **(Section 2.2).** Use of HIS and HMIS in reporting is introduced in relevance to this dissertation **(Section 2.3)**. As such, relevant details on the DHIS2 in relation to its role in this dissertation are described in **(Section 2.3.1).** EMRS are also briefly described **(Section 2.3.2)**, especially in their relation to HIV-reporting.

In **Chapter 3,** a theoretic background of the knowledge base used in this dissertation is provided. As such, concepts in data quality **(Section 3.1.1)** and data cleaning **(Section 3.1.2)** that contributed significantly in this dissertation, are outlined. Data visualization is also briefly discussed, given its use in representing the results of analyses in order to promote data use and decision-making **(Section 3.1.3)**. Knowledge discovery in databases is also described briefly in relation to its relevance in this dissertation **(Section 3.1.4)**.

In **Chapter 4,** the overall aim of this dissertation is outlined in **(Section 4.1)** as well as the three specific objectives to accomplish the aim. **(Section 4.2).**

In **Chapter 5,** the methodology used is described in **(Section 5.1.1)** as well as the Design Science Research (DSR) process model employed **(Section 5.1.2)**. The research paradigm applied in this dissertation is also described in this chapter **(Section 5.1.3).** The methods used in this dissertation are also described **(Section 5.2),** within the applied adopted DSR process model **(Section 5.3).**

In **Chapter 6**, results are described within the applied DSR process model **(Section 6.2).**

In **Chapter 7**, the key takeaways are discussed based on results and findings in each of the sub-cycles **(Section 7.1)**. Discussions are also presented on research validity and reliability **(Section 7.2)** and secondary analyses of existing data in relation to this dissertation **(Section 7.3).**

In **Chapter 8**, conclusions and the recommended future work are provided.

2. Chapter 2

2.1 HIV in Kenya

2.1.1 HIV Prevalence in Kenya

Kenya is among the countries with the highest HIV epidemic prevalence in the world. According to the 2018 Kenya HIV estimates report, the estimated total number of people living with HIV in 2017 was approximately 1.5 million [54]. The national adult (15-49 years) HIV prevalence was estimated at 4.9% in 2017, with prevalence among women (5.2%) higher than that of men (4.5%) [54]. In addition, HIV prevalence among Key Population (KP) was as follows: Sex workers (29.3%); Men who have sex with men (18.2%); and People who inject drugs (18.3%) [54] .

The HIV epidemic prevalence in Kenya has geographical disparities with some regions having high concentration of prevalence among key populations compared to others. Some of the top 10 counties with the highest adult HIV prevalence as at 2017 in descending order are as follows; Siaya (21.0%), Homa Bay (20.7%), Kisumu (16.3%), Migori (13.3%), Busia (7.7%), Nairobi (6.1%), Vihiga (5.4%), Kitui 4.5% , Kakamega (4.5%), and Kisii (4.4%) [54]. A notable progress is the decrease in new HIV incidences from 101,600 to 52,767 between 2014 and 2018 [55]. In addition, HIV prevalence between year 2014 and 2018 decreased from 6.04% to 4.9 %, and AIDS related deaths from 48,100 (2013) to 23,900 (2017) [55].

2.1.2 HIV Response in Kenya

Kenya has made substantial efforts towards meeting national and global targets with the aim of countering HIV and AIDS. One of Kenya’s national targets was to reduce annual new HIV adult infections by 75% as well as mother-to-child transmission of HIV to less than 5% by 2019 [56].

Challenges such as lack of coordination among the various donors had historically hampered efforts to counter the epidemic. In response to this, donors agreed to a strategy to harmonized their efforts, which led to the “three Ones” principle in September 2003 at the International Conference on AIDS and STIs in Africa (ICASA) held in Nairobi, Kenya [57]. The "three ones" consists of three core principles, namely: one agreed HIV/AIDS action framework that provides basis for coordinating work of all partners (principle I); one national AIDS coordinating authority (Principle II), with a broad multi-sector mandate; and one agreed country level monitoring and evaluation system (Principle III) [57]. In line with the first principle, Kenya developed the 'Kenya AIDS Strategic Framework' (KASF), which is often updated after a targeted period [56]. In line with the second principle, Kenya formed the National AIDS Coordination Council (NACC) as the national AIDS coordinating authority. In line with the third principle, Kenya established a national M&E system [30].

The most recent KASF (2014/15-2018/19) contains eight strategic directions to guide stakeholders in response to HIV with the aim of ensuring comprehensive HIV prevention, treatment and care [56]. The eight strategic directions of KASF include: (1) Reducing new HIV infections; (2) Improving health and wellness of all people living with HIV; (3) Using human rights approach to facilitate access to services for people living with HIV (PLHIV), KPs, and other priority groups in all sectors; (4) Strengthening integration of health and community systems; (5) Strengthening research and innovation to inform the KASF goals; (6) Promoting utilization of strategic information for research and monitoring and evaluation to enhance programming; (7) Increasing domestic financing for a sustainable HIV response; and (8) Promoting accountable leadership for delivery of KASF results and actors [56].

Data-driven decision- making, which is influenced by available and timely high-quality data, is a fundamental aspect in Kenya's response to HIV epidemic. M&E efforts depend on various data sources, which include routine and periodic collection and collation systems. These systems are maintained and supported by various stakeholders involved in HIV response. In addition, routine M&E systems have been established country wide as source for strategic information. This is aimed at promoting data collection at county levels in order to address county specific needs [24].

One of the limitations to Kenya's response to HIV is funding. Although PEPFAR continues to provide support, there is a potential risk of reduced support [58].

Moreover, Haakenstad et al. report in their study that LMICs such as Haiti, Kenya, Malawi and Uganda do not have the capacity to fill the funding gap and are less likely to replace even 10% on care and treatment if funding for development assistance declines [58]. Nonetheless, whilst recognizing the challenges Kenya targeted to increase domestic funding for sustainable HIV response to 50% by 2020 in its 2018 progress report [55].

2.1.3 HIV-Indicator data

Monitoring and evaluation systems have promoted availability of various HIV-indicator data gathered within Kenya's health facilities. The data used in this dissertation was based on the major summary form provided to facilities by the MoH in Kenya for purposes of collecting the HIV-indicator data. The summary form is referred to as '*MOH731-Comprehensive HIV/AIDS Facility Reporting Form'*. This form captures HIV-indicators for six programmatic areas, which are briefly discussed as follows:

I. **HIV Testing and Counselling (HTC)**: HTC is an important area in Kenya's HIV response as it creates awareness of HIV status. This also contributes to the first 90 of UNAIDS 90 90 90 targets, that aim at ensuring 90% of people living with HIV know their serological status [4]. As such, Kenya has made deliberate efforts to ensure HTC coverage among the general population, including introducing self-testing kits.

   HTC indicators need to be collected and submitted in order to monitor the number of people tested, and those who tested positive. These numbers can be disaggregated by age and gender. Of those tested positive, it is also important to know how many were linked to HIV care and treatment. This ensures that no one tested positive is left out of treatment (second 90).

II. **Prevention of Mother-to-Child Transmission of HIV (PMTCT)**: It is important that pregnant women are tested for HIV in order to prevent transmission of HIV to the child and to increase awareness of their HIV status during pregnancy. Moreover, HIV DNA polymerase chain reaction (PCR) tests ought to be performed on children born of HIV positive mothers within a specified time intervals in order to ensure and confirm that the child is not infected [59]. The PMTCT indicator data can be used in assessing the outcome of infants born to HIV-infected women. It is also important that the indicators for PMTCT are collected and submitted in time to guide decision-making that impact service provision, interventions and advocacy.

III. **HIV Care and Treatment (CRT)** : People testing positive for HIV are immediately linked to treatment in Kenya, in accordance to the WHO recommendations [60]. The recommendations states that ART should be initiated to everyone diagnosed with HIV regardless of their CD4 cell count [60]. Of those in treatment, the aim is to ensure that they are virally suppressed. Therefore, it is important that indicators for CRT are collected and submitted on time in order to monitor trends in the proportion of HIV positive persons receiving treatment.

IV. **Voluntary Medical Male Circumcision (VMMC):**
VMMC is conducted in HIV programmes in Kenya as a HIV prevention measure [61]. The VMMC indicator data collected can be used to assess the proportion of males (disaggregated by age), that are being offered and utilize the VMMC services.

V. **Post-Exposure prophylaxis (PEP):** PEP are ARVs administered to those exposed to HIV within 72 hours of exposure as a prevention measure. The indicator data collected can be used in identifying proportion of people exposed to HIV who received and utilized PEP services for HIV prevention.

VI. **Blood Safety:** The Blood Safety indicator was aimed at ensuring adequate supply of blood that has been screened for HIV and other transfusion-transmissible infections through measuring National Blood Transfusion Service's progress. This indicator was replaced with methadone assisted therapy in 2018.

2.2 Limitations and mechanisms used to promote data use

Given that health information is a key pillar in strengthening health systems [8, 10], availability of high quality HIV-data for the various programmatic areas are useful in measuring the progress to achieve the various strategic goals aimed at eradicating HIV. Good quality data supports decision-making, while reliability of data is adversely affected should the quality of the data be questionable. Recent studies also reveal that even after years of implementation of Routine Health Information Systems (RHIS) in LMISs, some of these issues persist and affect use of data [62].

2.2.1 Limitations of Data Use

Among the key issues are sub-optimal data quality, lack of culture of information use, and insufficient capacity, which affect accuracy, completeness, and timeliness of data. These issues are often associated with organizational, behavioral, and technical determinants of routine health information systems [63].

I. Sub-optimal Data Quality

High quality data are the cornerstone of health systems improvements leading to better information, better decision-making and better population health [10, 21, 64]. In LMICs, the HIS are rated as weak accompanied by data quality challenges. Hence, untimely, inaccurate and incomplete data contributes to lack of trust and credibility of information, which leads to inability to use data for evidence-based decisions [16, 20, 65]. In addition, lack of data exchange between systems remains a challenge to data quality and availability [36]. This issue can be depicted during indicator reporting process that requires health facilities to submit reports to a national-aggregate health management information system.

In most cases, the reporting process involves printing electronic data from one system, manual tallying of the data for each indicator and re-entering indicators manually to another system, which increases the chances of errors and delay [36]. Parallel and duplicate reporting channels also contribute to poor data quality in LMICs [17, 66]. In many settings, different stakeholders, including funders and MoH, usually have their own reporting tool, leading to facilities having to submit multiple reports to each system.

Timeliness in reporting results from lack of measures put in place to deter late reporting and leads to laxity in adhering to deadlines [29]. Issues such as stock-out of tools used for data collection, perceptions that data collection tools do not suffice in capturing necessary information (leading to use of notebooks for data collection), and frequent change in tools and indicators, lead to errors in data capture [29]. Data quality improvement feedback is sometimes irregular or delivered in a way that demotivates those collecting primary data or generating the reports [20, 29]. Such demotivation can contribute to data quality issues, given that motivation is a behavioral determinant that affects performance of routine health information systems [67].

II. Lack of culture of information use

Many LMICs settings do not have a culture of DIDM [16]. For instance, despite the high potential of improving patient outcomes by innovatively utilizing data where the data is generated, this is often not the case [22]. In contrast, the data is usually transmitted straight to sub-county, county or national levels for aggregate reporting [20, 30].

Unfortunately, oftentimes data producers have the perception that collection of data is only for reporting purposes [16, 30], reducing their vigilance in collecting high quality data for their own use. Although this perception is changing with data producers being trained on the importance of data, opportunities for them to be involved in data analysis as well as decision-making process in order to gain deeper understanding on what data is used and needed for is still lacking [22]. Hence, it is therefore unfortunate when decisions such as resource allocation, planning, and management of programs such as for HIV are not based on data considering the substantial financial, technical and organization resources channeled towards collecting such data [68].

III. Insufficient capacity

A contributing factor to data quality issues in LMICS is attributed to human resource challenges [23, 69, 70]. Studies reveal limited capacity of data producers to analyze, interpret and present information to decision-makers in LMICs [16, 20, 31]. Lack of human resources and basic competence for validating and recording data also hinders the quality of data. As an example, the routine health information system in Benin attributes its data quality problems to insufficient resources for training staff, poor supervision as well as low staff motivation [31]. This in effect can result to the needs of decision-makers not adequately represented in data collection effort. Another study conducted in Botswana reported varied levels of training and skills of key M&E data management personnel due to lack of M&E courses and training programs [29].

2.2.2 Mechanisms used to promote data use

Local and international health programs have made considerable efforts in recent years to increase use of data for decision-making. To date, projects such as MEASURE Evaluation have made substantial investments in promoting and improving demand and use of data in LMICs [21, 67]. The efforts to increase use of data have aimed at using strategies for improving the quality of data and increasing capacity of data use [71]. The following section describes the application of some of the strategies in countering the barriers of utilizing information for decision-making in LMICs.

I. Improving data quality and use

Improving data quality plays a significant role in strengthening performance of health information systems for decision-making. According to Ledikwe et al, in order to implement strategies to improve data quality, it is important to first identify and understand the strength and weaknesses of underlying factors within the health data management system that influence data quality [29]. Among these factors include; M&E structures, functions, and capabilities; indicator definitions and reporting guidelines; data collection and reporting forms and tools; data management processes; and links with the national reporting system [72].

Several approaches have been employed to improve the quality of data for decision-making. -Some of these approaches include use of routine data quality assessment (RDQA) tools [73]; use of the Performance of Routine Information System Management (PRISM) framework and tools [74] and application of the WHO data quality review toolkit [37, 39]. To improve data use, quarterly workshops have been implemented in countries such as Tanzania to improve the quality of data [17]. The Population Health and Implementation Training (PHIT) partnerships have also contributed to improving HIS decision-making in sub-Saharan Africa [66]. These partnerships are supported by the African Health Initiative launched by the Doris Duke

Charitable foundations to strengthen HIS. For instance, in the Ghana PHIT partnership, data capture has been simplified and reporting streamlined to direct more focus on data quality [66]. In the Mozambique PHIT partnership, there is ongoing feedback on missing data and outliers and assessments on data quality in district and provincial levels [66]. In the Zambia PHIT partnership, standardized protocols are used for data capture with real-time queries in data gaps [66].

II. Improving capacity

Various strategies have been used to address the shortage of human resources within health information systems. Some of these strategies include the use of task shifting to address the shortage of health information personnel; studying staff patterns to ensure data-related tasks are running well; on job training and mentorship [29, 75].

2.3 HIS and HMIS used in routine HIV-indicator reporting

2.3.1 The District Health Information Software Version 2 (DHIS2)

District Health Information Software (DHIS2) is an open-source web-based HMIS implemented in over 70 countries for data collection, reporting and analysis [53]. In Kenya, the DHIS2 utilizes a cloud-based infrastructure and is based on a central server, which simplifies technical support as a change made at any point of the system reflects to all the users [15]. This dissertation utilized HIV- indicator reporting data for health facilities in Kenya and extracted them from the DHIS2, which is the national aggregate data system.

DHIS2 also supports various activities and contains modules for processes such as data management and analytics, which contain features for data visualization, charts, pivot tables and dashboards [76]. In Kenya, DHIS2 was rolled out nationally in the year 2011 [15, 27]. Some of the features of DHIS2 that are of interest in relation to this dissertation are discussed as follows.

(i) Data quality mechanisms within DHIS2

Data quality is an important aspect in health management information systems as it promotes data use for decision-making. As such, various data quality mechanisms have been inbuilt within DHIS2 to ensure that data entered in the systems conforms to the pre-defined measures. Some of these approaches include: (a) validation during data entry in order to ensure data are captured using the right formats and within pre-defined ranges and constraint; (b) user-defined validation rules; (c) automated outlier analysis functions such as standard deviation outlier analysis (reveal data values that are numerically distant from the rest of the data), and minimum and maximum based outlier analysis (reveal data values outside the pre-set maximum and minimum values); and (d) automated calculations and reporting of data coverage and completeness [77].

In this dissertation, focus was particularly on facility reporting completeness and timeliness, which identifies the extent facilities submit the expected number of reports as well as the extent to which these reports are submitted on time. DHIS2 automatically calculates facility reporting completeness and timeliness. This facilitated extraction of reporting data in order to evaluate facility performance at meeting the completeness and timeliness reporting requirements. The variables contained within the summary report in DHIS2 are presented in Table 1.

Table 1. Variables within DHIS2 summary report

Organisation unit	Name of the organisation unit (health facility was used as the organization unit)
Actual reports	Actual reports that have been completed (submitted)
Expected reports	Expected number of reports that should have been completed (submitted), based on the organization units that have been *assigned* to the data set
Reporting Rate	Percentage of the expected reports that have actually been submitted
Actual reports on time	Number of the reports that have been completed (submitted) on time
Reporting Rate on time	Percentage of the expected reports that were submitted on time

These variables facilitate the calculation of facility reporting completeness (referred to as reporting rate [RR] in DHIS2) and facility reporting timeliness (referred to as reporting rate on time [RRT] in DHIS2). The RR and RRT are calculated as presented in Table 2.

Table 2. Calculation of Reporting Rate and Reporting Rate on Time

Variables	**Formula**
Reporting Rate (Completeness)	$\mathbf{RR} = \frac{\text{Number of actual reports submitted}}{\text{Expected number of reports}} \text{ X } 100$
Reporting Rate on Time (Timeliness)	$\mathbf{RRT} = \frac{\text{Number of actual reports submitted on time}}{\text{Expected number of reports}} \text{ X } 100$

The DHIS2 quality tool has also been developed to identify errors within the data in order to determine the next appropriate action [78]. The tool enables assessment of various data quality elements such as completeness and timeliness, consistency over time, analysis of consistency between indicators, consistency over time, analysis of missing data and outliers, completeness of reporting, and internal consistency of reported data [78].

(ii) Data warehousing in relation to DHIS2

Bill Inmon defines a data warehouse as "*subject-oriented, integrated, time variant and non-volatile collection of data in support of management's decision-making process*" [79]. Moreover, Biehl posit that a data warehouse is not a hardware or software product that can be bought off the shelf for purposes of providing strategic information [80]. On the contrary, it is a computing environment that provides users with strategic information and should be focused on what users need rather that how to collect more data [80]. Based on these descriptions, DHIS2 can somewhat fit into the description of a data warehouse.

Hence, DHIS2 is used in collecting aggregate level routine data for decision-making as illustrated in Figure 1, were reports are submitted to DHIS2 from various health facilities. It is worth noting that HIV-reports are among the many reports being submitted to DHIS2 by health facilities.

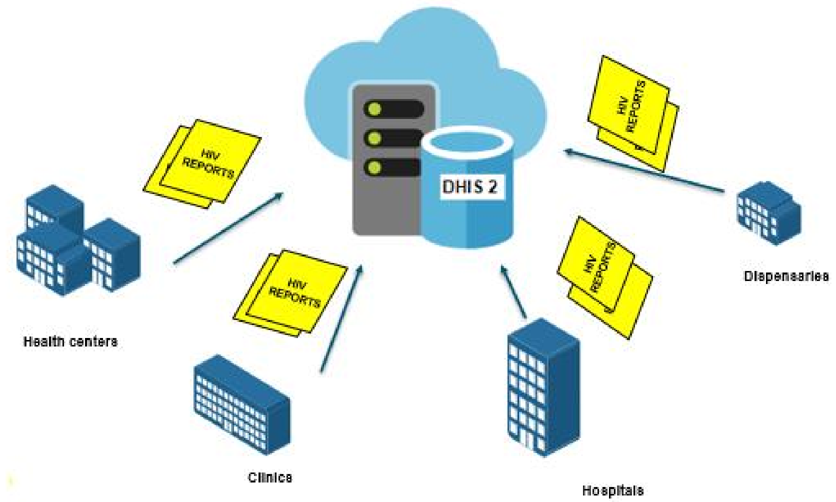

Figure 1. Submission of reports by various health facilities to DHIS2

Moreover, DHIS2 provides functionalities for data entry and validation (data quality mechanisms), and analysis and presentation of data using tools such as charts , maps, pivot tables and dashboards [81].

Data marts in DHIS2 contain data aggregated in time dimension (over different periods), space dimension and indicator formulas (for example mathematic expressions) [81].

2.3.2 Electronic Medical Record Systems (EMRS) in Reporting

Electronic Medical Record Systems (EMRS) have been implemented in numerous health facilities in developing countries in order to promote better healthcare and health services through data collection for use in decision-making [82]. These EMRS mainly support HIV programs funded by institutions such as PEPFAR, which also support EMRS implementation and use [26]. EMRS can be categorized as proprietary or open-source [82]. In Kenya four EMRS falling into either the proprietary or open-source category were selected based on a standardization assessment conducted in 2011 [83]. These were implemented between the year 2012 to 2014 in over 600 health facilities primarily for use in HIV care [83]. These include, Funsoft, C-PAD, OpenMRS and IQ-care [83]. A weakness commonly identified in selection of EMRS is the ability to electronically transmit aggregate information to DHIS2 [84].

Although the existing EMRS in Kenya scored a high mean score (71.8%) in health information and reporting, they attained the lowest mean score (14.3%) in interoperability [84]. In order to deal with the challenge of interoperability between DHIS2 and EMRS, automatic indicator reporting has been explored as a potential solution based on evidence from feasibility studies conducted [28, 36]. These studies also revealed improved quality of data in automatic indicator reporting [28, 36].

In as much as interoperability remains a problem, EMRS in Kenya are still able to generate MoH required reports, which are salient in indicator reporting as information can be retrieved electronically rather than searching among stacks of paper-based records. Currently, Kenya is transitioning all EMRS implementation to one open-source EMRS (KenyaEMR), whose platform is derived from OpenMRS. It is planned that KenyaEMR will have automatic data exchange with DHIS2. This has the potential to improve routine reporting of HIV indicator reporting from facility level to national level.

3. Chapter 3

3.1 Theoretical background

In this chapter, description is provided on the various concepts that informed the development of the systematic method in this dissertation.

3.1.1 Data Quality

Data quality is a complex multi-dimensional concept. Nonetheless, there is no consensus on the standard definition of data quality. The International Standards Organization defines data quality as "the totality of features and characteristics of an entity that bears on its ability to satisfy stated and implied needs" (ISO 8402-1986, Quality Vocabulary). As such, definitions for data quality revolve around the concept of "fit-for-use", and has been largely adopted by researchers whereby data quality is defined in the context of *data that are fit for intended purpose* [85–87]. Furthermore, there exist multiple data quality dimensions in the literature that often seem to overlap, and contain varied definitions depending on context [86, 88]. Some of the data quality dimensions listed by the World Health Organization (WHO) include: accuracy and validity, reliability, completeness, legibility, currency and timeliness, accessibility, meaning or usefulness, confidentiality and security [89]. Nevertheless, the most frequently assessed attributes of data quality especially in information systems in healthcare include completeness, accuracy and timeliness [41].

Various efforts have been made to develop frameworks that categorize important aspects for understanding data quality [85, 87, 90]. Wang and Strong categorized sixteen salient data quality dimensions into: intrinsic data quality, contextual data quality, representational and accessibility data quality [86]. Intrinsic data quality focus on features that are inherent to data itself such as accuracy and believability [86]. Contextual data quality focuses on features that are relevant in the context for the task for data use such as value-added, appropriate amount of data, and relevancy [86]. Representational and accessibility data quality highlights features that are salient within the role of the system such as interpretability, representational consistency, and accessibility [86].

Shanks and Price propose a theoretic-based framework using semiotic theory, database integrity theory as well as mapping cardinalities in developing an information quality framework [90]. As such, they categorize information quality as syntactic quality (degree to which stored data conforms to database rules (metadata)), semantic quality (degree to which stored data corresponds to external metadata) and pragmatic quality (degree to which stored data are suitable for a given purpose) [90]. Shanks and Corbitt further extended this framework by including social data quality, which entails shared understanding of meaning [87]. Kahn et al. on the other hand perceive information quality from both the product quality and service quality standpoints [91]. These were categorized into four levels of information quality, which include sound information and useful information under product quality, and usable information and effective information under service quality.

Based on the proposed frameworks, data quality can be perceived from two broad categories, which include objective perspective and subjective perspective. As such, subjective data quality assessments focus on the users' data needs and experience, whereas objective measures focus on assessment of conformance to pre-defined requirements and specified integrity rules. For instance, using the framework proposed by Shanks and Price, the syntactic and semantic categories lie on the objective standpoints whereas the pragmatic category lies on the subjective standpoint [90]. On the other hand, the product quality, and service quality categories proposed by Kahn et al. are also based on objective and subjective viewpoints respectively [91]. Furthermore, Shanks and Corbitt categorize data quality into intrinsic characteristics (objective) and extrinsic characteristics (subjective) [87].

In addition, there are a number of frameworks used in assessment of data quality in health information systems, which can be utilized by countries with DHIS2. The Data Quality Review (DQR) tool developed in collaboration with WHO, Global Fund, Gavi, USAID, and MEASURE Evaluation provides a standardized approach that aims at facilitating regular data quality checks [42]. As such, this tool provides approaches for conducting desk reviews, data verification or system assessment in conducting performance assessments in HMIS [42]. Some of the data quality dimensions used as indicators of performance, which are comprised in this tool include completeness, timeliness, internal consistency of reported data, external comparisons, and external consistency of population data [42]. Performance assessments conducted within DHIS2 have leveraged on the aforementioned approaches, ranging from desk reviews, data verification to system assessments [37–40], [43] Other tools for routine data quality assessments include the MEASURE Evaluation Routine Data Quality Assessment Tool (RDQA) [92] and WHO/IVB Immunization Data Quality Self-Assessment (DQS) [93].

3.1.2 Data Cleaning

Chapman defines data cleaning as "the process used to determine inaccurate, incomplete, or unreasonable data and then improving the quality through correction of detected errors and omissions [94]." Data cleaning is also one of the techniques comprised in data preparation that is concerned with analyzing raw data in order to obtain quality data for purposes such as analysis and data mining [95]. It is also a salient component in the Knowledge Discovery of Data (KDD) process [96]. Data quality problems due to replicated entries, missing information, or other invalid data are common in integrated data sources such as data warehouses [97]. In order to improve data quality, 'dirty' data needs to be cleaned. Therefore, the need for data cleaning increases significantly, when multiple data sources need to be integrated. Furthermore, data cleaning is essential in research studies in order to provide quality assurance, and is a determinant of study validity as advocated in other studies [98].

A substantial body of works exists on how to clean data [98–100]. Quantitative approaches (statistical methods such as outlier detection to identify these errors) [101] and qualitative approaches (use patterns, constraints, and rules to detect errors) [99] have also been employed in data cleaning. Moreover, there also exists a number of automated data cleaning tools that such as ARKTOS, AJAX, FraQL, Potter's Wheel and IntelliClean, which remove anomalies from data [102]. It is worth noting that data cleaning approaches largely depend on the data especially with the existing myriad of data quality problems [103], as well as the differences in data and its uses. Within HIS, Dziadkowiec et. al for instance employed Kahn et. al.'s framework to clean data extracted from a relational database of an Electronic Health Record system (EHRS) [104]. In order to prepare secondary data for analysis, Miao et. al proposed a data quality assessment and cleaning framework that can be used for EHR data. The aforementioned approaches largely focus on issues based on syntactic quality (conformance to database rules) and semantic quality (correspondence or mapping to external phenomena) [87].

In this dissertation, data cleaning was conducted in order to improve the quality of the data extracted from DHIS2 for purposes of secondary analyses. As such, data quality was defined from the basis of "fitness-for use" and addressed from the contextual perspective. Based on the data set extracted, Van den Broeck et al. 's conceptual framework [98], informed the data cleaning approach used as it provided a systematic data cleaning guideline that is able to be tailored towards cleaning of the data extracted from DHIS2.

3.1.3 Data Visualization

The implementation of HIS and HMIS in the health sector in LMICs has resulted in large amounts of data gathered within these systems [83, 105]. Hence, this sparks the need for data visualization for purposes of representing these data in ways that can promote decision-making [106]. As such, efforts have been made in LMICs to build capacity for data visualizations so as to promote data use for decision-making [19].

Within DHIS2, data visualization features have been added to enable users to represent data in a dashboard format [107]. Visualizations within DHIS2 include column chart, stacked column chart, bar chart, stacked bar chart, line chart, area chart, pie chart, radar chart and speedometer chart [76]. Maps can also be used within DHIS2 to visualize data within a region [107]. Nonetheless, as much as health information dashboards and data visualizations have become popular in LMICs [49], a concern that illuminates is insufficient skills for analysis and data visualization among users [49, 107], training, and maintenance costs [49, 108], and lack of adherence to visualization design principles [107, 109].

There exist various principles that guide the development of data visualizations as well as information dashboards. Some of these principles leverage on cognitive aspects in order to prevent information overload and promote deriving of insights. The WHO for instance has developed best-practice recommendations for graphical presentations of health information [110]. Few also proposed 13 mistakes that ought to be avoided when designing dashboards that is: exceeding the boundaries of a single screen; supplying inadequate context for the data; displaying excessive detail or precision; choosing a deficient measure; choosing inappropriate display media; introducing meaningless variety; using poorly designed display media; encoding quantitative data inaccurately; arranging the data poorly; highlighting important data ineffectively or not at all; cluttering the display with useless decoration; misusing or overusing color; and designing an unattractive visual display [109]. Sedig and Parson also proposed a pattern-based framework that aid in design thinking of visualizations for human-information interaction [111]. The framework enables developers in designing visualizations that promote representation of various facets of data at a glance, as opposed to one or two facets as is with the case of simple visualizations such as bar charts [112].

Bearing in mind the importance of utilizing good design principles for data visualizations, this dissertation utilized the various design principles in developing data visualizations for results analyzed using data extracted from DHIS2. The visualizations aim to enable monitoring and evaluation teams within the ministries of health as well as various stakeholders to identify reporting completeness and timeliness performance of facilities, and counties at various time periods.

3.1.4 Knowledge Discovery in Databases (KDD)

Fayyad et al. describe Knowledge Discovery in Databases (KDD) as the “overall process of discovering useful knowledge from data” [96]. Therefore, Fayyad et al. define KDD as “the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data” [96]. The KDD process involves an iterative sequence of various steps, which end with knowledge discovery and its implementation.

This dissertation entailed discovering knowledge from a large amount of data. As such, the various steps within the KDD process were applicable as follows:

1. **Obtaining prior knowledge and understanding of the application domain:** The application domain was the Ministry of Health, where HIV-indicator reporting data was used in assessing facility reporting performance. It was presumed that prior knowledge of the application domain will enable understanding of how to tackle decisions in the various KDD steps such as the algorithms to be used, and data representation.
2. **Selecting and creating a target data set:** This step involves identifying the data that is available based on the aim, and integrating it with necessary additional data to create a data set [113]. The data required for the aim for this dissertation was available in the District Health Information Software Version 2 (DHIS2). Therefore, health facility HIV-reporting data on all facilities in the 47 counties in Kenya was extracted from DHIS2. These data were then integrated with the necessary additional data to create the initial data set [44].
3. **Data cleaning and preprocessing:** In this dissertation, this process involved removing "dirty" data, using a systematic data cleaning approach [44].
4. **Data transformation:** This step involves developing better data in preparation for data mining, using methods such as dimension reduction, and attribute transformation [113]. This process is dependent on the aim of the task.
5. **Selecting the appropriate data-mining task:** This step involves deciding on the type of data mining to use (such as classification, regression and clustering), depending on the goal for data mining. These goals can be either prescriptive or descriptive. This dissertation employed a descriptive approach as the tasks performed in this study were aimed at obtaining knowledge on health facility reporting from the data set, such as identifying facilities that perform well and those that perform poorly.
6. **Choosing the data mining algorithm (s) to use:** This step involves identifying and selecting a suitable method to be used in searching patterns within the data [113]. The data-mining algorithm used predominantly in this study was k-means clustering.

7. **Implementing the selected data-mining algorithm:** The K-means clustering algorithm was used in identifying homogenous groups within the data to assess completeness and timeliness in facility HIV-indicator reporting.
8. **Pattern Evaluation and Interpretation:** This step involved deriving meaning from the patterns discovered and may require going back to the previous steps. Interpretation depended on the outcomes for facility reporting timeliness and completeness.
9. **Using the discovered knowledge:** This step involved use of the various knowledge representation techniques such as data visualization to present mined knowledge to users for purposes of deriving insights. The knowledge discovered concerning facility reporting performance was used in conducting further qualitative assessments. The knowledge discovered can also be used by monitoring and evaluation teams within the Ministry of Health to inform solutions to identified issues.

4. Chapter 4

4.1 Overall Aim

The overall aim of this dissertation is to develop and apply a method for conducting a rigorous evaluation of HIV-indicator data reporting by health facilities to DHIS2 using completeness and timeliness as reporting-performance indicators. This assessment is of relevance as it would be of benefit to the Ministry of Health, who have mandated reporting completeness and timeliness as some of the key requirements in DHIS2 reporting. The resultant method used in assessment contributes to strengthening HIV monitoring and evaluation efforts by identifying progress as well as weaknesses that require improvement.

4.2 Specific objectives

Specific Objectives

To achieve the overall aim, the objectives outlined below were implemented:

Specific objective 1: Evaluate performance of health care facilities at meeting HIV-indicator reporting requirements

Primary outcome: The primary outcome of interest consisted of identifying the performance in reporting by health facilities over time (2011-2018), with facilities put into various performance clusters and performance evaluated in the various programmatic areas.

Secondary outcome 1: The first secondary outcome was to obtain a clean data set prior to conducting analyses for specific objective 1.

Secondary outcome 2: The second secondary outcome was to determine the relationship between facility ownership and performance in reporting completeness and timeliness.

Specific objective 2: Identify factors contributing to success and deficiencies in HIV-indicator reporting

Primary outcome: Identify factors influencing timeliness and completeness of HIV-reporting from facilities to DHIS2.

Specific objective 3: Develop visualizations for reporting completeness and timeliness to aid in decision-making

Primary outcome: To develop a performance dashboard with visualizations that represent performance of facility reporting based on timeliness and completeness in reporting.

5. Chapter 5

5.1 Methodology

In this section, a description of the methodology used in this dissertation is provided. It was also considered salient to distinguish methodology and methods on the onset as these terms are often used interchangeably. Moreover, distinguishing these terms promotes better understanding of the dissertation.

Based on common definitions of methodology in their review, Mackenzie and Knipe define methodologies as an “overall approach to research linked to the paradigm or theoretic framework” [114]. In addition, they refer to methods as “systematic models, procedures or tools used for collection and analysis of data” [114]. In this dissertation, Design Science Research was selected as the methodology, as it enables development of artifacts relevant in providing solutions to problems in industry or academia.

5.1.1 Design Science Research

Hevner and Chatterjee define design science as a research paradigm that results to contribution of knowledge to the scientific body, and entails creation of innovative artifacts with the aim of answering questions that are relevant to human problem [115]. By observing the environment and identifying problems, a researcher can demonstrate the need to formalize or develop a relevant artifact that seeks to provide a solution to the problem. Design Research (DR) is distinguished from Design Science Research (DSR) in that whereas DR involves researching on design, DSR involves learning through building [116]. Thus, DSR entails bridging the gap between theory and practice, by prescribing solutions and designing IT artifacts with focus on relevance in the application domain [115].

These artifacts are categorized as constructs, models, methods, instantiations and better design theories [117, 118] The solutions developed from design science research need not be optimal but should be satisfactory [119]. The artifact in this dissertation are under the ‘method’ category, inform of descriptions on approaches for “best practice”, as they provide guidance on how problems can be solved [119].

5.1.2 Design Research Science (DSR) Process Frameworks

There are various frameworks for design science research process, which depict the various phases in developing an artifact [117–120]. The aim of these frameworks is to provide researchers with valuable guidelines to consider during artifact development. Peffers et al. for instance proposed a framework that presents various angles in which research can be initiated that is, problem-centered initiation, objective-centered initiation, design and development initiation and context initiation [120]. This framework consists of the following phases: problem identification and motivation, definition of objectives, design and development, demonstration, evaluation and communication [120].

Hevner et al. on the other hand proposed a framework consisting of three cycles that is relevance cycle, design cycle and rigor cycle, which take place within the environment and knowledge base as presented in Figure 2 [119].

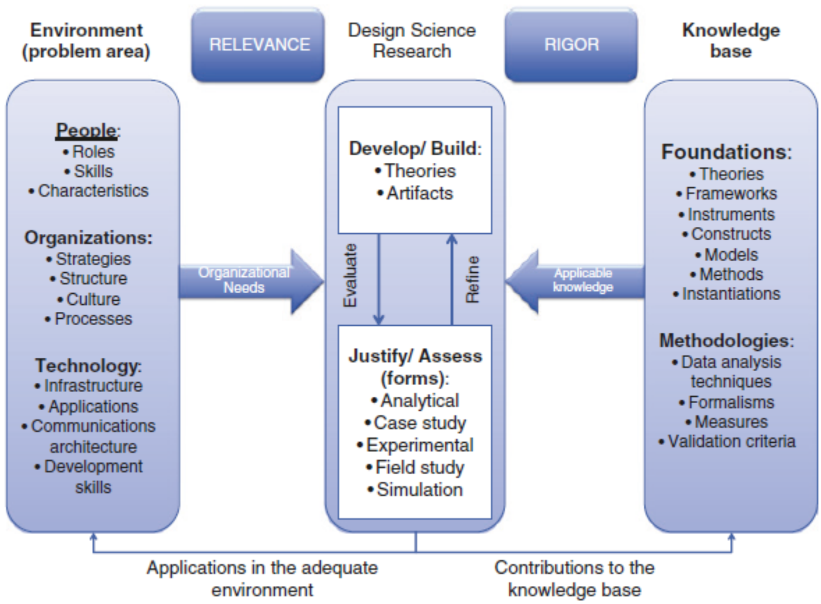

Figure 2. Framework for design science research (Source: Hevner et.al [119]).

The relevance, design and rigor cycles are briefly discussed as follows;

Relevance cycle: The environment consists of people, the organization, and technology [119]. Organizations should be able to use the knowledge generated during development of an artifact in order to solve practical problems. To be relevant to this environment, research must address the problems faced and the opportunities offered by the interaction of people, organizations, and information technology [121].

Design cycle: This cycle obtains components from relevance and rigor cycle to develop the artifact, and entails a number of iterations performed during construction of the artifact [121].

Rigor cycle: Hevner et al. define the knowledge base as "the location where raw material for development of new research and new artifacts are obtained" [122]. As such, this cycle entails selection and application of appropriate methods as well as theories for purposes of developing and evaluating the artifact. Research also needs to be valid and reliable [121].

Hevner et al. also proposed seven guidelines that are essential when conducting design science research: (i) Design as an artifact; (ii) Problem Relevance; (iii) Design Evaluation; (iv) Research Contributions; (v) Research Rigor; (vi) Design as a Search Process; and (vii) Communication of Research [122].

Vaishnavi et al. also proposed a DSR process model that is similar to existing models by extending Takeda et al.'s work [123], and provide detailed description of design science knowledge generation [116]. Moreover, Vaishnavi et al. pointed out that the key focus of DSR should be knowledge contribution as illustrated in their framework [116]. This model consists of the following phases: awareness of problem, suggestion, development, evaluation and conclusion (See Figure 3), which are described further in subsequent **Section 5.3**. It is worth noting that the DSR processes contain somewhat similar phases and can be selected by a researcher based on their suitability in providing guidance with regard to construction of the artifact. The DSR process used in this dissertation was informed by the model proposed by Vaishnavi et al. as it emphasizes on the detailed process for generating design science knowledge [116].

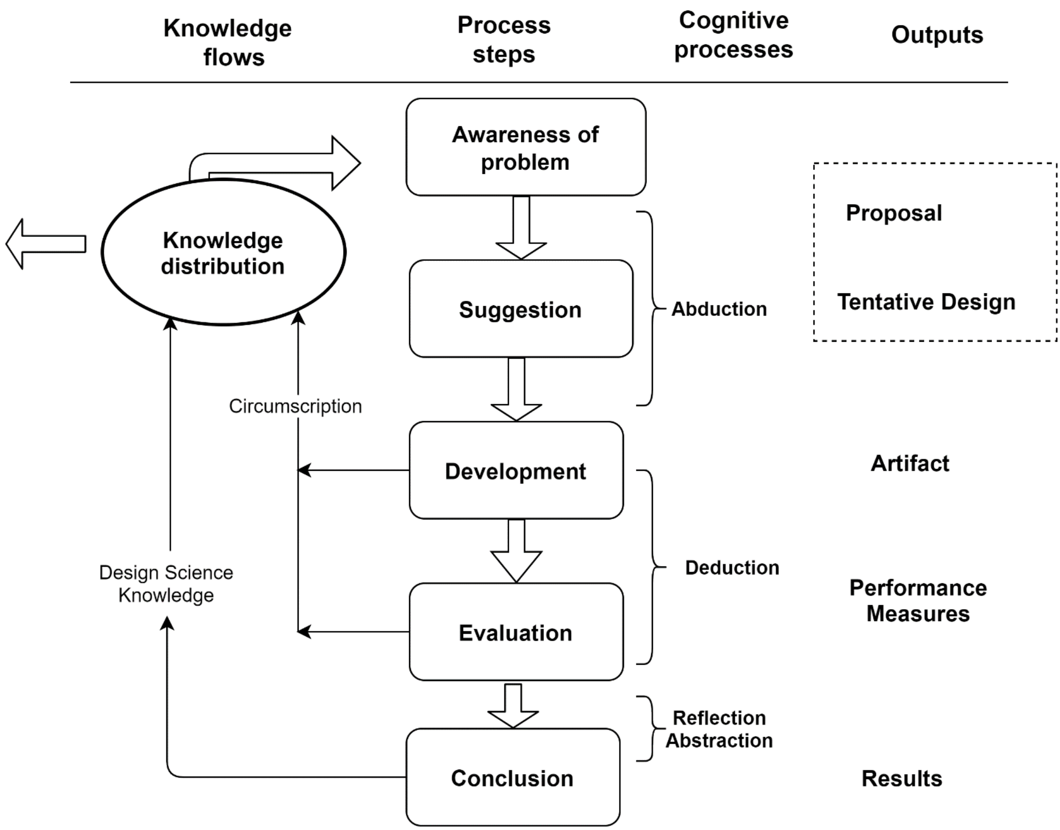

Figure 3.Design Science Research Process Model. (Source: Vaishnavi et al. [116].)

5.1.3 Research Paradigms

In this section, a brief description of the main paradigms in research are provided These include positivism, post-positivism, interpretivism, and pragmatism.

Positivist paradigm

This paradigm is at the objective end of the continuum as reality is thought to be independent of beliefs and perceptions, and therefore it can be obtained through observation and experiments [124]. It is often referred to as a 'scientific method' and its data collection and analysis procedures are often aligned with quantitative methods

[125]. Research conducted under this paradigm is value-free and seeks generalization when measuring social phenomena.

Post-positivism paradigm

Post-positivism on the other hand challenges the idea of a single reality or absolute truth especially when studying social phenomenon [126]. Hence, reality is not viewed as rigid and does not exist within a vacuum [125]. The definition for post-positivism provided by O'Leary ("what might be truth for one person or cultural group may not be truth for another") seems to describe the aspect of multiple realities [127], which in a sense leans towards the constructivism/interpretivism paradigm.

Interpretive/Constructivism paradigm

This paradigm falls at the subjective end of the continuum where reality exists in many forms and is made, not found [124]. Its data collection and analysis procedures are often aligned with qualitative methods [114]. Quantitative data may be used to supplement the qualitative data [114]. Researchers using the constructivism paradigm develop theory inductively throughout the research, which is not the case with post-positivism research as it begins with theory and collects data that either supports theory or not [128].

The pragmatic paradigm

This paradigm uses a mixed-method approach and is problem-centered [128]. As such, it is neither on the positivism or interpretivism research philosophies as it emphasizes on viewing research as a continuum. Hence, the objective and subjective perspectives comprised within the positivism and interpretivism paradigms respectively are viewed as not mutually exclusive [126]. Qualitative and quantitative data are used to better understand social reality [126].

Philosophical assumption applied

Hevner contends that design science research holds a pragmatic philosophy due to the emphasis on relevance, which entails identifying and defining a problem in the environment and making a clear contribution that seeks to provide solutions [121]. Hevner and Chatterjee also describe critical realist as a philosophical underpinning that can be used in design science research [115]. The design science research used in this dissertation leans on the pragmatic paradigm.

5.2 Methods

In this section, description is provided on the summary of methods used in developing the artifact in this dissertation as well as resulting papers (see Appendices) from the specific outcomes outlined for each objective (see Table 3). A detailed description of methods is found in the respective publications (see Appendices).

Table 3. Summary of methods and resultant papers in the dissertation

Paper	Methods	Sample Size	Expected Outcome
I	— Generic five-step data-cleaning sequence — Descriptive statistics (Frequency tables and crosstabulations) — Non-parametric tests (Friedman ANOVA, Wilcoxon Signed Rank Test)	— 93,179 facility-records from 11,446 health facilities	— Clean dataset suitable to be used for subsequent secondary analyses
II and III	— A retrospective observational study — K-means clustering	— **Counselling and Testing** =39,480 records — **Prevention of Mother to Child Transmission** =34,016 records — **Care and Treatment** =18,394 records — **Voluntary Medical Male Circumcision** =277 records — **Post-Exposure Prophylaxis** =9,283 records — **Blood Safety** =202 records	— The primary outcome of interest consisted of identifying the performance in reporting by health facilities over time (2011-2018) considering completeness and timeliness, with facilities put into various performance clusters and performance evaluated in the various programmatic areas.
IV	— A retrospective observational study — Non-parametric test (Mann Whitney U test)	— **Records varied by year (2011-2012), and by programmatic area**	— Determine the relationship between facility ownership and performance in reporting completeness and timeliness.
V	— Qualitative Case study	— **13 health facilities** — **13 participants**	— Identify factors influencing timeliness and completeness of HIV-reporting from facilities to DHIS2
VI	— Usability testing — Think aloud session — Interview	— **5 participants (IT experts)** — **1 participant (practitioner)** — **1 participant (researcher)**	— Develop a reporting performance dashboard

5.3 Application of Design Science Research (DSR) Process Model

The phases of the DSR process model proposed by Vaishnavi et al. [116] are described as follows:

5.3.1 Awareness of problem

This phase entails identification and defining of the research problem, which can originate within the environment (industry) or academia resulting to outputs such as research proposal. The problem identified in this dissertation originates within the environment (Ministry of Health) and articulated gaps on previous work concerning assessment of facility reporting to DHIS2.

5.3.2 Suggestion

This phase entails identifying or presenting a tentative suggestion for the problem solution, and immediately follows the proposal, with its connection denoted by the dotted line surrounding its output in Figure 3 (Proposal and Tentative Design) [116]. The problem solution suggestions are derived abductively from the existing knowledge base for which the problem exists. Hevner et al. lists examples of foundations and methodologies that can be derived from the knowledge base (see Figure 2) [119]. The tentative suggested solution is then implemented in creating an artifact in the development phase [116]. In this dissertation, a formal proposal was developed describing the identified problem as well as the proposed solution.

5.3.3 Development

This is a creative phase and entails design and development of the artifact, which can present a solution that is complete or one that needs further iterations [116]. The development phase can be likened to the design cycle in Hevners et al.'s DSR process model [119], which consists of iterations conducted during construction of the artifact.

As such, in this phase focus was on the steps used in creating the artifact. The development stage contained four sub-cycles illustrating the various steps. Various approaches were used within the four sub-cycles in the development phase. These are described below as well as the associated publications resulting from each sub-cycle.

I. Sub-cycle 1 – (Resultant publication: Paper I)

HIV-indicator data gathered from 2011 to 2018 were extracted from the DHIS2 in Kenya to conduct secondary analyses, which entailed evaluation of health facilities at meeting reporting requirements (reporting completeness and timeliness) mandated by the Ministry of Health (MoH).

The HIV-data was extracted from health facilities in all 47 administrative counties in Kenya (Figure 4). All the counties report a range of healthcare indicator data from care facilities and settings into the DHIS2 system.

Figure 4. Map of Kenya representing various counties.

Prior to using the data for analysis, it was salient for a data cleaning process to be conducted. The data cleaning process not only aimed at obtaining a final clean data set, but also report on methods and results of a comprehensive, systematic and replicable data cleaning approach applied on the HIV-data.

The reporting requirements of interest included facility reporting completeness and facility reporting timeliness, which were extracted per facility per year. Other data variables extracted include HIV-indicator data elements per facility per year. Therefore, these data were extracted for six programmatic areas for all care facilities in all counties in Kenya from 2011 to 2018: (i) HIV Counselling and Testing (HCT); (ii) Prevention of Mother-to-Child Transmission (PMTCT); (iii) Care and Treatment (CrT); (iv) Voluntary Medical Male Circumcision (VMMC); (v) Post-Exposure Prophylaxis (PEP); and (vi) Blood Safety (BS). Hence, 93,179 facility-records from 11,446 health facilities were extracted from year 2011 to 2018. A summary of creation of the data set is illustrated in Figure 5.

It is well recognized that real-world data like that found in DHIS2 are often “dirty” consisting of issues such as; incomplete, inconsistent, and duplicated data [129]. Failure to detect data quality issues and to clean these data can lead to inaccurate analyses outcomes [130]. Therefore, a systematic and replicable data cleaning approach was used on the extracted data and employed semi-automatically within a generic five-step data-cleaning sequence that was developed and applied in cleaning the extracted data.

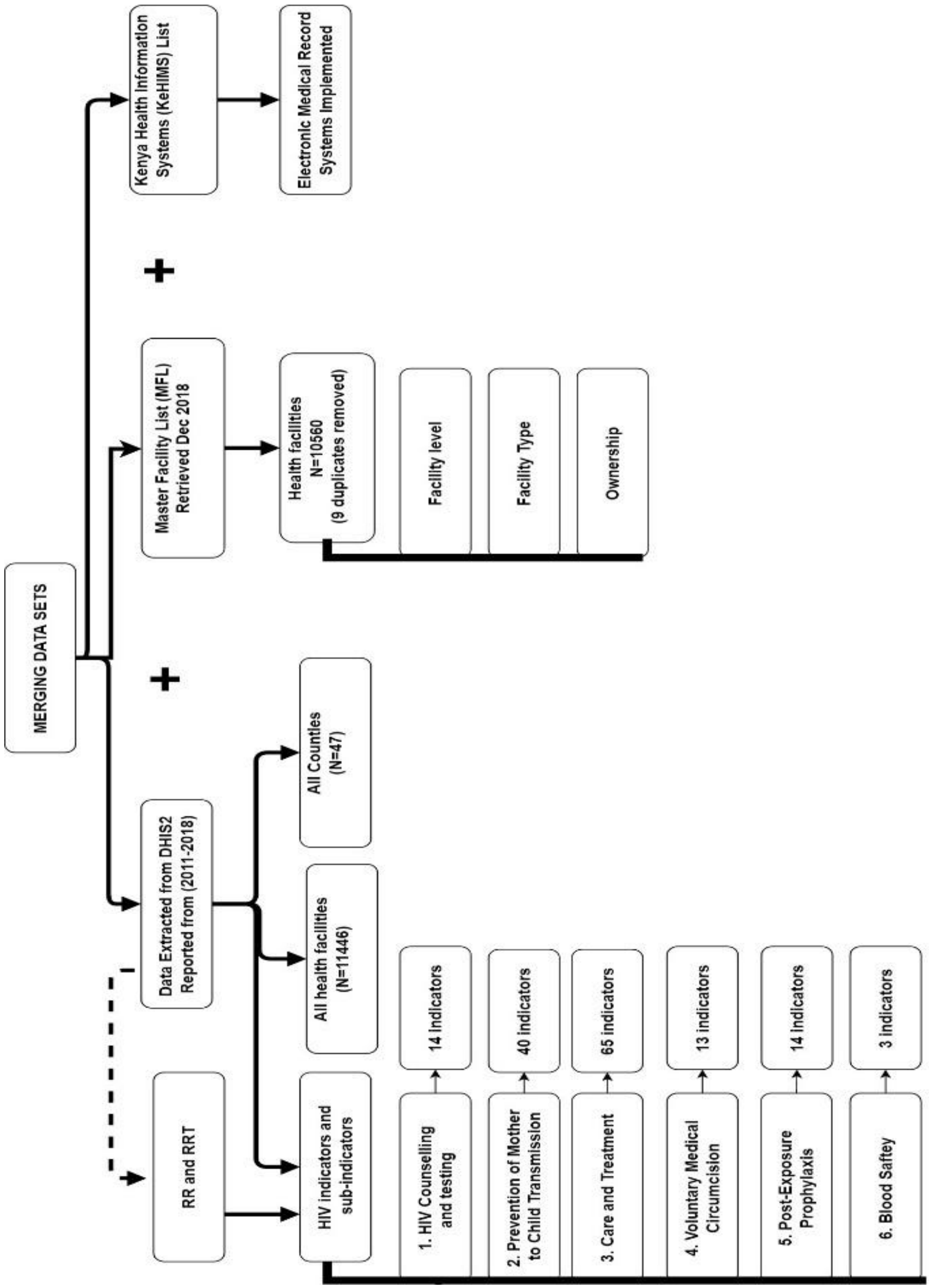

Figure 5. Creation of the data set to be used in secondary analyses (adopted from Gesicho et. al [44])

The generic five-step cleaning sequence comprised of the following steps: (step 1) Outline the analyses or evaluation questions; (step 2) Description of data and study variables; (step 3) Create the data set; (step 4) Apply the framework for data cleaning (*Van den Broeck et al's* framework, involving repeated cycles of a three-phase process data screening, data diagnosis and data treatment (see Figure 6)); and (step 5) Analyze the data.

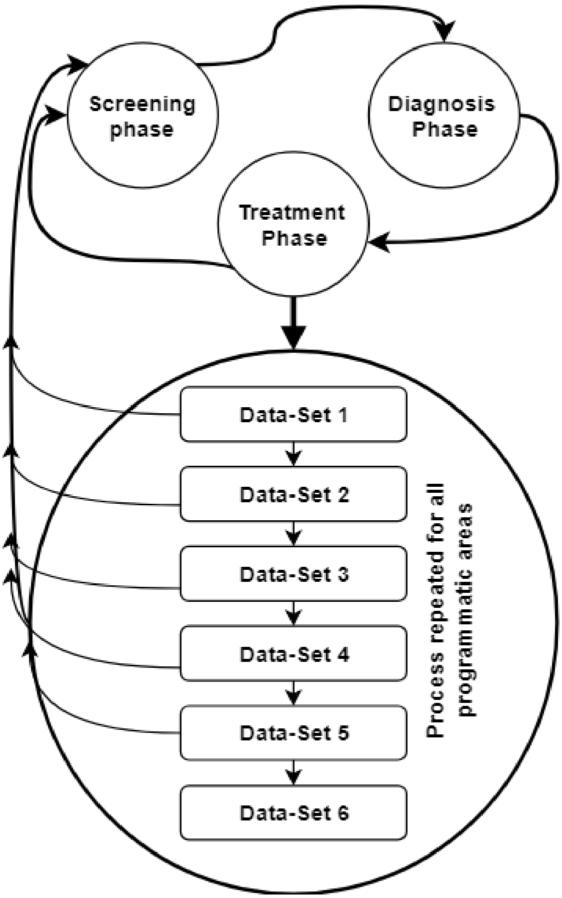

Figure 6. Repeated cycles of data cleaning (adopted from Gesicho et. al [44])

Descriptive statistics such as cross-tabulations and frequency tables were used in order to screen various issues within the data such as duplicates, and outliers. Friedman analysis of variance (ANOVA) was conducted to examine if there is a difference in distribution of facility-records by programmatic area across all years N=8 (2011 to 2018) for the selected situation types (facility records with empty reports, and non-empty reports with missing data for reporting completeness and timelines). The distribution of facility-records was measured in all the six programmatic areas across the eight years and categorized by situation type. Wilcoxon Signed Rank Test were carried out as post hoc tests to compare significances in facility report distribution within the programmatic areas.

II. Sub-cycle 2 – (Resultant publication: Paper II, III and IV)

Secondary data analyses

In this section, a description is provided for the secondary analyses performed on the clean data set obtained from data extraction and cleaning. These analyses aimed at assessing reporting performance of health facilities at meeting completeness and timeliness reporting requirements to DHIS2. The description is categorized according to the specific objectives and expected outcomes.

Specific objective 1: Evaluate performance of health care facilities at meeting HIV-indicator reporting requirement

The resultant final clean dataset from the data cleaning process was used in conducting secondary analyses aimed at evaluation of reporting performance by health facilities. Using the clean dataset, a comprehensive assessment of the reporting status for HIV-indicators was conducted, using Kenya as a case study. Specific objective 1 consisted of a primary outcome as well as secondary outcomes. Methods used in achieving these outcomes are described below.

Primary outcome for specific objective 1: *The primary outcome of interest consisted of identifying the performance in reporting by health facilities over time (2011-2018), with facilities put into various performance clusters and performance evaluated in the various programmatic areas.*

Method

A retrospective observational study was conducted to assess reporting performance of health facilities providing any of the HIV services in all 47 counties in Kenya between 2011 to 2018. Cluster analysis was selected as the approach to use in achieving the outcome. Well put by Kaufman and Rousseeuw, "cluster analysis is the art of finding groups in data" [131]. It is also a kind of unsupervised learning whereby similar groups are determined by separating data whose identities are not known in advance [132]. As such, cluster analysis was used to group health facilities based on their performance in meeting reporting requirements, and to identify patterns in the data [133]. K-means clustering algorithm was preferred due to its suitability in pattern recognition, ease of implementation, simplicity, efficiency and empirical success [133]. The variables selected for the analyses include facility reporting completeness, and timeliness, reporting year, and counties.

K-means algorithm is a nonhierarchical procedure where k represents the number of clusters, which need to be specified by the user prior to any clustering [134]. K-means clustering algorithm was used to identify homogeneous groups of health facilities based on their performance in meeting timeliness and completeness facility reporting requirements for each of the six programmatic areas. Determining the number of clusters is a challenge with no standard consensus on finding the "right" k-value in clustering [135, 136]. In addition, numerous evaluation techniques exist on determining clustering validity [134]. Nevertheless, some of the proposed methods are highly dependent on a data set and investigations on their comparative performance and properties are limited [135]. Detailed investigations are based on simulated artificial data rather than real data sets [135, 136]. According to Everitt et al., a limitation of simulated artificial data sets is the ability to generalize the findings [134].

In addition, hierarchical clustering methods have also been proposed in choosing k but when stopping rules are applied to this technique to produce a range of *one* to *n* cluster solutions, it can result to correct decisions or decision errors as presented in the comparative investigation of 30 procedures for determining the number of clusters conducted by Milligan and Cooper [135]. Examples of decision errors include scenarios where too many clusters are obtained, or fewer clusters than present are obtained. When fewer clusters are obtained, information can be lost because distinct clusters are merged causing a disadvantage when trying to derive insights from a data set. Based on the data set and purpose of this study, the average silhouette coefficient was used, which is an intrinsic method of measuring the quality of a cluster [137].

Secondary outcome for specific objective 1: *The secondary outcome was to determine the relationship between facility ownership and performance in reporting completeness and timeliness.*

Methods

To achieve this outcome, a retrospective observational study was conducted to identify the relationship between facility type and performance on HIV-indicator reporting in Kenya. Facilities with missing facility ownership information were excluded from the study. Mann-Whitney U tests were conducted given that the data was not normally distributed, based on normality tests conducted (Shapiro-Wilk tests and test of Homogeneity of Variance).

III. Sub-cycle 3 – (Resulting publication: Paper V)

Qualitative analyses

In this section, the qualitative inquiry based on results from sub cycle 2 is described. The descriptions are categorized by specific objective and expected outcomes.

Specific objective 2: Identify factors contributing to success and deficiencies in HIV indicator reporting

The results from the secondary analyses conducted in specific objective 1 were used in selecting facilities based on reporting performance in order to identify the facilitators and barriers in HIV-indicator reporting among facilities with varying performance in reporting.

Primary outcome: *Identify factors influencing timeliness and completeness of reporting from facilities to DHIS2.*

Method

Qualitative case study is among the most frequently used qualitative research approaches. Yin argues that case studies have come a long way in being accepted as a formal research method [138]. This is because case study research have been confused with "field work" or participant observation, which is a limited perspective of what case study research actually is [138]. Various types of case studies have also been suggested based on seminal works by key prominent authors within the case study research space (Yin, Merriam and Stake) [139]. Furthermore, based on the comparative study conducted by Yazan, it is clear these methodologists demonstrate certain philosophical orientations that shape their perspective/lens on how case study research ought to be conducted as well as its definition [139]. Yin posits the applicability of different epistemological orientations in case study research [138]. As such, although Yin's case study approach leans towards a realist perspective, he further posits that researchers can use a relativist or constructivist approach, as these too are accommodated well in case study research [138]. On the other hand, Stake [140], and Merriam lean towards constructivism epistemological orientation [139]. The various methodologists with seminal works define case study and case as follows:

> Yin defines case study as an "*empirical method that investigates a contemporary phenomenon (the "case") in-depth and within its real-world context, especially when the boundaries between phenomenon and context may not be clearly evident*" [138].
>
> Merriam defines case study as "*an intensive, holistic description and analysis of a bounded phenomenon such as a program, an institution, a person, a process, or a social unit".*
>
> Stake defines case study as "*the study of the particularity and complexity of a single case, coming to understand its activity within important circumstances*" [140].
>
> Miles and Huberman define a case as "*a phenomenon of some sort occurring in a bounded context*" [141].
>
> Creswell on the other hand defines a case as "*a program, organization, event, activity, process, or one or more individuals*" [142].

These descriptions enabled understanding and defining the “case” in this study. As such, the “case” in this study was defined as a health care facility offering HIV services. The cases in this study were bounded by context, which includes only HIV healthcare facilities that meet the following criteria (1) located in Nairobi (2) either use EMRS or paper in reporting, (3) performance based on reporting completeness and timeliness. In addition, a constructivist approach was applied in this qualitative case study.

A multiple qualitative case study design was employed. The criteria for case selection was based on performance in HIV-indicator facility reporting regarding completeness and timeliness. Questions of interest revolved around reporting procedures, organizational, behavioral, and technical factors. Purposive sampling was used to identify key informants in the study to conduct in-depth interviews. Data was collected using semi-structured interviews (13 participants), archival records, documentation, and informal direct observation.

IV. Sub-cycle 4 (Paper VI: Manuscript)

Data presentation and use

In this section, methods used in presenting visualization of some of the results in specific objective 1 are described.

Specific objective 3: Develop visualizations for reporting completeness and timeliness to aid in decision-making

For this objective, the analyzed HIV-indicator reporting data from specific objective 1 was utilized. The methods used in achieving these outcomes are described below.

Primary outcome: *To develop a performance dashboard with visualizations that represent performance of facility reporting based on timeliness and completeness in reporting.*

Method

(i) Application of design principles

An interactive dashboard of facility reporting performance was designed using Tableau, as it offers diverse dashboard features [143]. Designing of the performance dashboard was guided by dashboard design principles proposed by Few [109], and World Health Organization (WHO) guidelines for graphical presentations of health information, intended to monitor areas such as HIV [110]. A task based approach similar to that employed in our previous study [112], was used in selecting the tasks to be performed on the dashboard. Three tasks were identified with the aim of assessing the performance of facilities at meeting completeness and timeliness reporting requirements allowing for testing of dashboard usability. The designed dashboard is illustrated in Figure 7.

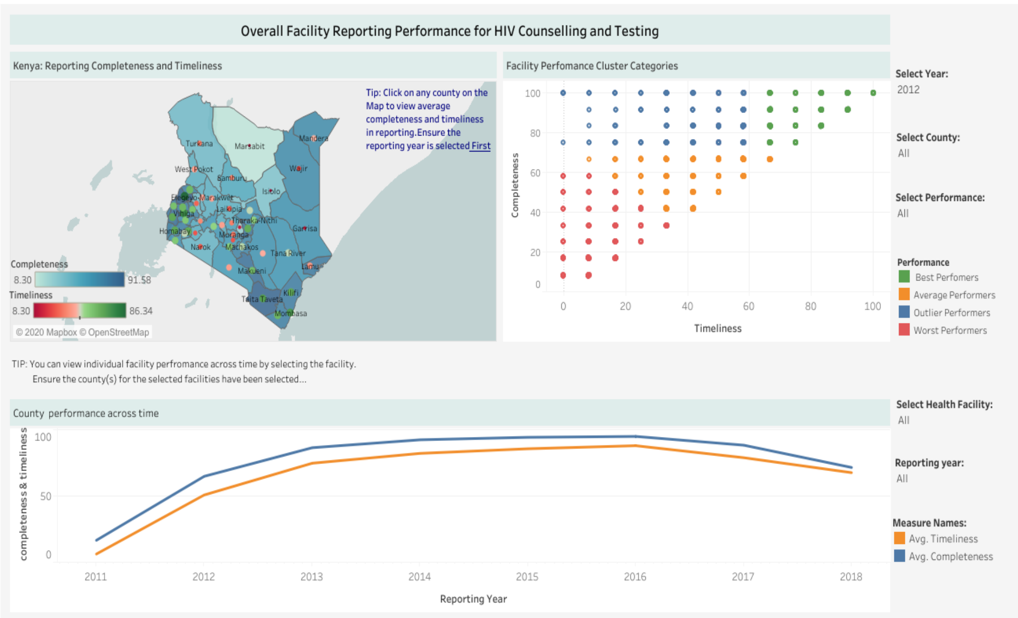

Figure 7. Facility reporting performance dashboard visualization

The visualizations are described as follows:

Map: The color code used for the various counties represents the reporting performance in a particular year whereby various shades of green represents good reporting whereas the shades of red-orange representing poor reporting, with red being very poor.

Scatter chart: This was used to identify performance (based on the four performance category clusters) by facility, year, or county.

Line graph: This was used to identify performance trend in reporting completeness and timeliness over the eight-year period by facility.

(ii) Usability tests

Brooke summarizes usability as being "a general quality of the appropriateness to a purpose of any particular artifact [144]." The usability testing was conducted on the developed visualizations using the System Usability Scale (SUS), which consists of a five-point Likert scale, ten-item questionnaire [144]. SUS was selected because it provides a global view of subjective assessments of usability [144].

5.3.4 Evaluation

This phase entails evaluating the developed artifact within a set of selected criteria. The iterations and feedback resulting from the evaluations are referred to as circumscription [116]. Three main questions proposed by Pries-Heje et al. informed the description of the evaluation that is, what is evaluated (the type of artifact); how it is evaluated (artificial or naturalistic); and when it was evaluated (ex-ante: before artifact development, or ex post: after artifact development) [145].

5.3.5 Conclusion

This phase entails consolidating and writing of the results [116]. As such, Vaishnavi et al. categorize the knowledge gained as either "firm" (can be applied repeatedly) or "loose ends" (need for further research) [116]. Moreover, this phase also entails communicating the research to relevant audience as similarly indicated by Hevner et al. [119], and is depicted by the arrow coming out of the knowledge contribution in Figure 8. (Section 5.1.2) [116].

6. Chapter 6

6.1 Results

This section provides a summary of key findings based on each of the specific objectives investigated in this dissertation. The results are presented based on the DSR process model proposed by Vaishnavi et al. [116], which guided the development of the artifact. A detailed explanation of the results is found in the respective publications (see Appendices).

6.2 Application of Design Science Research (DSR) Process Model

An illustration of how the model was mapped to the dissertation is illustrated in Figure 8.

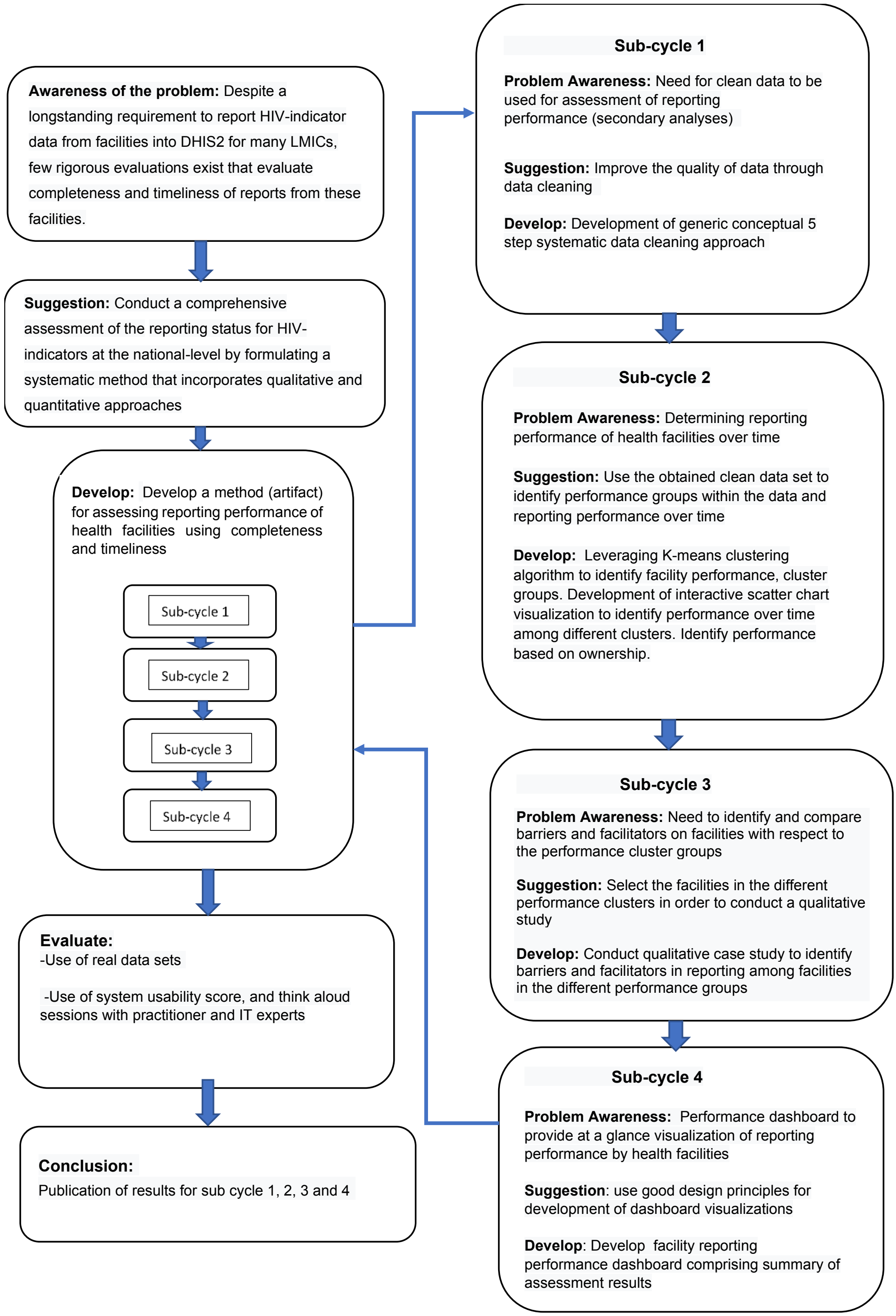

Figure 8. Expanded section of the adopted DSR process model as applied in the dissertation

6.2.1 Awareness of the problem

The ability to report complete, accurate and timely data by HIV care providers and other entities is a key aspect in monitoring trends in HIV prevention, treatment and care, hence contributing to its eradication. In many LMICs, aggregate HIV data reporting is done through the District Health Information Software (DHIS2), which has contributed to improved availability of routinely generated HIV data from health facilities to the national level. Nevertheless, despite a longstanding requirement to report HIV-indicator data to DHIS2 in LMICs, few rigorous evaluations exist to evaluate health facility reporting at meeting completeness and timeliness requirements overtime.

6.2.2 Suggestion

The proposal for this dissertation entailed formulating a method that incorporates the use of both quantitative and qualitative research approaches for assessing facility reporting performance using completeness and timeliness reporting requirements. As such, various approaches used in studies assessing reporting performance by facilities in DHIS2 were identified. An overview of the proposed approach was presented, which aimed at filling gaps from previous approaches.

6.2.3 Development

The resultant systematic method artifact based on the sub-cycles in the development phase is illustrated in Figure 9. The results for the individual sub-cycles are described in the subsequent sections.

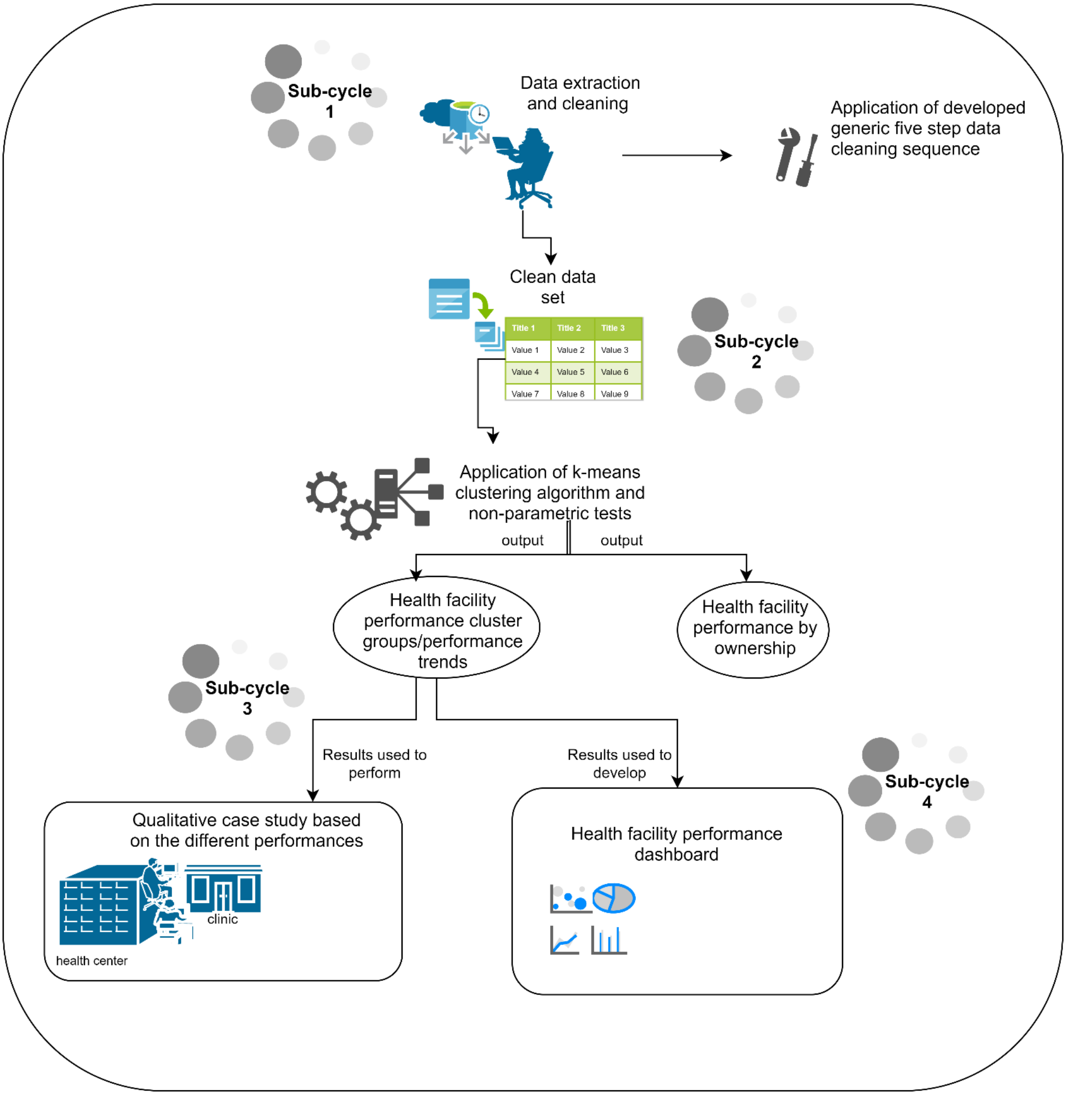

Figure 9. Expanded illustration of the developed systematic method artifact

I. Sub-cycle 1 – (Resultant publication: Paper I)

The findings reveal that data cleaning process involves repeated cycles of identifying and correcting issues within the data. Figure 10 illustrates the number (proportion) of various facility-records that were excluded as they did not meet the ideal criteria (only facilities submitting HIV-indicator data to DHIS2) and contained quality issues. A breakdown of reporting by proportion of facilities and programmatic areas in descending order based on facility records retained after cleaning in data set 4 is as follows; 93.98% were retained for HIV Counselling and Testing (HTC), 80.98% for Prevention of Mother to Child Transmission (PMTCT), 43.79% for Care and Treatment (CRT), 22.10% for Post Exposure Prophylaxis (PEP), 0.66% for Voluntary Medical Male Circumcision (VMMC), and 0.45% for Blood Safety (BS).

50.23% of facility-records had no data and were removed. 0.03% of the remaining had over 100% in reporting rates. Distribution of facility-records with selected quality issues varied significantly by programmatic area ($p<0.001$). Friedman Tests results for empty reports reveal that PEP had the highest mean rank of 6.00 compared to the other programmatic areas CT (3.50), PMTCT (4.88), CrT (2.00), VMMC (3.00), PEP and BS (1.63). On the other hand, Friedman Tests results for distribution of records with non-empty reports with issues reveal that PMTCT and CrT had the highest mean rank of 5.88 and 5.13 respectively compared to the other programmatic areas CT (3.00), VMMC (3.06), PEP (2.88) and BS (1.06).

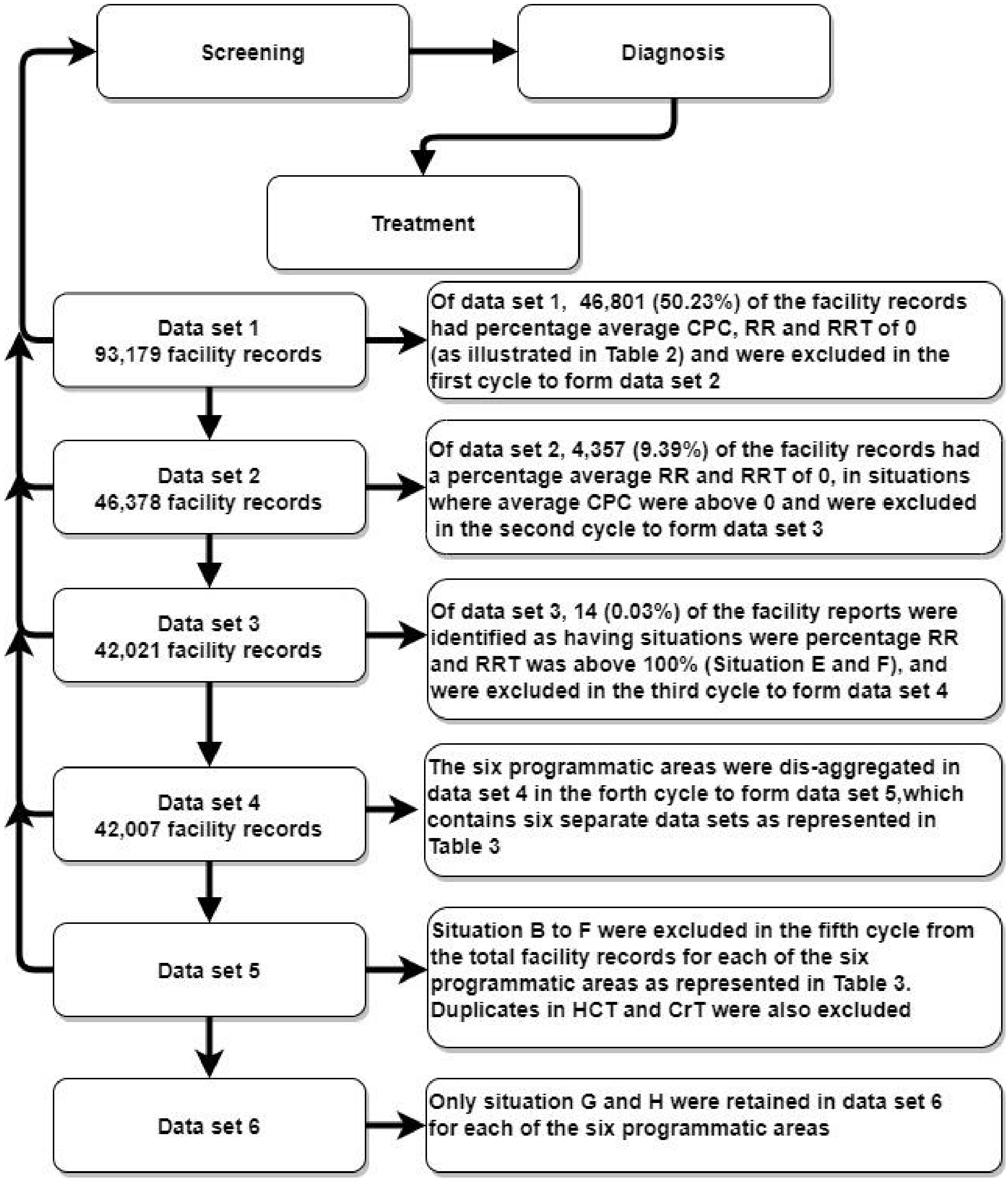

Figure 10. Data cleaning process (adopted from Gesicho et al. [44])

II. Sub-cycle 2 – (Resultant publication: Paper II, III and IV)

The resulting clean dataset was used in achieving specific objective 1.

Specific objective 1: Evaluate performance of health care facilities at meeting HIV-indicator reporting requirement

The findings for this specific objective are highlighted under each of the outcomes of interest.

Primary outcome for specific objective 1: *The primary outcome of interest consisted of identifying the performance in reporting by health facilities over time (2011-2018), with facilities put into various performance clusters and performance evaluated in the various programmatic areas.*

Results: Primary outcome for specific objective 1

Based on percentage average facility reporting completeness and timeliness, four homogeneous groups of facilities were identified namely: best performers, average performers, poor performers, and outlier performers [45]. The four clusters were characterized based on health facility performance as follows:

Best performers: This cluster consisted of health facilities that had the highest percentage in reporting completeness and timeliness in a particular reporting year.

Average performers: This cluster consisted of health facilities that had lower percentage in reporting completeness and timeliness compared to best performers in a particular year.

Poor performers: This cluster consisted of health facilities with lowest percentage in reporting completeness and timeliness in a particular year.

Outlier performers: This cluster consisted of health facilities with high percentage in completeness compared to average performers, but with low percentage in timeliness in that particular year.

Performance was therefore categorized per year by cluster. As such, the average percentage reporting completeness and timeliness for a particular cluster group varied by year.

Apart from blood safety reports, a distinct pattern was observed in five of the remaining reports, with the proportion of best performing facilities increasing and the proportion of poor performing facilities decreasing over time [45]. However, between 2016 and 2018, the proportion of best performers declined in some of the programmatic areas. Over the study period, no distinct pattern or trend in proportion changes was observed among facilities in the average and outlier groups. In this section, results for HIV Testing and Counselling (HTC) programmatic area were used to provide an illustration of the performance trend based on proportion of facilities by year (Figure 11).

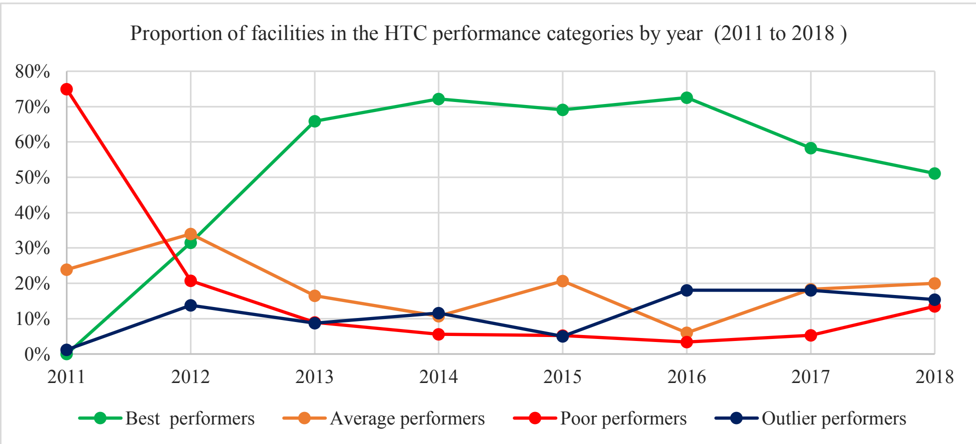

Figure 11. HTC performance trend based on proportion of facilities by year (adopted from Gesicho et al. [45])

Secondary outcome for specific objective 1: *Impact of facility ownership based on completeness and timeliness*

Results: Secondary outcome for specific objective 1

Results revealed that public facilities had statistically significant better performance compared to private facilities, with the exception of year 2017 in reporting of indicators for HIV testing and counselling, and prevention of mother-to-child transmission programmatic areas. In this section, results for HTC were used to illustrate the reporting performance by private and public health facilities from 2011 to 2018 [47].

Table 4. Results for Mann-Whitney U tests on reporting performance for HIV Testing and Counselling indicators by health facilities, highlighted are significant differences. (adopted from Gesicho et at. [47])

Year	**Ownership**	**N**	**P-Value**		**Mean Rank**	
			Completeness	Timeliness	Completeness	Timeliness
2011	Public	487	0.601	0.072	319.46	310.93
	Private	147			311.00	339.26
2012	Public	2404	0.000	0.427	1663.63	1636.26
	Private	852			1529.37	1606.60
2013	Public	3002	0.927	0.254	2063.97	2050.18
	Private	1123			2060.39	2097.26
2014	Public	3189	0.003	0.342	2254.11	2212.81
	Private	1258			2147.67	2252.37
2015	Public	3276	0.000	0.001	2408.86	2379.74
	Private	1404			2181.00	2248.93
2016	Public	3419	0.000	0.004	2479.73	2477.57
	Private	1463			2352.16	2357.21
2017	Public	3494	0.000	0.000	2443.43	2421.70
	Private	1559			2714.30	2762.99
2018	Public	3589	0.304	0.360	2608.49	2607.40
	Private	1600			2564.73	2567.19

III. Sub-cycle 3 – (Resulting publication: Paper V)

The results obtained in specific objective 1 were used in achieving specific objective 2 by selecting facilities performing well and those performing poorly.

Specific objective 2: Identify factors contributing to success and deficiencies in HIV indicator reporting

The findings for this specific objective are highlighted under each of the outcomes of interest.

Primary outcome: *Identify factors influencing timeliness and completeness of reporting from facilities to DHIS2.*

Results

Findings revealed that facilitators and barriers in reporting were based on the following factors: availability of teamwork and skilled personnel, time constraints, availability and access to national aggregate system, complexity of reports, staff rotation, trainings and mentorship, use of standard operating procedures and availability of resources. There was less variation in barriers and facilitators faced by facilities performing well and those performing poor.

Figure 12 provides a summary of the factors in relation to the reporting process. The factors are color coded to represent the category they belong to. In addition, red and green were used to illustrate whether timeliness was affected positively (green) or negatively (red).

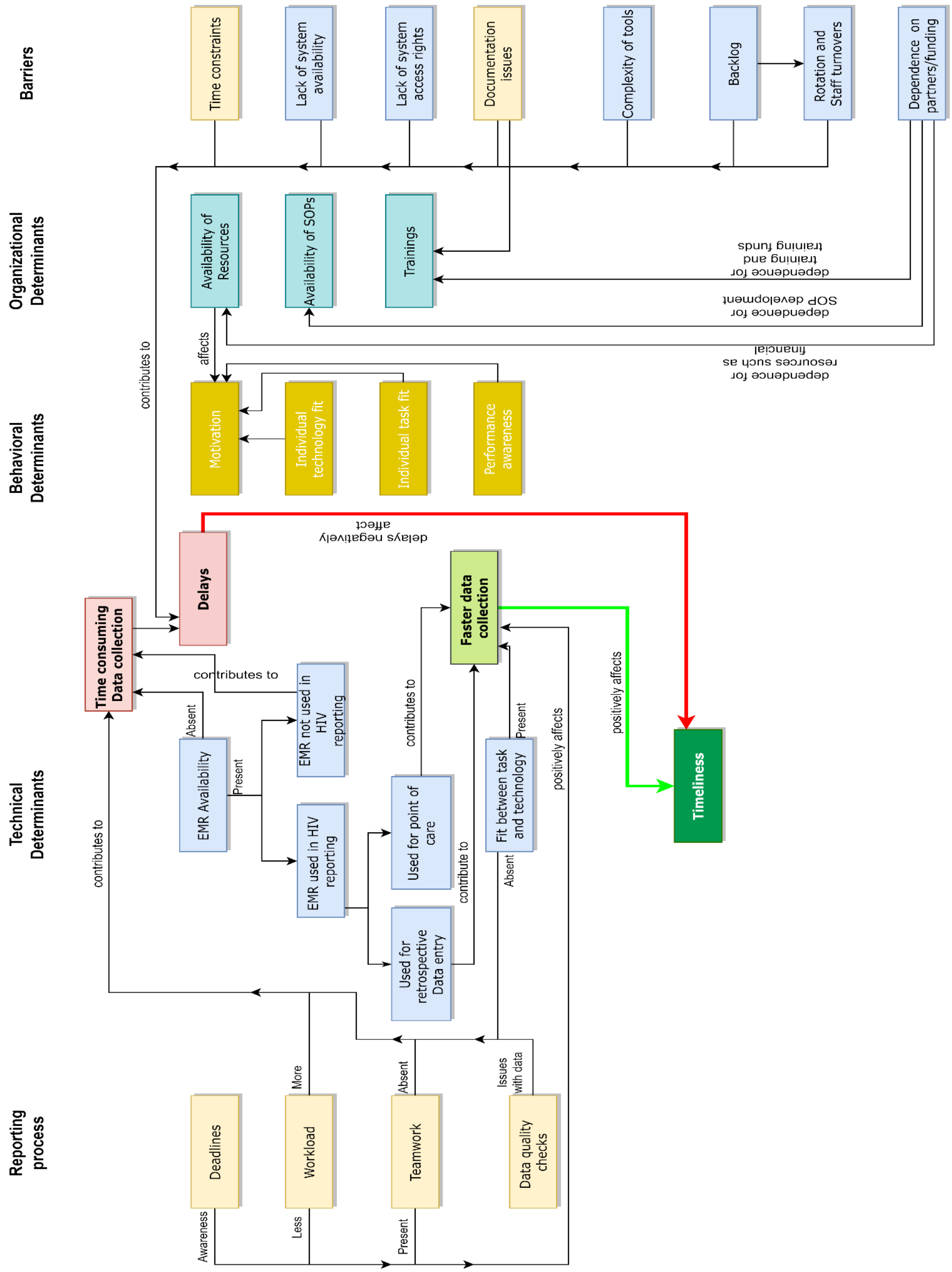

Figure 12. Barriers and facilitators in HIV-indicator data reporting

IV. Sub-cycle 4 (Paper VI: Manuscript)

The results obtained in specific objective 1 were also used in achieving specific objective 3

Specific objective 3: Develop visualizations for reporting completeness and timeliness to aid in decision-making

Primary outcome: *To develop a performance dashboard with visualizations that represent performance of facility reporting based on timeliness and completeness in reporting.*

Results

Users were able to drill-down the data using filters and managed to perform the assigned tasks well. Some of the remarks concerning the dashboard from the users include, “I could understand quickly and gain insight (TE3)”; “The colors are easy to read (TE1)”; “I cannot see from the dashboard what timeliness means, but finding the information was very easy and clear (TE4).”The usability score of the designed dashboard was found to be 87 (Figure 13), which is above average 68 [144].

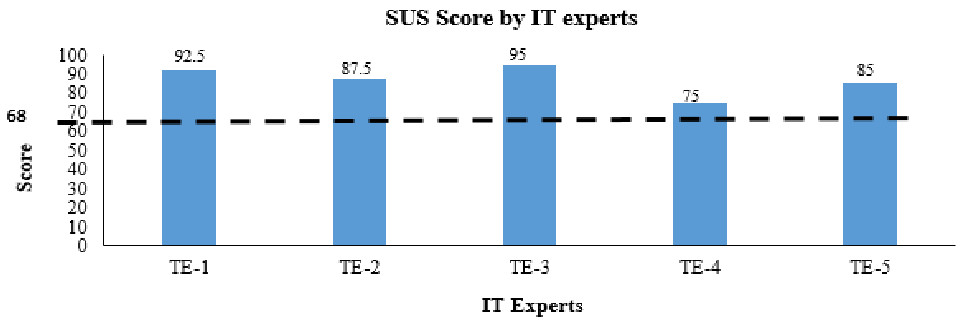

Figure 13. Results of the system usability scale of five IT experts

6.2.4 Evaluation

Describing the '*artifact*' has been a point of debate among researchers in DSR [146]. The artifact is thus required to provide relevant solution to the identified problem by fulfilling a specific utility, which is assessed by the artifact's ability to meet pre-defined criteria. There are various criteria proposed in literature that IT artifacts can be valuated upon such as completeness, performance, generality and consistency [119, 117] .

Prat et al. proposed a holistic approach of evaluating an artifact, which applies a general systems theory in presenting generic evaluation methods as well as the various evaluation criteria that these methods assess [147]. Moreover, based on their view of IS artifacts as systems, Prat et al. grouped the various interrelated criteria into five dimensions of systems that is; goal, environment, structure, activity and evolution [147].

Hevner et al. also propose various methods of evaluating artifacts, which are categorized as observational, analytical, experimental, testing, and descriptive. According to Hevner et al. [119], evaluation method used should match with the artifact as well as the evaluation criteria. Pries-Heje et al. proposed a strategic framework framework can be used to describe evaluation strategies used in a DSR [145]. The strategic DSR evaluation framework is in from of a quadrant and seeks to answer three main questions (Figure 14) [145]. These three main questions that need to be considered include, what is evaluated (the type of artifact); how it is evaluated (artificial or naturalistic); and when it was evaluated (ex-ante-before artifact development, or ex post- after artifact development) [145].

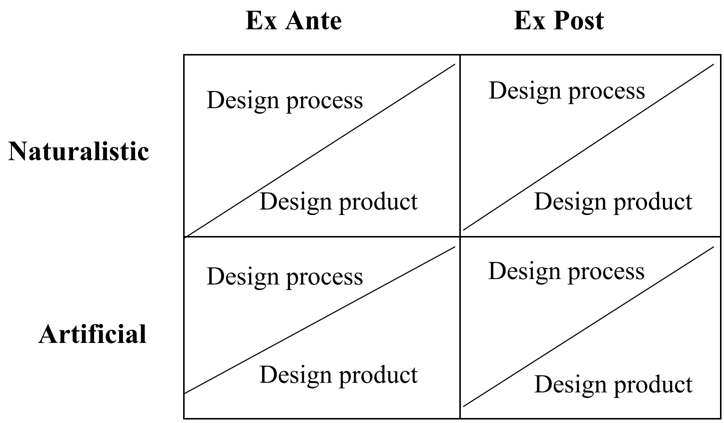

Figure 14. Strategic DSR evaluation framework adopted from Priers-Heje et al. [145]

In this dissertation, description of evaluation of the artifact was structured using three questions comprised within the strategic DSR evaluation framework [145], as follows:

What was actually evaluated? - A method of assessing reporting performance of health facilities using facility reporting completeness and timeliness as reporting requirements. This included the reporting performance dashboard, which is a part of the artifact (method) in this dissertation.

How it is evaluated? - The evaluation was a natural evaluation as real-world data sets were used as well as use of system usability scale with IT experts, and interview with practitioner and researcher. The evaluation criteria assessed is goal (ability to meet intended objective, which involved assessing reporting performance of facilities), and efficacy (ability to satisfy users in terms of usability).

When was it evaluated? -The artifact was evaluated ex post.

6.2.5 Conclusion

The results were published in peer review publications and presented at international conferences. In addition, the dissertation will be available in open-source university site (BORA).

7. Chapter 7

7.1 Discussion

In this section, a discussion of the main findings for each of the sub-cycles is provided. The systematic method of assessing facility reporting performance resulted from the combination of four sub-cycles within the development step.

(i) Sub-cycle 1: Data extraction and cleaning process

This dissertation entailed extraction of a voluminous amount of data gathered for a period of eight years for all facilities in all 47 counties in Kenya within the DHIS2 national-aggregate system. As such, it was important for data cleaning to be performed on the extracted data set. Nonetheless, data cleaning is not a trivial endeavor. As described in **Paper I** [44]**,** a number of iterations were performed to identify and deal with "dirty" data. Moreover, prior to embarking on the task of data cleaning, it was necessary to provide a background on data quality in order to establish a basis for data cleaning with respect to data quality. As such, data cleaning in this dissertation was viewed as a process of improving the quality of data for use in secondary analyses. As it turned out, no off-the-shelf solution for cleaning this dataset existed that could fit the intended use. Thus, the challenge was to develop a tailored solution for the aggregate data.

Providing background on data quality also aimed at avoiding confusion of data cleaning activity with data quality assessment. Whereas data quality assessments and data cleaning go hand in hand, there are contrasting aspects that distinguish them. Data quality assessments in our view entail selection of one or more predefined data quality dimensions such as accuracy and examining whether the quality dimensions were met within the data. Data cleaning on the other hand is a process of detecting quality issues within the data and deciding on the appropriate action to take for the identified issues, thereby improving the quality of data. Moreover, data cleaning consists of a broad category of issues that need to be defined and identified, given the broad categories of "dirty data" that exists [148] as well as the objective aspects and subjective aspects of data quality [149].

The conceptual framework proposed by Wang and Strong helped frame the focus for data cleaning [86]. Their framework categorized data quality dimensions into four categories: intrinsic data quality, contextual data quality, representational and accessibility data quality [86]. As such, given that the focus was to retain data that was fit for use, a contextual data quality perspective was deemed fit for identifying quality issues in the data cleaning process. Hence, the remaining three categories were beyond the scope of the study. Moreover, it was assumed that the data quality mechanisms with the DHIS2 had dealt with syntactic data quality (database rules), which is an aspect of intrinsic data quality.

With this in mind, it was presumed that there is no "one size fits all" solution to data cleaning, especially given the varying contexts and tasks for which the data is to be used for. For instance, this dissertation encountered facility-records with incomplete reporting, which can be viewed from various perspectives. This can include a situation where a report for a service provided has been successfully submitted but is incomplete [150–152]; or missing reports (expected reports have not been submitted consistently for all 12 months), hence making it difficult to identify whether services were provided or not, in months were reports were missing [151]. Whereas decisions made during the data cleaning in this body of work entailed retaining such scenarios, other studies opted to adjust for incomplete reporting based on the aim of each study [151].

Furthermore, this dissertation presented transparent and systematic documentation of procedures used in data cleaning, which are valuable in increasing validity in research, which is advocated in literature [98, 153, 154]. Therefore, the generic five-step approach developed in this study provides a systematic sequence that can be adopted for cleaning data extracted from DHIS2.
A limitation experienced in cleaning the data was that it was labor intensive as well as time-consuming. In addition, there are also limitations with human augmented procedures, through human-generated errors especially when dealing with extremely

large data sets [155]. Nonetheless, the advantage of using a semi-automatic approach employed in this dissertation was that it provided a deeper understanding of the data and insights into approaches that will be needed to subsequently develop automated data cleaning procedures in future. This will be of potential relevance to the many countries already using DHIS2.

(ii) Sub cycle 2: Evaluating performance of health care facilities at meeting HIV-indicator reporting requirement

The clean data obtained from the cleaning process performed in sub cycle 1 enabled conducting evaluations on reporting performance of facilities as described in **Paper II** [45]. Use of k-means clustering algorithm enabled identification of four cluster groups as well as patterns, which might have gone unidentified. These cluster groups include best performers, average performers, poor performers and outlier performers. The categorization as well as interpretation of the cluster groups was based on the average percentage completeness and timeliness reporting performance by health facilities. As such, interpretation of the cluster groups was also based on yearly performance as percentage completeness and timeliness varied by year in each cluster group. For example, best performer group in the care and treatment programmatic area had an average completeness and timeliness of 97.67% and 90.81% in 2013 respectively, and 98.67%, and 82.01% in 2014.

Moreover, apart from the blood safety programmatic area, a distinct pattern observed in five of the other programmatic areas was that as the proportion of best performing facilities increased, the proportion of poor performing facilities decreased. In addition, the proportion of facilities in the best performing cluster was higher over time, compared to the proportion of facilities in the other performance clusters. These observations denote improvements in reporting over time in Kenya.

A common attribute among average, poor and outlier performance is the discrepancy between completeness and timeliness as described in **Paper III** Gesicho et. al [46]. Therefore, results from this study reveal the need for investigating issues that bring about delays more so in the outlier performance group, were facility reporting timeliness was low.

Moreover, a contribution in this dissertation was the application of machine learning algorithms in evaluation of reporting performance by facilities, which is a novel approach that can be replicated in settings using DHIS2. In addition, this approach to our knowledge has not been utilized in evaluating performance in the context of DHIS2. The scope of the study can be relevant for many countries dealing with HIV reporting in aggregate-level HMIS. However, the limitation in this study is that data have been collected and analyzed for one country only. Nonetheless, the indicators used (completeness and timeliness) could also be relevant in other contexts.

Moreover, whilst assessing the relationship between facility ownership and reporting performance as in Paper IV [47], it was observed that public institutions performed better in the timeliness and completeness of the MoH-mandated HIV indicator reports in Kenya. This seemed paradoxical, given the general perception that private institutions are often 'better' at meeting quality and reporting requirements [47]. The better performance by the public sector could likely be due to significant investments in care and reporting in the public sector by programs such as Global Fund and PEPFAR. Conducting this analysis demonstrated the potential of how aggregate reporting data can be used to inform decision-making. Nonetheless, only yearly dimensions were assessed, thus further subsequent analyses could be done by county, facility level, and facility type.

(iii) Sub cycle 3: Identifying factors contributing to success and deficiencies in HIV indicator reporting

Facilities in the four performance cluster groups in sub-cycle 3 were purposively selected for a qualitative case study, which identified barriers and facilitators in HIV reporting that were linked to the RHIS process and determinants of RHIS (technical, behavioral, and organizational determinants) as in **Paper V** [48]. The qualitative study revealed that whereas facilities may demonstrate differences in reporting performance, they are likely to face similar barriers and facilitators regardless of the contextual differences [48]. For instance, all the facilities that performed poorly in timeliness (outlier performers), had access to EMRS and were in fairly good locations (not in slums where accessibility can be a challenge).

As revealed in one study [156], EMRS do not directly contribute to improved performance in reporting. This is due to lack of EMRS interoperability with DHIS2. Nonetheless, this qualitative inquiry revealed that utilization of EMRS in reporting contributed to easing the data collection burden, as data was retrieved from EMRS rather than multiple registers, which is time consuming. Hence, one would expect that these facilities would at least perform well in reporting due to the investments allotted to them. However, it seemed that there were other underlying issues that affected their performance. For instance, one of the facilities had been going through a pilot phase of point of care EMRS implementation. It was assumed that this might have also contributed to its performance in reporting.

Moreover, it emerged that some of the facilities had no access to DHIS2 during the study period (2017 and 2018), which might have affected their reporting timeliness. As revealed in some of the key findings, poor performance in timeliness by facilities submitting reports by hand was in part due to insufficient human resources, lack of availability of the DHIS2, and slow internet connection at the sub-county level resulting to late entries of hand submitted reports. Another key finding is time constraints, which were aggravated by issues such as, reporting period falling on a weekend [48]. In addition, funded facilities were at an advantage due to continuous on job training as opposed to non-funded facilities who depended on trainings by MoH as well as supportive supervisions [48].

A limitation of this assessment was that there had been staff turnovers and rotations among health records officers in some of the facilities hence not able to give conclusive descriptions of happenings of previous years. In addition, though cases were in one county, it is expected that the findings revealed in this assessment are transferrable in other counties and in LMICs in similar contexts.

(iv) Sub cycle 4: Developing visualizations of reporting requirements to aid in decision-making

Recommendations for designing dashboards using results from sub cycle two came about due to the need of utilizing good design principles in developing a performance dashboard instance of facility reporting of HIV-indicators over time. The intention was to improve the user experience by giving possibility to interact with the data by selecting a period of analysis and selecting county or facility of interest. As such, good design principles were utilized in designing an interactive dashboard, which incorporated three visualizations, and fit on a single screen. Moreover, the performance dashboard designed aimed at promoting use of data by enabling users explore and derive insights on facility reporting performance at a glance.

Moreover, the application of data visualizations also provided novel ways in which reporting performance data within DHIS2 can be represented. Furthermore, due to the volume, velocity and veracity of health data consolidated from various sources, representing various facets of data in a way that makes sense is a challenge. Hence, another strength of this dissertation was the ability to summarize voluminous amount of retrospective data gathered within eight years into simplistic representations. These promote at a glance view of reporting performance by facilities over a period of time, which could be of benefit in gaining of insights as revealed in additional interviews conducted involving a researcher in the domain area as well as an M&E practitioner.

7.2 Discussion of Research Rigor

7.2.1 Validity and Reliability

Reliability and validity are salient in research as they demonstrate rigor and trustworthiness in research [157]. Basing on common definition used in texts, Dyson and Brown describe validity as "how close what is being measured in practice is, to what we intend to measure in theory", and reliability as the 'consistency' of a measure [157]. Furthermore, Dyson and Brown describe validity as a back and forth between two realms; theory and observation, where a researcher links their ideas of how the world operates to the 'real world' where the ideas are operationalized. According to Roberts and Priest [158], trustworthiness revolves around the original research question, data collection methods, data analysis and conclusions drawn.

This dissertation consisted of both qualitative and quantitative approaches, which lie on rationalistic and naturalistic paradigms respectively. Consequently, various researchers in qualitative studies have argued on the need for using alternative terms for reliability and validity as these were considered to be more appropriate for quantitative studies [159]. As such, paradigm specific criteria are needed for addressing 'rigor' (common term used in rationalistic paradigm) or trustworthiness (common term used in naturalistic paradigm) [157]. Therefore, in this dissertation, these concepts are addressed in respect to the approach used.

Quantitative approach

To determine the reporting performance of facilities at meeting HIV-indicator reporting requirements, data was extracted from the DHIS2 in order to conduct secondary analyses.

Approaches used to enhance validity

External validity: HIV-indicator data was extracted from all facilities in all the 47 counties in Kenya gathered within a period of eight years. Thus, all available data in DHIS2 at the time of study was extracted, and statistical analyses conducted. Given that DHIS2 is implemented in over 70 countries, approaches used in this dissertation can be applicable in similar settings especially given that the measures applied (reporting completeness and reporting timeliness) are used in countries reporting data into DHIS2.

Statistical conclusion validity: Threats to this validity can occur due to violation of assumptions of statistical tests. As such, analyses in this dissertation were conducted based on the appropriate statistical assumptions. For instance, non-parametric tests such as Mann Whitney U tests were conducted given that the data was not normally distributed. This was concluded based on normality tests conducted using Shapiro-Wilk tests and test of Homogeneity of Variance**.**

Construct validity: Dyson and Brian denote validity of measures as being applicable across space and time so as to facilitate comparisons over long or different periods of time [157]. Facility reporting completeness and timeliness are considered as good measures as they enabled identifying reporting performance over a long period of time

(2011-2018). Determining relationships between measures was beyond the scope of this study.

Approaches used to enhance reliability

Standardized approaches were used in determining reporting performance of facilities. The formular used in automatically calculating facility reporting completeness and facility reporting timeliness are outlined by the WHO among other partners [160]. In addition, the k-means algorithm as applied in this dissertation is a popular approach used for purposes of grouping data [133]. Hence, various performance groups were identified within the data. Further still, transparent, and systematic data cleaning approaches were used in order to ensure only data that is ideal for secondary analysis was retained. This promoted understanding of how the data sets used to conduct analyses were generated and obtained.

Qualitative approach

To identifying barriers and facilitators in HIV-indicator reporting for different health facility performances, a qualitative case study approach was used. Guba and Lincoln proposed replacement of concepts of reliability and validity with trustworthiness, in qualitative enquiries [161].

Approaches used to enhance trustworthiness

Various approaches were used in order to ensure trustworthiness of the study. Guba and Linoln proposed four concepts of ensuring trustworthiness which include credibility, transferability, dependability and confirmability [161]. These are discussed as follows:

Credibility: To ensure credibility of the study, prolonged engagement with participants, and triangulation of data sources was conducted long enough. Peer debriefing was also carried out in sessions after conducting two to three interviews.

Dependability and confirmability: To ensure dependability and confirmability, an audit trail using Nvivo 11, was used to manage and store data. Reflexivity was achieved on the basis that the researcher was a stranger in the setting hence had no established prior familiarity with participants, which may be a threat to validity (bias).

Transferability: Transferability was attained through provision of thick descriptions hence enabling applicability of findings in other similar settings.

7.3 Discussion of secondary analysis of existing data

Data that are collected or extracted from sources such as databases or registries for various purposes are regarded as 'secondary analysis of existing data'. Cheng and Philips posit that using the term 'secondary analysis of existing data' other than 'secondary data analysis' prevents confusion of differentiating whether the data in use is primary data or secondary data [162]. It is important to note that this dissertation utilized existing data for analyses. Nonetheless a broad term 'secondary analyses' was used in denoting secondary analyses of existing data.

Cheng and Philips propose two approaches for analysing existing data: research question-driven approach and data-driven approach, which are used iteratively or jointly in research [162]. In this dissertation, a question-driven approach was used given that the variables selected in the data set were only those that aimed to address the research question.

Cheng and Philips also mention the need for having a clear understanding of the strengths and weaknesses on the dataset [162]. As much as their work seemed to focus on survey data, some of the approaches could also be suitable for data extracted from large databases. As such, the data cleaning process conducted for the dataset used in this dissertation promoted a deeper understanding of the data set hence uncovering weaknesses as well as strengths within the dataset.

The issue of 'imposition' also needs to be considered by researchers when analysing data. As such, the results should not be viewed as being imposed as to how the world ought to be seen [157]. In contrast, the results ought to be viewed as bringing different perspectives or awareness to a certain phenomenon, which may lead to various causes of action. Hence, the results in this dissertation provide a different perspective in which HIV-indicator data can be used in assessing facility reporting performance.

8. Chapter 8

8.1 Conclusions

This dissertation contributes towards strengthening HIV monitoring and evaluation efforts, with the goal of enabling epidemic control. As such, proper use of available data is necessary if at all the various goals such as the ambitious UNAIDS 90 90 90 goals [4], as well as the millennium development goals are to be achieved [68]. As such, this dissertation achieved the overall aim of developing and applying a systematic method for conducting a rigorous evaluation of HIV-indicator data reporting by health facilities to DHIS2 using completeness and timeliness as reporting-performance indicators.

Data cleaning proved to be a worthwhile undertaking prior to conducting any analyses as data anomalies as well as data considered as “dirty” were removed in order to retain the intended clean data set. In addition, the data cleaning exercise brought to perspective the need for more robust and systematic automated processes to be integrated within DHIS2 in order to complement the existing data quality mechanisms. Moreover, the developed generic five step sequence approach enabled comprehensive, transparent, and systematic reporting on the approaches as well as results obtained for the cleaning exercise. As such, issues that may trigger further evaluation as well as improvements within the system are disclosed. Aside from that, validity of study that is largely advocated through transparent reporting [98, 153] was promoted as researchers, monitoring and evaluation teams as well as other stakeholders are able to replicate the approaches used. In addition, the results obtained using the data can be easily understood as the approach used in obtaining the dataset was clearly and elaborately described.

Furthermore, by leveraging on machine learning algorithms (such as the k-means clustering) on the resulting data set, a novel approach is provided that can be used by monitoring and evaluation teams to evaluate routine reporting. Another contribution also entailed the identification and interpretation of four resulting clusters, which revealed the various performance groups that facilities fall into. The outlier group more so revealed issues with poor performance in timeliness that require further investigations and resources in order to facilitate improvements. In addition, using a retrospective observational study to evaluate HIV routine reporting enabled comprehensive assessment as few rigorous evaluations exist despite a long-standing requirement to report HIV-indicators to DHIS2 in LMICs. Given that the resultant data set contained facility attributes such as ownership type (private/public), this enabled ascertaining relationship between facility ownership and performance in reporting completeness and timeliness for a period of eight years (2011-2018). These results are of benefit to ministries of health as well as stakeholders as it reveals the place of ownership in reporting. In addition, this study demonstrated other ways in which reporting data from DHIS2 can be used.

Further still, qualitative inquiries supplemented quantitative analysis as seen in this dissertation. For instance, some of the factors that contribute to poor performance in reporting timeliness among facilities in the various performance groups were identified. Some of the issues identified that influence reporting include factors such as organizational factors (staff turnovers, training, volume of work, resources allocated to reporting), technical factors (DHIS2 was not linked to EMRS, which hinders the potential of EMRS in reporting). These findings would be of benefit to ministries of health as well other stakeholders for purposes of finding solutions on barriers identified. Hence, this dissertation reveals that qualitative analysis need to be considered to supplement quantitative studies.

The recommendations for data visualization provided in this dissertation further revealed the importance of utilizing design principles in developing visualizations. Given that data visualization is salient especially when representing large amounts of data in order to derive insight and decision-making, this work provided recommendation on how retrospective data can be presented on a dashboard while adhering to design principles. This work could thus benefit those in charge of developing visualizations for routine reporting as it provides them with other ways in which data can be represented and interacted with.

8.2 Future Work

Based on this dissertation, future work is recommended in the following areas:

Leveraging data mining on DHIS2 data

Machine learning algorithms can be used for data mining. The machine learning algorithms such as those used in this work can be extended to the data stored within DHIS2. This work only used HIV data within DHIS2. Nonetheless, DHIS2 contains data for diseases such as Malaria as well as administrative data. As such, there exists a plethora of knowledge awaiting to be unveiled. Future work should therefore leverage on various machine learning algorithms as well as other approaches in data mining in order to obtain insights within the data. In addition, deriving of insights from retrospective data for routine reporting as well as the indicator data is particularly useful in identifying patterns, improvements as well as predictions where possible depending on the data.

Integration of statistical tools for data quality

Based on the data cleaning exercise conducted in this dissertation, it emerged that integration of statistical tools to compliment mechanisms used for checking data quality would be of benefit within DHIS2. For example, use of descriptive statistics to identify the number of empty reports submitted by a facility in a given year for a specific programmatic area. This would enable decision-makers to identify frequency of provision of certain services offered by facilities.

As such, if a facility is not offering a certain service, the system should flag based on calculations of the frequency of empty reports within a certain period. Moreover, the system should restrict retrieval of reporting data from such facilities as they are not useful for decision-making, and only add on to the work for data cleaning. Furthermore, the existing visualizations within DHIS2 can be used for purposes of evaluating data quality. For example, cross tabulations were used to identify reporting rates that are over 100%. Use of visualizations within DHIS2 can be used to identify such issues and clean the data by either correcting or removing so that it can be suitable for analysis by researchers as well as interested stakeholders.

Exploring data visualization

Simple visualizations such as charts and bar graphs (see Figure 15) are the most commonly used in representing data. Nonetheless, more sophisticated visualizations have been advocated for purposes utilizing space when representing various facets of data [163]. This also avoids overcrowding especially when using a dashboard.

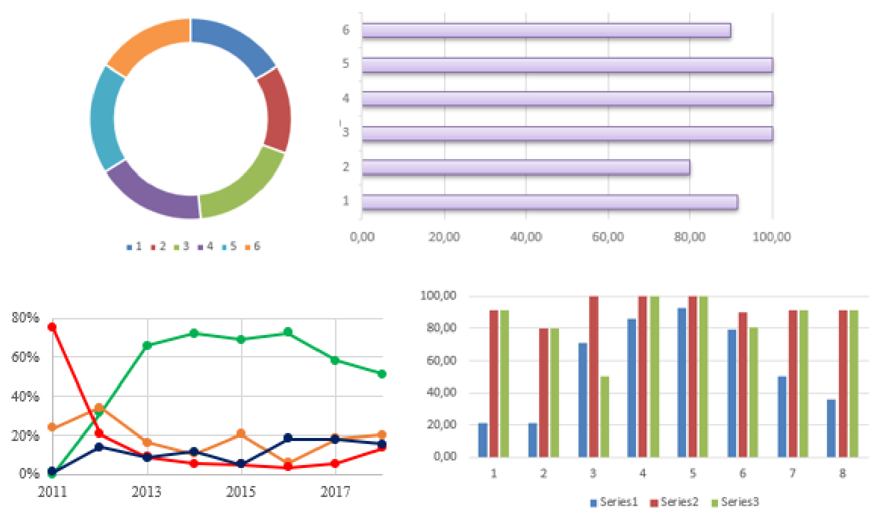

Figure 15 . Examples of simple visualizations

Future work can use data within DHIS2 to develop sophisticated visualizations, which utilize space. Whereas DHIS2 contains a number of simple visualizations, future developments can explore utilization of sophisticated visualizations [163].

Figure 16. is an example of how sophisticated visualizations can be used for data extracted within DHIS2 [112]. Such visualizations can be used in representing the number of people tested for HIV within a certain age group. This example only provides an at a glance representation in order to derive insights.

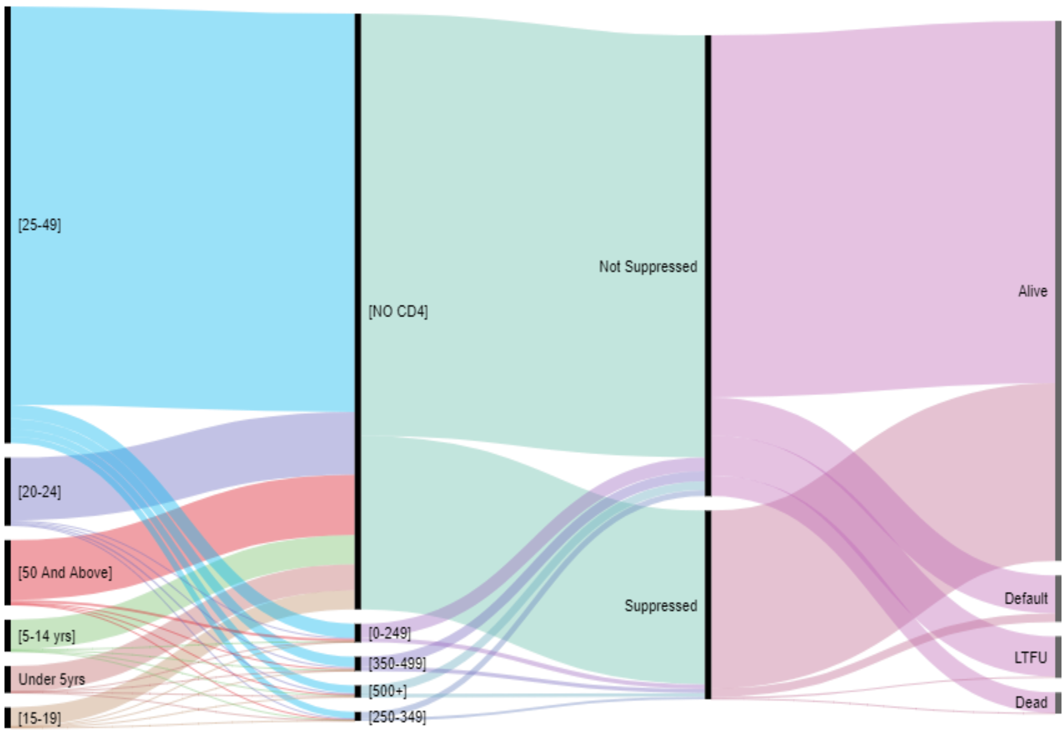

Figure 16. Sophisticated visualization of patient outcomes adopted from Gesicho and Babic [112]

Figure 16 reveals ages 24-49 as having majority of people living with HIV (PLHV). In addition, it can be viewed at a glance that CD4 count was not recorded for majority of the patients and suppression rate was low.

Moreover, existing visualization within DHIS2 can also be utilized more in representing various types of data for purposes of gaining insights. Figure 17 illustrates the proportion of expected HIV reports in 2016.

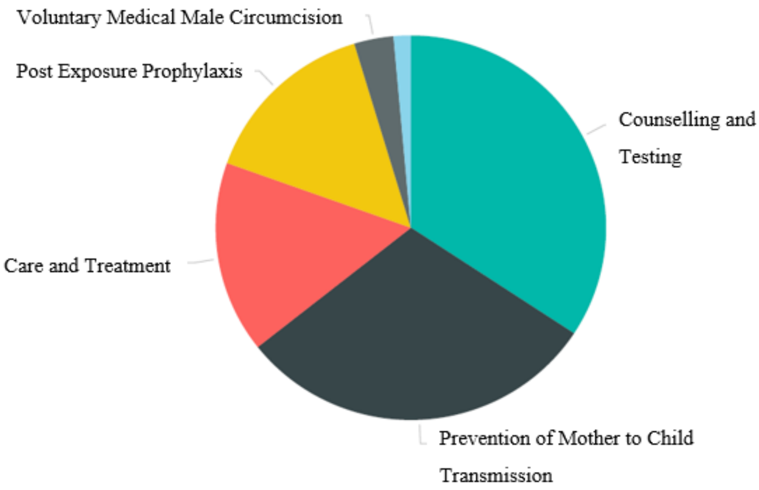

Figure 17. Proportion of expected HIV-indicator reports in 2016

On the other hand, Figure 18 provides an example of using maps to identify the actual reports submitted by counties in 2016 using the size of the circles. The bigger the circle, the more the reports submitted and vice versa. For illustrative purposes, the visualizations for both Figure 17 and 18 were developed using tools such as power BI. Nonetheless visualizations within DHIS2 can as well be used.

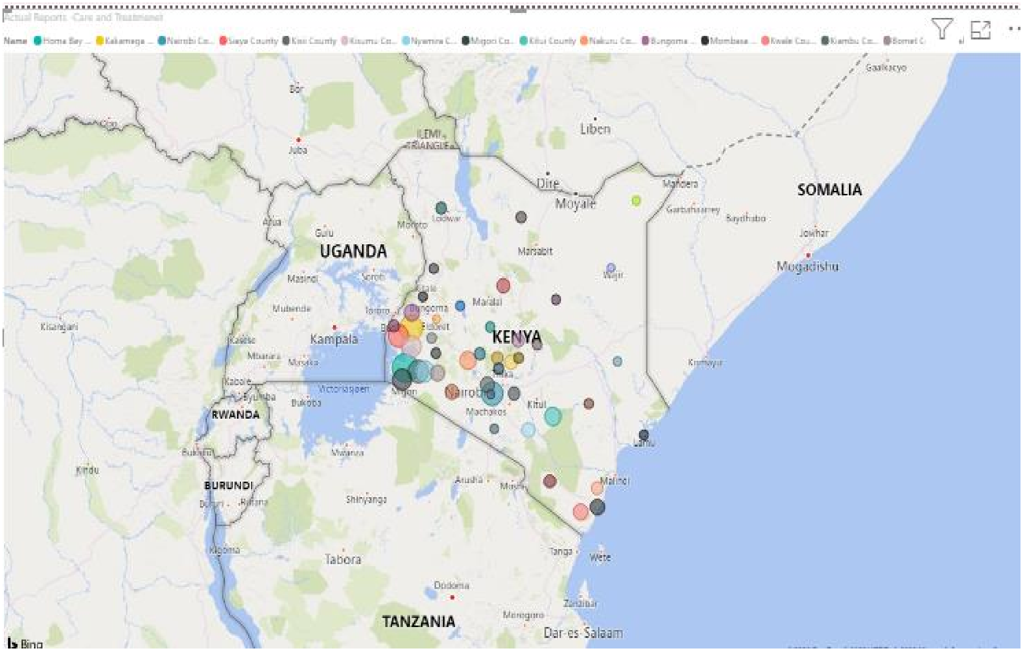

Figure 18. Care and treatment actual reports submitted by counties (2016)

There is potential to utilize the work in this dissertation by researchers and decision-makers, since the dissertation demonstrates the use of data and approaches that could be further applied in different environments that collect and utilize reporting data for other diseases.

References

1. UNAIDS. Global hiv statistics. 2019. p. 6. Available from: https://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf

2. UNAIDS. AIDSinfo | UNAIDS. 2018. Available from: http://aidsinfo.unaids.org/

3. UNAIDS. Communities at the centre: Global AIDS Update . 2019. Available from: https://www.unaids.org/sites/default/files/media_asset/2019-global-AIDS-update_en.pdf

4. UNAIDS. 90-90-90 An ambitious treatment target to help end the AIDS epidemic . 2014 . Available from: http://files.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2014/90-90-90_en.pdf

5. WHO/UNAIDS/UNICEF. Towards Universal access . UNAIDS Annual Report. 2010. Available from: www.who.int/hiv

6. Schneider MT, Birger M, Haakenstad A, Singh L, Hamavid H, Chapin A, et al. Tracking development assistance for HIV/AIDS: the international response to a global epidemic. AIDS . 2016;30:1475–9.

7. Dieleman JL, Haakenstad A, Micah A, Moses M, Abbafati C, Acharya P, et al. Spending on health and HIV/AIDS: domestic health spending and development assistance in 188 countries, 1995–2015. Lancet. 2018 May 5;391(10132):1799–829.

8. Lippeveld T, Sauerborn R, Bodart C. Design and implementation of health information systems Edited by. World Heal Organ. 2000;280.

9. World Health Organization. Everybody's Business: Strengthening Health Systems to Improve Health Outcomes: WHO's Framework for Action. GenevaWorld Heal Organ . 2007 Available from: http://www.who.int/healthsystems/strategy/everybodys_business.pdf

10. AbouZahr C, Boerma T. Health information systems: the foundations of public health. Bull World Health Organ . 2005 ;83(8):578–83.

11. Mpofu M, Semo B-W, Grignon J, Lebelonyane R, Ludick S, Matshediso E, et al. Strengthening monitoring and evaluation (M&E) and building sustainable health information systems in resource limited countries: lessons learned from an M&E task-shifting initiative in Botswana. BMC Public Health.2014;14.

12. Ekouevi DK, Karcher S, Coffie PA. Strengthening health systems through HIV monitoring and evaluation in Sub-Saharan Africa . Current Opinion in HIV and

AIDS Jul, 2011 p. 245–50.

13. Wagenaar BH, Hirschhorn LR, Henley C, Gremu A, Sindano N, Chilengi R, et al. Data-driven quality improvement in low-and middle-income country health systems: Lessons from seven years of implementation experience across Mozambique, Rwanda, and Zambia. BMC Health Serv Res. 2017;17.

14. Fraser HSF, Biondich P, Moodley D, Choi S, Mamlin BW, Szolovits P. Implementing electronic medical record systems in developing countries. Inform Prim Care. 2005;13(2):83–95.

15. Manya A, Braa J, Øverland L, Titlestad O, Mumo J, Nzioka C. National Roll Out of District Health Information Software (DHIS 2) in Kenya, 2011–Central Server and Cloud Based Infrastructure. IST-Africa 2012 . 2012;1–9.

16. Nicol E, Bradshaw D, Uwimana-Nicol J, Dudley L. Perceptions about data-informed decisions: an assessment of information-use in high HIV-prevalence settings in South Africa. BMC Health Serv Res . 2017;17(S2):765.

17. 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 May 1;90(5):379–84.

18. Nutley T, McNabb S, Salentine S. Impact of a decision-support tool on decision making at the district level in Kenya. Heal Res Policy Syst . 2013 Dec 8 ;11(1):34.

19. Courtenay-Quirk C, Spindler H, Leidich A, Bachanas P. Building Capacity for Data-Driven Decision Making in African HIV Testing Programs: Field Perspectives on Data Use Workshops. AIDS Educ Prev . 2016 ;28(6):472–84.

20. Suleman Qazi M, Ali M. Health Management Information System utilization in Pakistan: Challenges, pitfalls and the way forward. Biosci Trends . 2011;5(6):245–54.

21. Nutley T, Reynolds HW. Improving the use of health data for health system strengthening. Glob Health Action . 2013;6:20001.

22. Measure Evaluation. Barriers to Use of Health Data in Low- and Middle-Income Countries — A Review of the Literature — MEASURE Evaluation . 2018. Available from: https://www.measureevaluation.org/resources/publications/wp-18-211

23. Nash D, Elul B, Rabkin M, Tun M, Saito S, Becker M, et al. Strategies for more effective monitoring and evaluation systems in HIV programmatic scale-up in resource-limited settings: Implications for health systems strengthening. J Acquir Immune Defic Syndr JAIDS. 2009;52:58–62.

24. Nyaga Karani F, Okibo Bichanga W, Guandaru Kamau C. Effective use of

monitoring and evaluation systems in managing HIV/AIDS related projects: A case study of local NGOS in Kenya Effective Use of Monitoring and Evaluation Systems in Managing HIV/AIDS Related Projects: A Case Study of Local NGOS in. Kenya Sci J Bus Manag . 2014;2(2):67–76.

25. Unaids. Indicator Standards: Operational Guidelines for Selecting Indicators for the HIV Response. 2010;28. Available from: http://www.unaids.org/sites/default/files/sub_landing/files/4_3_MERG_Indicator_Standards.pdf

26. Porter LE, Bouey PD, Curtis S, Hochgesang M, Idele P, Jefferson B, et al. Beyond Indicators. JAIDS J Acquir Immune Defic Syndr . 2012;60:S120–6.

27. Karuri J, Waiganjo P, Orwa D, Manya A. DHIS2: The Tool to Improve Health Data Demand and Use in Kenya. J Health Inform Dev Ctries .2014;8(1):38–60.

28. Kariuki JM, Manders E-J, Richards J, Oluoch T, Kimanga D, Wanyee S, et al. Automating indicator data reporting from health facility EMR to a national aggregate data system in Kenya: An Interoperability field-test using OpenMRS and DHIS2. Online J Public Health Inform . 2016;8(2):e188.

29. Ledikwe JH, Grignon J, Lebelonyane R, Ludick S, Matshediso E, Sento BW, et al. Improving the quality of health information: a qualitative assessment of data management and reporting systems in Botswana. Heal Res Policy Syst . 2014;12(1):7.

30. Mbondo M, Scherer J, Aluoch GO, Sundsmo A, Mwaura N. Organizational HIV monitoring and evaluation capacity rapid needs assessment: The case of Kenya. Pan Afr Med J . 2013 ;14:1–7.

31. Glèlè Ahanhanzo Y, Ouedraogo LT, Kpozèhouen A, Coppieters Y, Makoutodé M, Wilmet-Dramaix M. Factors associated with data quality in the routine health information system of Benin. Arch Public Health . 2014;72(1):25.

32. Cahill K, Fleming D, Conway M, Gupta S. Global Health Partnerships: Assessing Country Consequences. Tuberculosis . 2005;41(3):92–120.

33. Bernardi R, De Chiara F. ICTs and monitoring of MDGs: a case study of Kenya HIV/AIDS monitoring and evaluation in a donor multi-agency context. Inf Technol Dev . 2011;17(1):24–41.

34. Souteyrand Y, Banda MA, Kaufman J, Perriëns JH, Yu D. Investment in HIV/AIDS programs: Does it help strengthen health systems in developing countries? Global Health . 2008 ;4.

35. Manya A, Nielsen P. Reporting practices and data quality in health information systems in developing countries: an exploratory case study in Kenya. J Health Inform Dev Ctries . 2016 ;10(1):114–26.

36. Kariuki J, Manders E, Richards J, Oluoch T, Mutiso J. Automating Indicator Data Reporting from an EMR to Aggregate Data System Using OpenMRS and DHIS2. J Heal Informatics Africa. 2013;1(1).

37. Bhattacharya AA, Umar N, Audu A, Allen E, Schellenberg JRM, Marchant T. Quality of routine facility data for monitoring priority maternal and newborn indicators in DHIS2: A case study from Gombe State, Nigeria. Bazzano AN, editor. PLoS One . 2019;14(1):e0211265.

38. Adokiya MN, Awoonor-Williams JK, Beiersmann C, Müller O. Evaluation of the reporting completeness and timeliness of the integrated disease surveillance and response system in northern Ghana. Ghana Med J. 2016;50(1):3–8.

39. 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(1).

40. Kiberu VM, Matovu JK, Makumbi F, Kyozira C, Mukooyo E, Wanyenze RK. Strengthening district-based health reporting through the district health management information software system: The Ugandan experience. BMC Med Inform Decis Mak . 2014;14(1):40.

41. 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 Heal Int J Environ Res Public Heal Int J Environ Res Public Heal.2014;11:5170–207.

42. WHO. Data Quality Review (DQR) Toolkit . WHO. World Health Organization; 2019 . Available from:who.int/healthinfo/tools_data_analysis/en/

43. Githinji S, Oyando R, Malinga J, Ejersa W, Soti D, Rono J, et al. Completeness of malaria indicator data reporting via the District Health Information Software 2 in Kenya, 2011-2015. BMC Malar J. 2017;16(1):1–11.

44. Gesicho MB, Were MC, Babic A. Data cleaning process for HIV-indicator data extracted from DHIS2 national reporting system: a case study of Kenya. BMC Med Inform Decis Mak . 2020;20(1):293.

45. Gesicho MB, Were MC, Babic A. Evaluating performance of health care facilities at meeting HIV-indicator reporting requirements in Kenya: an application of K-means clustering algorithm. BMC Med Inform Decis Mak . 2021;21(1):6.

46. Gesicho MB, Babic A, Were MC. K-means Clustering in Monitoring Facility Reporting of HIV Indicator data:Case of Kenya. In: Studies in Health Technology and Informatics . IOS Press; 2020. p. 143–6.

47. Gesicho M, Babic A, Were M. A Retrospective Observational Study of Health Facility Ownership Type and Performance on HIV Indicator Data Reporting in Kenya. In: IFMBE Proceedings . Springer Science and Business Media Deutschland GmbH; 2021. p. 38–44.

48. Gesicho MB, Babic A. Identifying barriers and facilitators in HIV-indicator reporting for different health facility performances: A qualitative case study. Gopichandran V, editor. PLoS One . 2021;16(2):e0247525.

49. Etamesor S, Ottih C, Salihu IN, Okpani AI. Data for decision making: using a dashboard to strengthen routine immunisation in Nigeria. BMJ Glob Heal. 2018;3(5):e000807.

50. Centers for Disease Control and Prevention. Principles of Epidemiology in Public Health Practice, Third Edition: An Introduction to Applied Epidemiology and Biostatistics. Atlanta, GA: CDC; 2006.

51. Cohen MS, Hellmann N, Levy JA, Decock K, Lange J. The spread, treatment, and prevention of HIV-1: Evolution of a global pandemic . Vol. 118, Journal of Clinical Investigation. American Society for Clinical Investigation; . p. 1244–54.

52. World Health Organization (WHO). Global HIV, Hepatitis and STIs Programme . HIV Update. 2020 . Available from: https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/data-use/hiv-data-and-statistics

53. Health Information Systems Program (HISP). DHIS2 in Action . 2015. Available from: https://www.dhis2.org/in-action

54. National AIDs Control Council. Kenya HIV Estimates Report 2018 . 2018. Available from: https://nacc.or.ke/wp-content/uploads/2018/12/HIV-estimates-report-Kenya-20182.pdf

55. National AIDS Control Council NACC. Kenya AIDS Response Progress Report 2018 . 2018. Available from: www.nacc.or.ke

56. National AIDS Control Council NACC. Kenya AIDS Strategic Framework 2014/2015-2018/2019 . 2014 . Available from: http://www.undp.org/content/dam/kenya/docs/Democratic Governance/KENYA AIDS STRATEGIC FRAMEWORK.pdf

57. UNAIDS. "Three Ones" Key Principles. "Coordination of National Responses to HIV/AIDS" Guiding principles for national authorities and their partners. In: International Conference on AIDS and STIs . 2004. p. 1–4. Available from: http://data.unaids.org/una-docs/three-ones_keyprinciples_en.pdf

58. Haakenstad A, Moses MW, Tao T, Tsakalos G, Zlavog B, Kates J, et al.

Potential for additional government spending on HIV/AIDS in 137 low-income and middle-income countries: an economic modelling study. Lancet HIV. 2019 Jun 1;6(6):e382–95.

59. Nuwagaba-Biribonwoha H, Werq-Semo B, Abdallah A, Cunningham A, Gamaliel JG, Mtunga S, et al. Introducing a multi-site program for early diagnosis of HIV infection among HIV-exposed infants in Tanzania. BMC Pediatr . 2010;10:44.

60. WHO. Guideline on When to Start Antiretroviral Therapy And On Pre-Exposure Prophylaxis For Hiv Guidelines. 2015.

61. Cork MA, Wilson KF, Perkins S, Collison ML, Deshpande A, Eaton JW, et al. Mapping male circumcision for HIV prevention efforts in sub-Saharan Africa. BMC Med . 2020 ;18(1):189.

62. Hoxha K, Hung YW, Irwin BR, Grépin 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 Information Management Journal. SAGE Publications Inc.; 2020

63. 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(3):217–28.

64. World Health Organization (WHO). Framework and Standards for Health Information Systems . 2nd ed. Geneva: WHO Press; 2008. Available from: http://www.who.int/

65. Hahn D, Wanjala P, Marx M. Where is information quality lost at clinical level? A mixed-method study on information systems and data quality in three urban Kenyan ANC clinics. Glob Health Action . 2013;6:21424.

66. Mutale W, Chintu N, Amoroso C, Awoonor-Williams K, Phillips J, Baynes C, et al. Improving health information systems for decision making across five sub-Saharan African countries: Implementation strategies from the African Health Initiative. BMC Health Serv Res . 2013;13(Suppl 2):S9.

67. Measure Evaluation. Data Demand and Information Use in the Health Sector. 2007;126.

68. Peersman G, Rugg D, Erkkola T, Kiwango E, Yang J. Are the Investments in National HIV Monitoring and Evaluation Systems Paying Off? JAIDS J Acquir Immune Defic Syndr . 2009 ;52:S87–96. 1

69. 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(3):176–82.

70. Chaulagai CN, Moyo CM, Koot J, Moyo HB, Sambakunsi TC, Khunga FM, et al. Design and implementation of a health management information system in Malawi: issues, innovations and results. Health Policy Plan. 2005. 1;20(6):375–84.

71. Tiberio J, Said C, Ndayongeje J, Kivenule D, Komba T, Welty S. Promoting a culture of data use for evidence-based decisions in HIV programming : regional HIV / AIDS triangulation in Tanzania. East African J Appl Heal Monit Eval. 2010 ;1–7.

72. MEASURE Evaluation. Data Quality Audit Tool. 2008.

73. 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(4):1–14.

74. 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 Inform . 2016 ;95:60–70.

75. Ledikwe JH, Reason LL, Burnett SM, Busang L, Bodika S, Lebelonyane R, et al. Establishing a health information workforce: Innovation for low- and middle-income countries. Hum Resour Health . 2013 ;11(1):35.

76. Health Information Systems Program (HISP). DHIS 2 . Available from: https://www.dhis2.org/

77. Dhis2 Documentation Team. Control data quality . DHIS2 User Manual. 2020. Available from: https://docs.dhis2.org/2.31/en/user/html/dhis2_user_manual_en_full.html#control_data_quality

78. Haugen JÅ, Hjemås G, Poppe O. Manual for the DHIS2 quality tool.Understanding the basics of improving data quality. 2017.

79. W. H. Inmon. Building the Data Warehouse. 3rd ed. Wiley Computer Publishing; 2002.

80. Biehl RE. Data warehousing for biomedical informatics. CRC Press; 2016.

81. DHIS2 Team. DHIS2 as Data Warehouse - DHIS2 Documentation . 2008. Available from: https://docs.dhis2.org/2.34/en/dhis2_implementation_guide/dhis2-as-data-warehouse.html

82. Kalogriopoulos NA, Baran J, Nimunkar AJ, Webster JG. Electronic medical record systems for developing countries: Review. Proc 31st Annu Int Conf IEEE Eng Med Biol Soc Eng Futur Biomed EMBC 2009. 2009;53706:1730–3.

83. Muinga N, Magare S, Monda J, Kamau O, Houston S, Fraser H, et al. Implementing an open source electronic health record system in kenyan health care facilities: Case study. J Med Internet Res. 2018;20(4):1–23.

84. Kang' S, Puttkammer N, Wanyee S, Kimanga D, Madrano J, Muthee V, et al. A national standards-based assessment on functionality of electronic medical records systems used in Kenyan public-Sector health facilities. Int J Med Inform . 2017;97:68–75.

85. Wang RY. A Product Perspective on Total Data Quality Management. Commun ACM. 1998;41(2):58–65.

86. Wang RY, Strong DM. Beyond Accuracy: What Data Quality Means to Data Consumers. Vol. 12, Source: Journal of Management Information Systems. 1996.

87. Shanks G, Corbitt B. Understanding Data Quality : Social and Cultural Aspects. In: 10th Australasian Conference on Information Systems,. 1999. p. 785–97.

88. Liaw ST, Taggart J, Dennis S, Yeo A. Data quality and fitness for purpose of routinely collected data--a general practice case study from an electronic practice-based research network (ePBRN). AMIA Annu Symp Proc. 2011;2011:785–94.

89. Who. Improving Data Quality: A Guide for Developing Countries. 2003;1–74. Available from: http://iris.wpro.who.int/bitstream/handle/10665.1/5421/9290610506_eng.pdf

90. Price R, Shanks G. A semiotic information quality framework: Development and comparative analysis. J Inf Technol. 2005;20(2):88–102.

91. Kahn BK, Strong DM, Wang RY. A model for delivering quality information as product and service. In: Proceedings of the 1997 Conference on Information Quality . 1997. p. 80–94.

92. Measure Evaluation. User Manual Routine Data Quality Assessment RDQA User Manual . 2015. Available from: https://www.measureevaluation.org/resources/tools/data-quality/rdqa-guidelines-2015

93. World Health Organization. The immunization data quaity self-assessment (DQS) tool . World Health Organization. 2005. Available from: www.who.int/vaccines-documents/

94. Chapman AD. Principles and Methods of Data Cleaning Primary Species Data. 1st ed. Report for the Global Biodiversity Information Facility. GBIF; 2005. 72 p.

95. Zhang S, Zhang C, Yang Q. Data preparation for data mining. Appl Artif Intell . 2003 May ;17(5–6):375–81.

96. Fayyad U, Piatetsky-Shapiro G, Smyth P. The KDD Process for Extracting Useful Knowledge from Volumes of Data. Commun ACM. 1996;39(11):27–34.

97. Rahm E, Do HH. Data Cleaning: Problems and Current Approaches Erhard. IEEE Trans Cloud Comput . 2014 ;2(1):1–1.

98. Van Den Broeck J, Cunningham SA, Eeckels R, Herbst K. Data cleaning: Detecting, diagnosing, and editing data abnormalities . Vol. 2, PLoS Medicine. Public Library of Science; 2005 . p. 0966–70.

99. Chu X, Ilyas IF, Krishnan S, Wang J. Data cleaning: Overview and emerging challenges. In: Proceedings of the ACM SIGMOD International Conference on Management of Data . New York, New York, USA: ACM Press; 2016 . p. 2201–6.

100. Ilyas IF, Chu X. Trends in cleaning relational data: Consistency and deduplication. Found Trends Databases . 2015;5(4):281–393.

101. Hellerstein JM. Quantitative Data Cleaning for Large Databases. United Nations Econ Comm Eur . 2008 ;42.

102. Müller H, Freytag J-C. Problems, Methods, and Challenges in Comprehensive Data Cleansing Challenges. 2003 . p. 1–23. Available from: http://www.dbis.informatik.hu-berlin.de/fileadmin/research/papers/techreports/2003-hub_ib_164-mueller.pdf

103. Oliveira P, Rodrigues F, Galhardas H. A Taxonomy of Data Quality Problems. 2nd Int Work Data Inf Qual. 2005;(April 2014):219–33.

104. Dziadkowiec O, Callahan T, Ozkaynak M, Reeder B, Welton J. Using a Data Quality Framework to Clean Data Extracted from the Electronic Health Record: A Case Study. 2016 Jun 24 ;4(1).

105. Mamlin BW, Biondich PG, Wolfe BA, Fraser H, Jazayeri D, Allen C, et al. Cooking up an open source EMR for developing countries: OpenMRS - a recipe for successful collaboration. AMIA . Annu Symp proceedings AMIA Symp . 2006 ;2006:529–33.

106. Measure Evaluation. Data Visualization That Works Facilitating HIV Program Targeting : Case Examples and Considerations. Meas Eval . 2016. Available from: https://www.measureevaluation.org/resources/publications/wp-16-162

107. Chrysantina A, Sæbø JI. Assessing user-designed dashboards: a case for developing data visualization competency. In: IFIP Advances in Information and Communication Technology. Springer New York LLC; 2019. p. 448–59.

108. Ghazisaeidi M, Safdari R, Torabi M, Mirzaee M, Farzi J, Goodini A. Development of performance dashboards in healthcare sector: Key practical issues. Acta Inform Medica. 2015;23(5):317–21.

109. Few S. Information Dashboard Design: The Effective Visual Communication of Data. O’Reilly Media, Inc; 2006.

110. WHO. Analysis and Use of Health Facility Data: General Principles . 2018. Available from: https://www.who.int/healthinfo/FacilityAnalysis_GeneralPrinciples.pdf?ua=1

111. Parsons P, Sedig K. Adjustable properties of visual representations: Improving the quality of human-information interaction. J Assoc Inf Sci Technol . 2014;65(3):455–82.

112. Gesicho M, Babic A. Task-based approach recommendations to enhance data visualization in the Kenya national health data warehouse. In: IFMBE Proceedings. Springer Verlag; 2019. p. 467–70.

113. Maimon O, Rokach L. Introduction to Knowledge Discovery and Data Mining. In: Data Mining and Knowledge Discovery Handbook . Springer US; 2009. p. 1–15.

114. Mackenzie N, Knipe S. Research dilemmas: Paradigms, methods and methodology. Vol. 16, Issues in Educational Research. 2006.

115. Hevner A, Chatterjee S. Design Research in Information Systems. Theory and Practice. Springer Science and Business Media LLC; 2010.

116. Vaishnavi V, Kuechler B., Petter S,.Design Science Research in Information Systems. Des Sci Res Inf Syst . 2004; Available from: http://www.desrist.org/design-research-in-information-systems/

117. March ST, Smith G. Design and natural science research on information technology. Decis Support Syst. 1995;15:251–66.

118. Purao S. Design Research in the Techonology of Information Systems: Truth or Dare. School of Information Sciences and Technology, The Pennsylvania State University, University Park, State College, PA. 2002.

119. Hevner AR, March ST, Park J, Ram S. Design science in information systems research. MIS Q Manag Inf Syst. 2004;28(1):75–105.

120. Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. J Manag Inf Syst. 2007;24(3):45–77.

121. Hevner AR. A Three Cycle View of Design Science Research. Vol. 19, Scandinavian Journal of Information Systems. 2007.

122. Hevner A, March ST, Park J, Ram S. Design Science In Informaton Systems Research. AI Soc. 1996 ;10(2):199–217.

123. Takeda H, Veerkamp P, Tomiyama T, Yoshikawa H. Modeling design processes. AI Mag. 1990;11(4):37–48.

124. Rahi S. Research Design and Methods: A Systematic Review of Research Paradigms, Sampling Issues and Instruments Development. Int J Econ Manag Sci. 2017;06(02).

125. Crossan F. Research philosophy:Towards an understanding. Nurse Res. 2003;11(1):46.

126. Wahyuni D. The Research Design Maze : Understanding Paradigms , Cases , Methods and Methodologies. JMAR. 2012;10(1):69–80.

127. o'leary Z. The Essential Guide to doing research . 3rd ed. London: Sage Publications; 2017.

128. Creswel JW. The Selection of a Research Approach. Res Des Qual Quant Mix methods approaches . 2008;3–22.

129. Fan W, Geerts F. Foundations of Data Quality Management. Vol. 4, Synthesis Lectures on Data Management. Morgan & Claypool Publishers LLC; 2012. 1–217 p.

130. Maïga A, Jiwani SS, Mutua MK, Porth TA, Taylor CM, Asiki G, et al. Generating statistics from health facility data: The state of routine health information systems in Eastern and Southern Africa. Vol. 4, BMJ Global Health. BMJ Publishing Group; 2019. p. e001849.

131. Kaufman L, Rousseuw PJ. Finding Groups in Data: An Introduction to Cluster Analysis. Vol. 47, Biometrics. New Jersey: John Wiley & Sons, Inc.; 2005. 788 p.

132. Wilks DS. Cluster Analysis. Int Geophys . 2011;100:603–16. Available from: https://www.sciencedirect.com/science/article/pii/B9780123850225000154

133. Jain AK. Data clustering: 50 years beyond K-means. Pattern Recognit Lett . 2010;31(8):651–66.

134. Everitt BS, Landau S, Leese M, Stahl D. Cluster Analysis. Cluster Analysis: Fifth Edition. wiley; 2011. 1–330 p.

135. Milligan GW, Cooper MC. An examination of procedures for determining the number of clusters in a data set. Psychometrika. 1985;50(2):159–79.

136. Pham DT, Dimov SS, Nguyen CD. Selection of K in K-means clustering. Proc Inst Mech Eng Part C J Mech Eng Sci. 2005(1):103–19.

137. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20(C):53–65.

138. Yin RK. Case Study Research and Applications. Sage Publications; 2018.

139. Yazan B. The Qualitative Report Three Approaches to Case Study Methods in Education: Yin, Merriam, and Stake . Vol. 20, Teaching and Learning. Available from: http://www.nova.edu/ssss/QR/QR20/2/yazan1.pdf

140. Stake RE. The art of case study research . Thousand Oaks : Sage Publications,; 1995. 175 p.

141. Miles MB, Huberman AM, Saldaña J. Qualitative data analysis : a methods sourcebook . Thousand Oaks, CA : Sage; 2013. 381 p.

142. Creswell JW, Preceded by: Creswell JW. Qualitative inquiry & research design : choosing among five approaches . Third edition. 2013. 448 p.

143. Murray D, Chabot C. Tableau Your Data!: Fast and Easy Visual Analysis with Tableau Software . Canada: John Wiley & Sons, Inc.; 2013. 528 p.

144. Brooke J. SUS-A quick and dirty usability scale. Usability Eval Ind. 1996;189(194):4–7.

145. Pries-Heje J, Baskerville R, Venable J. Strategies for design science research evaluation. 16th Eur Conf Inf Syst ECIS 2008. 2008;

146. Offermann P, Blom S, Schönherr M, Bub U. Artifact types in information systems design science - A literature review. In: International Conference on Design Science Research in Information Systems . Springer, Berlin, Heidelberg; 2010. p. 77–92.

147. Prat N, Comyn-wattiau I, Cnam C. DESIGN-SCIENCE RESEARCH – A HOLISTIC VIEW. In: PACIS. 2014. p. 23.

148. Li L, Peng T, Kennedy J. A Rule Based Taxonomy of Dirty Data. GSTF Int J Comput. 2011;1(2).

149. Pipino LL, Lee YW, Wang RY, Lowell Yang Lee MW, Yang RY. Data Quality Assessment. Commun ACM April . 2002 ;45.

150. Thawer SG, Chacky F, Runge M, Reaves E, Mandike R, Lazaro S, et al. Sub-national stratification of malaria risk in mainland Tanzania: a simplified assembly of survey and routine data. Malar J . 2020;19:177.

151. Maina I, Wanjal P, Soti D, Kipruto H, Droti B, Boerma T. Using health-facility data to assess subnational coverage of maternal and child health indicators, Kenya. Bull World Health Organ . 2017 ;95(10):683–94.

152. Wilhelm JA, Qiu M, Paina L, Colantuoni E, Mukuru M, Ssengooba F, et al. The impact of PEPFAR transition on HIV service delivery at health facilities in Uganda. Rockers P, editor. PLoS One . 2019;14(10):e0223426.

153. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLOS Med . 2015;12(10):e1001885.

154. Leahey E, Entwisle B, Einaudi P. Diversity in everyday research practice: The case of data editing. Sociol Methods Res. 2003 Aug;32(1):64–89.

155. Langouri MA, Zheng Z, Chiang F, Golab L, Szlichta J. Contextual Data Cleaning. 2018 IEEE 34th Int Conf Data Eng Work. 2018;21–4.

156. Ngugi PN, Gesicho MB, Babic A, Were MC. Assessment of HIV data reporting performance by facilities during EMR systems implementations in Kenya. Stud Health Technol Inform. 2020;272:167–70.

157. Dyson S, Brown B. Social Theory and Applied Health Research. Bryman A, editor. Maidenhead,England: McGraw-Hill Education; 2006.

158. Roberts P, Priest H. Reliability and validity in research. . Vol. 20, Nursing standard (Royal College of Nursing (Great Britain) : 1987). Nurs Stand; 2006. p. 41–5.

159. Cypress BS. Rigor or reliability and validity in qualitative research: Perspectives, strategies, reconceptualization, and recommendations. Dimens Crit Care Nurs. 2017;36(4):253–63.

160. WHO. Module 3 Data verification and system assessment DATA QUALITY REVIEW . 2017. Available from: http://apps.who.int/bookorders.

161. Lincoln YS, Guba EG. Trustworthiness and Naturalistic Evaluation. New Dir Progr Eval. 1986;73–84(30).

162. Cheng HG, Phillips MR. Secondary analysis of existing data: opportunities and implementation. Heal Shanghai Arch Psychiatry . 2014 ;26(6):371–5.

163. Ola O, Sedig K. Beyond simple charts: Design of visualizations for big health data. Online J Public Health Inform . 2016;8(3):e195.

Appendicies

Publications

- Paper I
- Paper II
- Paper III
- Paper IV
- Paper V

Errata

Page	Changes made	Action
3	*…and my parents* ***Mr.*** *Robert O. Gesicho and* ***Mrs****. Esther K. Gesicho*	*Bolded words were missing*
17	…rendering some of the countries data rich but information poor [11–13] **.**	Full stop missing
20	**Comprehensive**, systematic, and transparent procedures for data cleaning were presented,	*Bolded word was missing as it appears in other subsequent texts*
28	Limitations and mechanisms ~~uesd~~ **used** to promote data use	*Misspelling: uesd to used*
30	… trained on **the** importance	*Bold word was missing*
32	…HIV-indicator reporting data for health facilities in Kenya and extracted **them** from the DHIS2	*Bold word was missing*
33	…reports that have been completed **(submitted)** on time	*Bold word was missing as appears in other texts*
42	integrated together with the necessary additional data to create the initial data set	*Strikethrough word was removed as is repetitive*
60	…the qualitative **inquiry**	*Misspelling: Enquiry to inquiry*
62/76	…Sub-cycle 4 (Paper **VI**: Manuscript)	*Typo correction of paper number V to VI*
83	…this qualitative **inquiry** revealed that utilization of EMRS in reporting	*Misspelling: Enquiry to inquiry*
89	.. validity of study that is largely advocated **through** transparent reporting	*Misspelling: though to through*
91	Further still, **qualitative** inquiries supplemented quantitative analysis as…	*Misspelling: qualitive to qualitative*

Gesicho et al. BMC Med Inform Decis Mak (2020) 20:293
https://doi.org/10.1186/s12911-020-01315-7

BMC Medical Informatics and Decision Making

RESEARCH ARTICLE Open Access

Data cleaning process for HIV-indicator data extracted from DHIS2 national reporting system: a case study of Kenya

Milka Bochere Gesicho[1,4*], Martin Chieng Were[2,4] and Ankica Babic[1,3]

Abstract

Background: The District Health Information Software-2 (DHIS2) is widely used by countries for national-level aggregate reporting of health-data. To best leverage DHIS2 data for decision-making, countries need to ensure that data within their systems are of the highest quality. Comprehensive, systematic, and transparent data cleaning approaches form a core component of preparing DHIS2 data for analyses. Unfortunately, there is paucity of exhaustive and systematic descriptions of data cleaning processes employed on DHIS2-based data. The aim of this study was to report on methods and results of a systematic and replicable data cleaning approach applied on HIV-data gathered within DHIS2 from 2011 to 2018 in Kenya, for secondary analyses.

Methods: Six programmatic area reports containing HIV-indicators were extracted from DHIS2 for all care facilities in all counties in Kenya from 2011 to 2018. Data variables extracted included reporting rate, reporting timeliness, and HIV-indicator data elements per facility per year. 93,179 facility-records from 11,446 health facilities were extracted from year 2011 to 2018. *Van den Broeck* et al.'s framework, involving repeated cycles of a three-phase process (data screening, data diagnosis and data treatment), was employed semi-automatically within a generic five-step data-cleaning sequence, which was developed and applied in cleaning the extracted data. Various quality issues were identified, and Friedman analysis of variance conducted to examine differences in distribution of records with selected issues across eight years.

Results: Facility-records with no data accounted for 50.23% and were removed. Of the remaining, 0.03% had over 100% in reporting rates. Of facility-records with reporting data, 0.66% and 0.46% were retained for voluntary medical male circumcision and blood safety programmatic area reports respectively, given that few facilities submitted data or offered these services. Distribution of facility-records with selected quality issues varied significantly by programmatic area ($p < 0.001$). The final clean dataset obtained was suitable to be used for subsequent secondary analyses.

Conclusions: Comprehensive, systematic, and transparent reporting of cleaning-process is important for validity of the research studies as well as data utilization. The semi-automatic procedures used resulted in improved data quality for use in secondary analyses, which could not be secured by automated procedures solemnly.

Keywords: Data-cleaning, dhis2, HIV-indicators, Data management

*Correspondence: milcagesicho@gmail.com
[1] Department of Information Science and Media Studies, University of Bergen, Bergen, Norway
Full list of author information is available at the end of the article

Background

Routine health information systems (RHIS) have been implemented in health facilities in many low-and middle-income countries (LMICs) for purposes such as facilitating data collection, management and utilization [1]. In order to ensure effectiveness of HIV programs, accurate, complete and timely monitoring and evaluation (M&E) data generated within these systems are paramount in decision-making such as resource allocation and advocacy [2]. Monitoring and Evaluation (M&E) plays a key role in planning of any national health program. De Lay et al. defined M&E as "acquiring, analyzing and making use of relevant, accurate, timely and affordable information from multiple sources for the purpose of program improvement [2]."

In order to provide strategic information needed for M&E activities in low- and middle-income countries (LMICs), reporting indicators have been highly advocated for use across many disease domains, with HIV indicators among the most common ones reported to national-level facilities in many countries [3–5]. As such, health facilities use pre-defined HIV-indicator forms to collect routine HIV-indicator data on various services provided within the facility, which are submitted to the national-level [6].

Over the years, national-level data aggregation systems, such as the District Health Information Software 2 (DHIS2) [7], have been widely adopted for use in collecting, aggregating and analyzing indicator data. DHIS2 has been implemented in over 40 LMICs with the health indicator data reported within the system used for national- and regional-level health-related decision-making, advocacy, and M&E [8]. Massive amounts of data have been collected within health information systems such as DHIS2 over the past several years, thus providing opportunities for secondary analyses [9]. However, these analyses can only be adequately conducted if the data extracted from systems such as DHIS2 are of high quality that is suitable for analyses [10].

Furthermore, data within health information systems such as DHIS2, are only as good as their quality, as this is salient for decision-making. As such, various approaches have been implemented within systems like DHIS2 to improve data quality. Some of these approaches include: (a) validation during data entry in order to ensure data are captured using the right formats and within pre-defined ranges and constraint; (b) user-defined validation rules; (c) automated outlier analysis functions such as standard deviation outlier analysis (identifies data values that are numerically extreme from the rest of the data), and minimum and maximum based outlier analysis (identifies data values outside the pre-set maximum and minimum values); and (d) automated calculations and reporting of data coverage and completeness [11]. WHO data quality tool has also been incorporated with DHIS2 to identify errors within the data in order to determine the next appropriate action [12]. Given that this tool is a relatively new addition to the DHIS2 applications, it is still being progressively improved and implemented in countries using DHIS2 [13].

Despite data quality approaches having been implemented within DHIS2, data quality issues remain a thorny problem, with some of the issues emanating from the facility level [14]. Real-life data like that found in DHIS2 are often "dirty" consisting of issues such as; incomplete, inconsistent, and duplicated data [15]. Failure to detect data quality issues and to clean these data can lead to inaccurate analyses outcomes [13]. Various studies have extracted data from DHIS2 for analyses [16–20]. Nonetheless, few studies attempt to explicitly disclose the data cleaning strategies used, resulting errors identified and the action taken [16–18]. In addition, some of these studies largely fail to exhaustively and systematically describe the steps used in data cleaning of the DHIS2 data before analyses are done [19, 20].

Ideally, data cleaning should be done systematically, and good data cleaning practice requires transparency and proper documentation of all procedures taken to clean the data [21, 22]. A closer and systematic look into data cleaning approaches, and a clear outlining of the distribution or characteristics of data quality issues encountered in DHIS2 could be instructive in informing approaches to further ensure higher quality data for analyses and decision-making. Further, employment of additional data cleaning steps will ensure that good quality data is available from the widely deployed DHIS2 system for use in accurate decision-making and knowledge generation.

In this study, data cleaning is approached as a process aimed at improving the quality of data for purposes of secondary analyses [21]. Data quality is a complex multidimensional concept. Wang and Strong categorized these dimensions as: intrinsic data quality, contextual data quality, representational and accessibility data quality [23]. Intrinsic data quality focuses on features that are inherent to data itself such as accuracy [23]. Contextual data quality focuses on features that are relevant in the context for the task for data use such as value-added, appropriate amount of data, and relevancy [23]. Representational and accessibility data quality highlights features that are salient within the role of the system such as interpretability, representational consistency, and accessibility [23]. Given that data quality can be subjective and dependent on context, various studies have specified context in relation to data quality [24–26]. Bolchini et al. specify context by tailoring data that are relevant for a given particular use case [27]. Bolchini et al. further posit that the process of separating noise (information not relevant to a specific task) to obtain only useful information, is not an easy task [27]. In this study, data cleaning is approached from a contextual standpoint, with the intention of retaining only relevant data for subsequent secondary analyses.

Therefore, the aim of this study is to report on the method and results of a systematic and replicable data cleaning approach employed on routine HIV-indicator data reports gathered within DHIS2 from 2011 to 2018 (8 year period), to be used for subsequent secondary analyses, using Kenya as a reference country case. This approach has specific applicability to the broadly implemented DHIS2 national reporting system. Our approach is guided by a conceptual data-cleaning framework, with a focus on uncovering data quality issues often missed by existing automated approaches. From our evaluation, we provide recommendations on extracting and cleaning data for analyses from DHIS2, which could be of benefit to M&E teams within Ministries of Health and by researchers to ensure high quality data for analyses and decision-making.

Methods

Data cleaning and data quality assessment approaches

Data cleaning is defined as "the process used to determine inaccurate, incomplete, or unreasonable data and then improving the quality through correction of detected errors and omissions" [28]. Data cleaning is essential to transform raw data into quality data for purposes such as analyses and data mining [29]. It is also an integral step in the knowledge discovery of data (KDD) process [30].

There exists various issues within the data, which necessitate cleaning in order to improve its quality [31–33]. An extensive body of work exists on how to clean data. Some of the approaches that can be employed include quantitative or qualitative methods. Quantitative approaches employ statistical methods, and are largely used to detect outliers [34–36]. On the other hand, qualitative techniques use patterns, constraints, and rules to detect errors [37]. These approaches can be applied within automated data cleaning tools such as ARKTOS, AJAX, FraQL, Potter's Wheel and IntelliClean [33, 37, 38].

In addition, there are a number of frameworks used in assessment of data quality in health information systems, which can be utilized by countries with DHIS2. The Data Quality Review (DQR) tool developed in collaboration with WHO, Global Fund, Gavi and USAID/MEASURE Evaluation provides a standardized approach that aims at facilitating regular data quality checks [39]. Other tools for routine data quality assessments include the MEASURE Evaluation Routine Data Quality Assessment Tool (RDQA) [40] and WHO/IVB Immunization Data Quality Self-Assessment (DQS) [41].

Some of the data quality categories (intrinsic, contextual, representational and accessibility) [23], have been used in cleaning approaches as well as the data quality frameworks developed. A closer examination of the aforementioned approaches reveals focus on assessing intrinsic data quality aspects, which can be categorized further to syntactic quality (conformance to database rules) and semantic quality (correspondence or mapping to external phenomena) [42].

Moreover, while tools and approaches exist for data quality assessments as well as data cleaning, concerted efforts have been paced on assessment of health information system data quality [39, 40], as opposed to cleaning approaches for secondary analyses, which are largely dependent on the context for data use [24]. Wang and Strong posited the need for considering data quality with respect to context of the tasks, which can be a challenge as tasks and context vary by user needs [23]. Therefore, specifying the task and relevant features for the task, can be employed for contextual data quality [23, 43].

With this in mind and based on our knowledge, no standard consensus-based approach exists to ensure that replicable and rigorous data cleaning approaches and documentation are applied on extracted DHIS2 data to be used in secondary analyses. As such, ad hoc data cleaning approaches have been employed for the extracted data prior to analyses [16–18]. Moreover, whereas some studies provide brief documentation of data cleaning procedures used [19], others lack documentation, leaving the data cleaning approaches used undisclosed and behind-the-scenes [20]. Failure to disclose approaches used makes it difficult to replicate data cleaning procedures, and to ensure that all types of anomalies are systematically addressed prior to use of data for analysis and decision-making. Furthermore, the approach used in data extraction and cleaning affects the analysis results [21].

Oftentimes, specific approaches are applied based on the data set and the aims of the cleaning exercise [10, 44, 45]. Dziadkowiec et al. used Khan's framework to clean data extracted from relational database of an Electronic Health Records (EHR) (10). In their approach, intrinsic data quality was in our view considered in data cleaning with focus on syntactic quality issues (such as conforming to integrity rules). Miao et al. proposed a data cleaning framework for activities that involve secondary analysis of an EHR [45], which in our view considered intrinsic data quality with focus on semantic quality (such as completeness and accuracy). Savik et al. approached data cleaning in our view from a contextual perspective, which entailed preparing the dataset that is appropriate for the intended analysis [44].

In this study, we approach data cleaning from a contextual perspective, whereby only data fit for subsequent analyses is retained. Based on our data set, our study's data cleaning approach was informed by a conceptual data-cleaning framework proposed by Van den Broeck et al. [21]. Van den Broeck et al.'s framework was used because it provides a deliberate and systematic data cleaning guideline that is amenable to being tailored towards cleaning data extracted from DHIS2. This framework presents data cleaning as a three-phase process involving repeated cycles of data screening, data diagnosis, and data editing of suspected data abnormalities. The screening process involves identification of lacking or excess data, outliers and inconsistencies and strange patterns [21]. Diagnosis involves determination of errors or missing data and any true extremes and true normal [21]. Editing involves correction or deleting of any identified errors [21]. The various phases in Van den Broeck et al.'s framework have also been applied in various settings [46, 47]. Human-driven approaches complemented by automatic approaches were also used in the various data cleaning phases in thus study. Human-involvement in data cleaning has also been advocated in other studies [35].

Study setting

This study was conducted in Kenya, a country in East Africa. Kenya adopted DHIS2 for use for its national reporting in 2011 [7]. The country has 47 administrative counties, and all the counties report a range of healthcare indicator data from care facilities and settings into the DHIS2 system. For the purposes of this study, we focused specifically on HIV-indicator data reported within Kenya's DHIS2 system, given that these are the most comprehensively reported set of indicators into the system.

Kenya's DHIS2 has enabled various quality mechanisms to deal with HIV data. Some of these include data validation rules, outlier analysis and minimum and maximum ranges, which have been implemented at the point of data entry. DHIS2 data quality tool is also an application that was included in DHIS2 to supplement the in-built data quality mechanisms [12]. Nonetheless it was not actively in use during our study period 2011–2018. The quality mechanisms as well as the DHIS2 quality tool consider intrinsic data quality aspects.

Data cleaning process

Adapting the *Van den Broeck* et al.*'s* framework, a step-by-step approach was used during extraction and cleaning of the data from DHIS2. These steps are generic and can be replicated by others conducting robust data cleaning on DHIS2 for analyses. These steps are outlined below:

i **Step 1—**Outline the analyses or evaluation questions: Prior to applying the *Van den Broeck* et al.*'s* conceptual framework, it is important to identify the exact evaluations or analyses to be conducted, as this helps define the data cleaning exercise.

j **Step 2—**Description of data and study variables: This step is important for defining the needed data elements that will be used for the evaluation data set.

k **Step 3—**Create the data set: This step involves identifying the data needed and extracting data from relevant databases to generate the final data set. Oftentimes, development of this database might require combining data from different sources.

l **Step 4—**Apply the framework for data cleaning: During this step, the three data cleaning phases (screening, diagnosis, and treatment) in *Van den Broeck* et al.*'s* framework are applied on the data set created.

m **Step 5—**Analyze the data: This step provides a summary of the data quality issues discovered, the eliminated data after the treatment exercise, and the retained final data set on which analyses can then be done.

Application of data cleaning process: Kenya HIV-indicator reporting case example

In this section, we present the application of the data cleaning sequence above using Kenya as case example. It is worth noting that in this study, the terms 'programmatic area report' and 'report' are used interchangeably as they contain the same meaning given that a report represents a programmatic area, and contains a number of indicators.

Step 1: Outline the analyses or evaluation questions and goals

For this reference case, DHIS2 data had to undergo the data cleaning process prior to use of the data for an evaluation question on 'Performance of health facilities at meeting the completeness and timeliness facility reporting requirements by the Kenyan Ministry of Health (MoH)'. The goal was to identify the best performing and poor performing health facilities at reporting within the country, based on completeness and timeliness in submitting their reports into DHIS2.

This study only attempts to clean the data for further subsequent analyses. Thus, the actual analyses and evaluation will be conducted using the final clean data in a separate study.

Step 2: Description of data and study variables

HIV-indicator data in Kenya are reported into DHIS2 on a monthly basis by facilities offering HIV services using the MOH-mandated form called "*MOH 731- Comprehensive HIV/AIDS Facility Reporting Form*" (MOH731). As of 2011–2018, MOH 731 consisted of six programmatic areas representing six independent reports containing HIV-indicators to be reported [see Additional file 1]. The six reports and the number of indicators reported in each include: (1) HIV Counselling and Testing (HCT)—14 indicators; (2) Prevention of Mother-to-Child transmission (PMTCT)—40 indicators; (3) Care and Treatment (CrT)—65 indicators; (4) Voluntary Medical Male Circumcision (VMMC)—13 indicators; (5) Post-Exposure Prophylaxis (PEP)—14 indicators; and (6) Blood Safety (BS)—3 indicators.

Each facility offering HIV services is expected to submit reports with indicators every month based on the type(s) of services offered by that facility. Monthly due date for all reports are defined by the MoH, and the information on the expected number of reports per facility.

For our use case, we wanted to create a data set for secondary analyses, which was to determine performance of facilities at meeting the MoH reporting requirements (facility reporting completeness and timeliness of reporting). Hence, retain only facilities offering services for any of the six programmatic areas. Completeness in reporting by facilities within Kenya's DHIS2 is measured as a continuous variable starting at 0% to 100% and identified within the system by a variable called '*Reporting Rate (RR)*'. The percentage RR is calculated automatically within DHIS2 as the actual number of reports submitted by each facility into DHIS2 divided by the expected number of reports from the facility multiplied by100 (Percentage RR = actual number of submitted reports/ expected number of reports * 100). Given that MOH731 reports should be submitted by facilities on a monthly routine, the expected number of monthly reports per programmatic area per year is 12 (one report expected per month). It should be noted that this Reporting Rate calculation only looks at report submission and not the content within the reports. Given that facilities offering any of the HIV services are required to submit the full MOH731 form containing six programmatic area reports, zero (0) cases are reported for indicators where services are not provided, which appear as blank reports in DHIS2. As such, a report may be submitted as blank or have missing indicators but will be counted as complete (facility reporting completeness) simply because it was submitted. Timeliness is calculated based on whether the reports were submitted by the 15th day of the reporting month as set by the MoH. Timeliness is represented in DHIS2 as '*Reporting Rate on Time (RRT)*' and is also calculated automatically. The percentage RRT for a facility is measured as a percentage of the actual number of reports submitted on time by the facility divided by the expected number of reports multiplied by 100 (Percentage RRT = actual number of reports submitted on time/ expected number of reports * 100). Annual reports were therefore generated from DHIS2 consisting of percentage Reporting Rate and Reporting Rate on Time, which were extracted per facility, per year.

Step 3: Create the data set

After obtaining Institutional Review and Ethics Committee (IREC) approval for this work, we set out to create our database from three data sources as outlined below:

(1) *Data Extracted from DHIS2*: Two sets of data were extracted from DHIS2 to Microsoft Office Excel (version 2016). For the first data set, we extracted variables from DHIS2 for all HIV programmatic area reports submitted from all health facilities in all 47 counties in Kenya between the years 2011 and 2018, with variables grouped by year. Variables extracted from DHIS2 by year included: facility name, programmatic area report (e.g. Blood Safety), expected number of reports, actual number of submitted reports, actual number of reports submitted on time, cumulative Reporting Rate by year (calculated automatically in DHIS2) and cumulative Reporting Rate on Time by year (calculated automatically in DHIS2) [see Additional file 2]. The extracted data for Reporting Rate and Reporting Rate on Time constituted to the annual reports in the six programmatic areas for years 2011–2018, for the respective health facilities.

For the second data set, we extracted the HIV-indicator data elements submitted within each annual programmatic area report by the health facilities for all the six programmatic areas for every year under evaluation [see Additional file 1].The annual report contained cumulative HIV-indicator data elements gathered in each programmatic area per facility, per year.

In addition, extracting the aforementioned datasets from 2011 to 2018 resulted to repeated occurrence of the facility variable in the different years. For example, facilities registered in DHIS2 in 2011 will appear in subsequent years resulting to eight occurrences within the 8 years (2011–2018) per programmatic area report (e.g. Blood Safety). These resulted to a facility containing the following variables per row: facility name, year, percentage Reporting Rate, and percentage Reporting Rate on Time for the six programmatic area reports. In this study, the facility data per row was referred to as 'facility record'.

(2) *Facility Information*: We augmented the DHIS2 data with detailed facility information derived from Kenya Master Facility List (KMFL). This information included facility level (II–VI), facility type (such as dispensary, health center, medical clinic) and facility ownership (such as private practice, MoH-owned, owned by a non-governmental organization).
(3) *Electronic Medical Record Status:* We used the Kenya Health Information Systems (KeHIMS) list, which contains electronic medical records (EMR) implemented in health facilities in Kenya, to incorporate information on whether the facility had an EMR or not. Information from these three sources were merged into a single data set as outlined in Fig. 1.

Step 4: Application of the framework for data cleaning

Figure 2 outlines the iterative cleaning process we applied adapting *Van den Broeck* et al.*'s* framework. Data cleaning involved repeated cycles of screening, diagnosis, and treatment of suspected data abnormalities, with each cycle resulting in a new data set. Details of the data cleaning process is outlined in Fig. 2.

a) Screening phase

During the screening phase, five types of oddities need to be distinguished, namely: lack or excess of data; outlier (data falling outside the expected range); erroneous inliers; strange patterns in distributions and unexpected analysis results [21].

For determining errors, we used Reporting Rate and Reporting Rate on Time as key evaluation variables. Reporting Rate by itself only gives a sense of the proportion of expected reports submitted but does not evaluate whether exact HIV-indicator data elements are included within each report. To evaluate completion of HIV-indicator data elements within each of the programmatic area reports that were submitted, we created a new variable named 'Cumulative Percent Completion (CPC)'. Using the annual report extracted for HIV-indicator data elements per facility, Cumulative Percent Completion was calculated by counting the number of non-blank values and dividing this by the total number of indicators for each programmatic area. As such, if a facility has

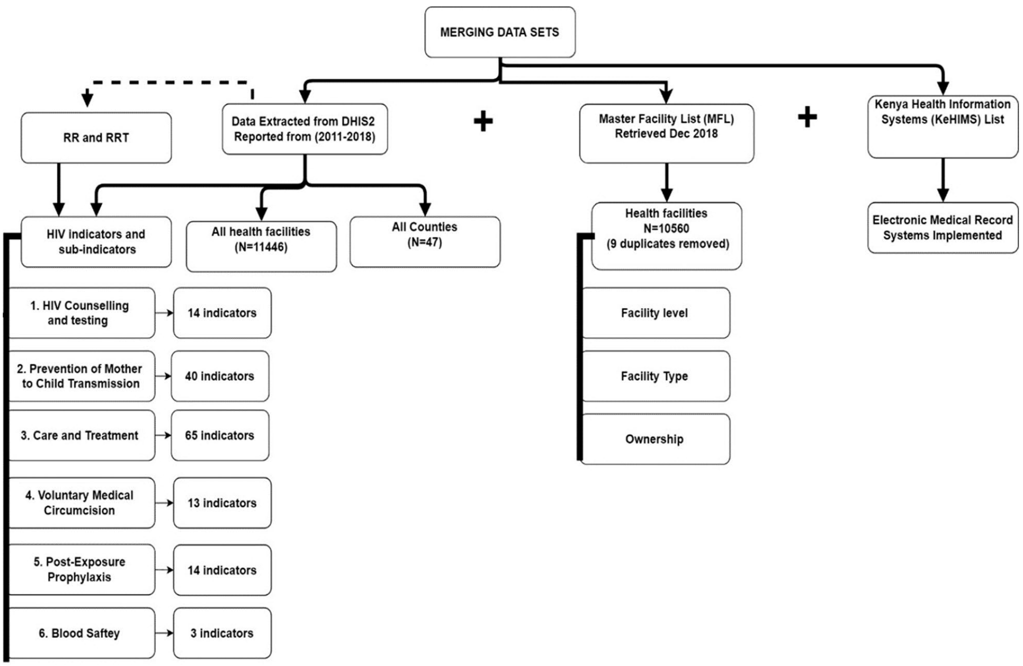

Fig. 1 Creation of the evaluation data set

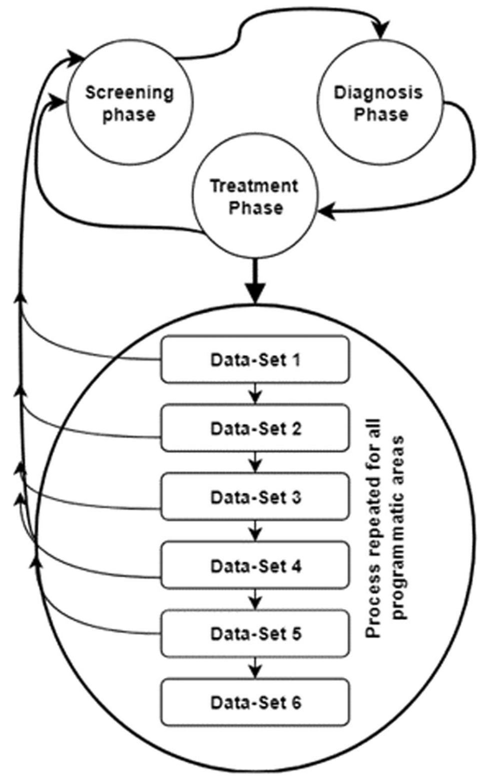

Fig. 2 Repeated cycles of data cleaning

reported on 10 out of 40 indicators in an annual report, it will have 25 percent on completeness. Therefore, Cumulative Percent Completion provides an aggregate annual summary of the proportion of expected indicator values that are completed within submitted reports. The results for Cumulative Percent Completion were then included as variables in the facility-records, described in step 3, section 1. This resulted to a facility-record containing the following variables per row: facility name, year, percentage Reporting Rate, percentage Reporting Rate on Time and Cumulative Percent Completion for the six programmatic areas.

b **Diagnostic phase**

The diagnostic phase enables clarification of the true nature of the worrisome data points, patterns, and statistics. Van den Broeck et al. posits possible diagnoses for each data point as: erroneous, true extreme, true normal or idiopathic (no diagnosis found, but data still suspected to having errors) [21]. We used a combination of Reporting Rate, Reporting Rate on Time and Cumulative Percent Completion to detect various types of situations (errors or no errors) for each facility per annual report (Table 1). Using the combination of Cumulative Percent Completion, Reporting Rate, and Reporting Rate on Time we were able to categorize the various types of situations to be used in diagnosis for every year a facility reported into DHIS2 (Table 1). In this table, "0" represents a situation where percentage is zero; "X" represents a situation where percentage is above zero; and "> 100%" represents a situation where percentage is more than 100. This data points

Table 1 Categorization of the various situations within DHIS2 and actions taken

Situation	CPC[a]	RR[b]	RRT[c]	Diagnosis	Action
A	0	0	0	Nothing was reported by facilities during this period, signifying that the facility does not report to DHIS2. This could be a true normal	Facility records excluded
B	0	X	X	Submitted reports might be on time, but are empty. Can result from programs wanting to have full MOH731 submission even though they do not offer services in all the 6 programmatic areas—hence submitting empty reports from non-required programmatic areas (Report is useless to decision-maker as it is empty)	Facility records excluded
C	0	X	0	Submitted reports are empty and not on time (Report is useless to decision-maker as it is empty and not on time)	Facility records excluded
D	X	0	0	No values present for RR and RRT. However, the reports are not empty	Facility records excluded
E	X	> 100%	X	Erroneous records as percentage RR cannot go beyond 100 as this is not logically possible	Facility records excluded
F	X	> 100%	> 100%	Erroneous records percentage RR and RRT cannot go beyond 100 as this is not logically possible	Facility records excluded
G	X	X	X	Reports submitted on time with relevant indicators included. Ideal situation	Facility records included
H	X	X	0	Submitted reports with data elements in them, but not submitted in a timely manner	Facility records included

[a] *CPC* cumulative percent completion, [b] *RR* reporting rate, [c] *RRT* reporting rate on time

were considered as erroneous records as the percentage reporting rate cannot go beyond 100 as this is not logically possible. Based on the values per each of the three variables, it was possible to diagnose the various issues within DHIS2 (Diagnosis Column).

For each programmatic area report (e.g. Blood Saftey) we categorized facilities by year and variables. All health facilities with an average Cumulative Percent Completion, Reporting Rate, and Reporting Rate on Time of zero (0) across all reports were identified as not having reported for the year and were henceforth excluded – as demonstrated by examples of Facility A and B in Table 2.

Beyond categorization of the various situations by report type, facility and year as defined above, errors related to duplicates were also identified using two scenarios. The first scenario of duplicates included a situation where health facilities had similar attributes such as year, name and county, with different data for Reporting Rate and Reporting Rate on Time. The second scenario of duplicates involves a situation where health facilities had similar attributes such as year, name and county, with similar data for Reporting Rate, and Reporting Rate on Time.

c **Treatment phase**

This is the final stage after screening and diagnosis, and entails deciding on the action point of the problematic records identified. Van den Broeck et al. limit the action points to correcting, deleting or leaving unchanged [21]. Based on the diagnosis illustrated in Table 1, facility-records in situation A-F were deleted hence excluded from the study. Duplicates identified in the scenarios mentioned were also excluded from the study. As such, for duplicates where health facilities had similar attributes such as year, name, and county, with different data for Reporting Rate, and Reporting Rate on Time, all entries were deleted. For duplicates where health facilities had similar attributes such as year, name, and county, with similar data for Reporting Rate, and Reporting Rate on Time, only one entry was deleted. Only reports in situation G and H were considered ideal for the final clean data set.

Step 5: Data analysis

The facility-records were then disaggregated to form six individual data sets representing each of the programmatic areas containing the following attributes: facility name, year, Cumulative Percent Completion, percentage Reporting Rate and percentage Reporting Rate on Time, as well as the augmented data on facility information and EMR status. The disaggregation was because facilities offer different services and do not necessarily report indicators for all the programmatic areas. SPSS was used to analyze the data using frequency distributions and cross tabulations in order to screen for duplication and outliers. Individual health facilities with frequencies of more than eight annual reports for a specific programmatic area were identified as duplicates. The basis for this is that the maximum annual reports per specific programmatic area for an individual health facility has to be eight, given that data was extracted within an eight-year period. From the cross tabulations, percentage Reporting Rate and percentage Reporting Rate on Time that were above 100% were identified as erroneous records.

After the multiple iterations of data cleaning as per Fig. 2, where erroneous data were removed by situation type (identified in Table 1), a final clean data set was available and brought forward to be used in a separate study for subsequent secondary analyses (which include answering the evaluation question in step 1). At the end of the data cleaning exercise, we determined the percentage distribution of the various situation types that resulted in the final data set. The percentages were calculated by dividing the number of facility-records in each situation type by the total facility-records in each programmatic area respectively, which was then multiplied by 100. As such, only data sets disaggregated into the six programmatic areas were included in the analysis. Using this analysis and descriptions from Table 1, we selected situation B, and situation D, in order to determine if there is a difference in distribution of facility records containing the selected situation types in the six programmatic areas across the 8 years (2011–2018).

This will enable comparing distribution of facility records by programmatic area categorized by situation B

Table 2 Example of sectional illustration of first data set containing facility records

Year	Organisation unit	CPC-HCT	RR-HCT	RRT-HCT	CPC-BS	RR-BS	RRT-BS	**	Avg-CPC	Avg-RR	Avg-RRT
2016	Facility A	0	0	0	0	0	0	0	0	0	0
2016	Facility B	0	0	0	0	0	0	0	0	0	0
2017	Facility C	10	90	80	100	90	80	0	50	60	50

CPC cumulative percentage completion, *RR-HCT* reporting rate HIV counselling and testing, *RRT* reporting rate on time, *BS* blood safety, *Avg* average, ** remaining four reports with the same variable sequence and situation D. The data contains related samples and is not normally distributed. Therefore, a Friedman analysis of variance (ANOVA) was conducted to examine if there is a difference in distribution of facility reports by programmatic area across all years N = 8 (2011–2018) for the selected situation types. As such, the variables analyzed include year, situation type, programmatic area, and unit of analysis include number of records in each situation type for a programmatic area. The distribution of facility-records was measured in all the six programmatic areas across the eight years and categorized by situation type. Wilcoxon Signed Rank Test were carried out as post hoc tests to compare significances in facility report distribution within the programmatic areas.

Below, we report on findings from the iterative data cleaning exercise and the resulting clean data set. The results further illustrate the value of the data cleaning exercise.

Results

Figure 3 reports the various facility records at each cycle of the data cleaning process and the number (proportion) of excluded facility-records representing data with errors at each cycle.

The proportion of the resultant dataset after removal of the various types of errors from the facility records is represented in Table 3. A breakdown of reporting by facilities in descending order based on facility records retained after cleaning in dataset 4 is as follows; 93.98% were retained for HIV Counselling and Testing (HTC), 83.65% for Prevention of Mother to Child Transmission (PMTCT), 43.79% for Care and Treatment (CRT), 22.10% for Post Exposure Prophylaxis (PEP), 0.66% for Voluntary Medical Male Circumcision (VMMC), and 0.46% for Blood Safety (BS).

Situations where data was present in reports, but no values present for Reporting Rate and Reporting Rate on Time (Situation D); and scenarios with empty reports (Situation B) were analyzed (Fig. 4). This was in order to examine whether there are differences in distribution of facility records by programmatic area across the eight years, categorized by situation type. Most facilities submitted PEP empty reports (18.04%) based on data set 4 as shown in Fig. 4.

Overall Friedman Tests results for distribution of records with situation B and situation D in the various programmatic areas reveal statistically significant differences in facility record distribution (p = 0.001) across the eight years. Specific mean rank results categorized by error type are described in subsequent paragraphs.

Friedman Tests results for empty reports (Situation B) reveal that PEP had the highest mean rank of 6.00 compared to the other programmatic areas CT (3.50), PMTCT (4.88) CrT (2.00), VMMC (3.00), PEP and BS (1.63). Post hoc tests presented in Table 4 also reveal that PEP had higher distribution of facility records in situation B (0XX) in all the eight years.

Friedman Tests results for distribution of records with situation D (X00) reveal that PMTCT and CrT had the highest mean rank of 5.88 and 5.13 respectively compared to the other programmatic areas CT (3.00), VMMC (3.06), PEP (2.88) and BS (1.06). Post hoc tests presented in Table 5 reveal that PMTCT and CrT had higher distribution of facility records in situation D (X00) in all the 8 years.

Discussion

Systematic data cleaning approaches are salient in identifying and sorting issues within the data resulting to a clean data set that can be used for analyses and decision-making [21]. This study presents the methods and results of systematic and replicable data cleaning approach employed on routine HIV-indicator data reports in preparation for secondary analyses.

For data stored in DHIS2, this study assumed that the inbuilt data quality mechanisms dealt with the predefined syntactical data quality aspects such as validation rules. As such, the contextual approach to data cleaning was employed on extracted data from DHIS2 with the aim of distinguishing noise (data that are not relevant for intended use or of poor quality), from relevant data as presented by the various situations in Table 1. As demonstrated in this study, identifying various issues within the data may require a human-driven approach as inbuilt data quality checking mechanisms within systems may not have the benefit of a particular knowledge. Furthermore, these human augmented processes also facilitated diagnosis of the different issues, which would have gone unidentified. For instance, our domain knowledge about health facility HIV reporting enabled us to identify the various situations described in Table 1. This entailed examining more than one column at a time of manually integrated databases and using the domain knowledge in making decisions on actions to take on the data set (treatment phase). Similarly, Maina et al. also used domain knowledge on maternal and child bearing programmes in adjusting for incomplete reporting [48].In addition, descriptive statistics such as use of cross tabulations and frequency counts complemented the human-driven processes, in order to identify issue within the data such as erroneous records (screening phase).

The use of Cumulative Percent Completeness (CPC) in this study facilitated screening and diagnosis of problematic issues highlighted in similar studies that are consistent with our findings. These include identifying and dealing with non-reporting facilities (situation A), and

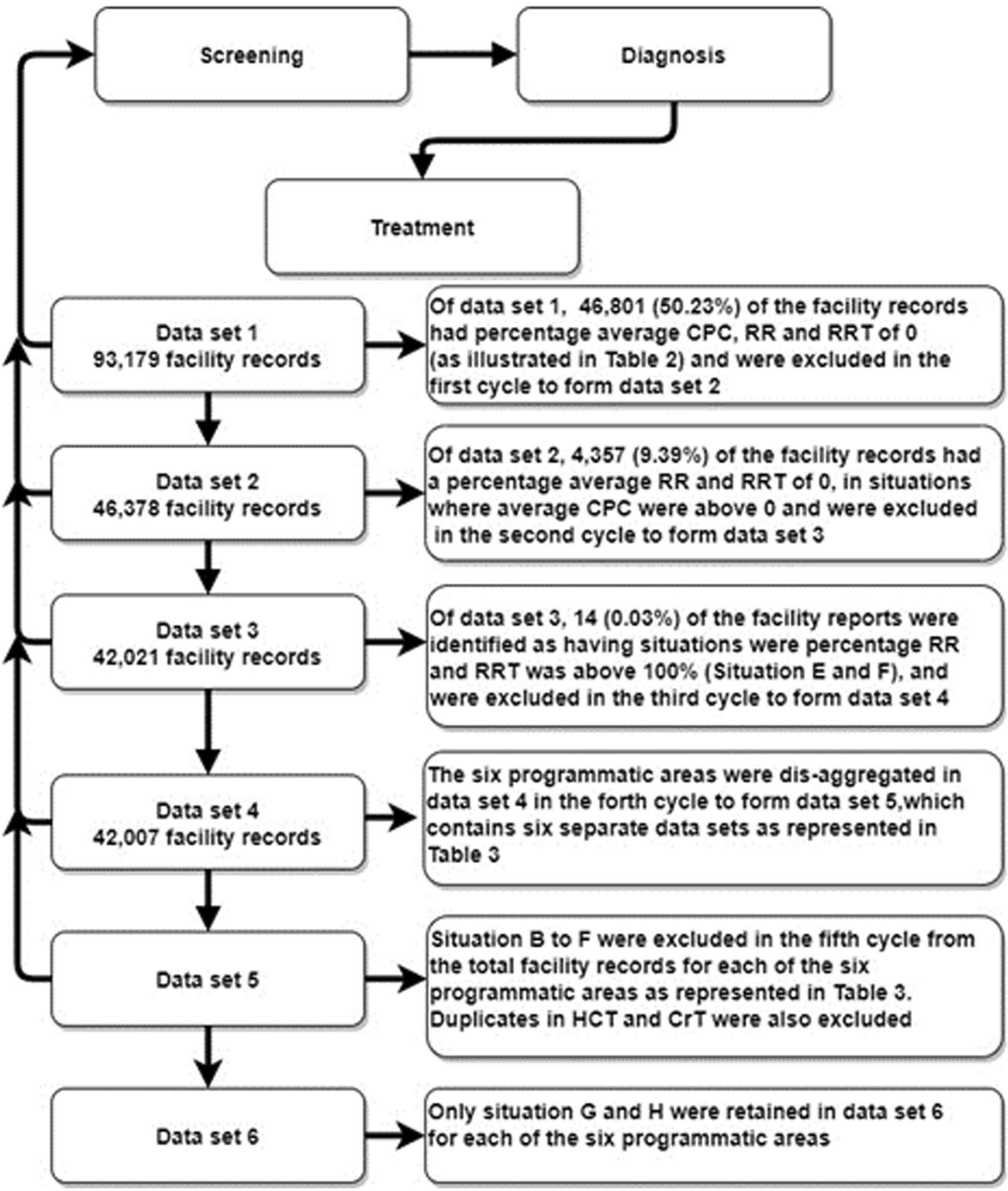

Fig. 3 Data cleaning process

non-service providing facilities (situation B and C) in a data set [19, 48]. This comes about as some of the reports extracted contain blanks, as DHIS2 is unable to record zeros as identified in other studies [16–19, 49]. As such, DHIS2 is unable to distinguish between missing values and true zero values. Therefore, facilities containing such

Table 3 Proportion of facility records (2011–2018) by programmatic area in the various situations based on facility records in dataset 4 (n = 42,007)

Situation	Facility records by programmatic area					
	HCT (%)	PMTC (%)	CrT (%)	VMMC (%)	PEP (%)	BS (%)
B(0XX)	2.68	6.15	1.32	2.81	18.04	1.70
C(0X0)	0.75	0.75	0.32	1.13	0.76	0.19
D(X00)	0.66	1.97	1.66	0.78	0.71	0.09
G(XXX)	92.44	81.52	42.60	0.63	21.82	0.45
H(XX0)	1.57	2.13	1.20	0.03	0.28	0.01
Duplicates	0.02	0.00	0.01	0.00	0.00	0.00
Total facility records (based on data set 4)	100.00	100.00	100.00	100.00	100.00	100.00
Total facility records removed	6.02	16.35	56.21	99.34	77.90	99.54
Total facility records retained	93.98	83.65	43.79	0.66	22.10	0.46

Situation-Detailed explanation of the various reporting situations within DHIS2 can be found in Table 1

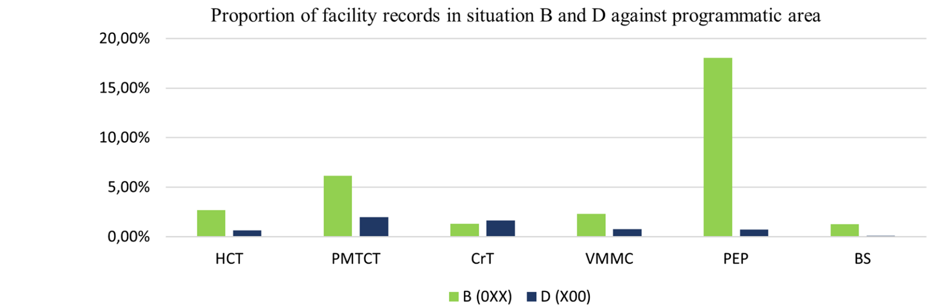

Fig. 4 Distribution of facility records based on situation B (empty reports) and situation D against programmatic area

records either are assumed to not be providing the particular service in question or are non-reporting facilities (providing services but not reporting or not expected to provide reports).

In most cases, such records are often excluded from the analyses [19, 48], as was the approach applied in this study. Furthermore, non-service providing facilities were excluded on the basis that they may provide inaccurate analyses for the evaluation question described in step1. This is on the basis that analyses may portray facilities as having good performance in facility reporting completeness and timeliness; hence give a wrong impression as no services were provided in a particular programmatic area (situation B and C). As such, even though a report was submitted on time by a facility, it will not be of benefit to a decision-maker as the report has no indicators (is empty). Nonetheless, it is worth noting that reporting facilities considered to be providing HIV services but had zero percent in timeliness were retained as these records were necessary for the subsequent analyses.

Maiga et al. posit that non-reporting facilities are often assumed not to be providing any services given that reporting rates are often ignored in analyses [13]. With this in mind, this study considered various factors prior to exclusion of non-reporting facility records. This include identifying whether there were any successful report submissions in the entire year, and whether the submitted reports contained any data in the entire year. Therefore, facilities with records that did not meet this criteria (situation A, B, and C) were considered as non-service providing in the respective programmatic areas.

Further still, another finding consistent with similar studies is that of identifying and dealing with incomplete reporting, which can be viewed from various perspectives. This can include a situation where a report for a service provided has been successfully submitted but is incomplete [17, 19, 48]; or missing reports (expected reports have not been submitted consistently for all 12 months), hence making it difficult to identify whether services were provided or not, in months were

Table 4 Results for Wilcoxon signed rank test for distribution of records in situation B

Situation B -Empty reports (0XX)			
Pairwise comparison by programmatic area	**Wilcoxon signed ranks test (*P* value)**	**Wilcoxon signed ranks test (Z value)**	**Distribution of records in situation B based on pairwise comparison by programmatic area**
PMTCT—HCT	0.012	− 2.521	Higher in PMTCT for 8 years
CrT—HCT	0.036	− 2.100	Lower in CrT for 6 years
PEP—HCT	0.012	− 2.521	Higher in PEP for 8 years
BS—HCT	0.012	− 2.524	Lower in BS for 8 years
CrT—PMTCT	0.017	− 2.521	Lower in CrT for 7 years
VMMC—PMTCT	0.012	− 2.521	Lower in VMMC for 8 years
PEP—PMTCT	0.012	− 2.521	Higher in PEP for 8 years
BS—PMTCT	0.012	− 2.524	Lower in BS for 8 years
VMMC—CrT	0.050	− 1.960	Higher in VMMC for 6 years
PEP—CrT	0.012	− 2.521	Higher in PEP for 8 years
PEP—VMMC	0.012	− 2.521	Higher in PEP for 8 years
BS—VMMC	0.012	− 2.524	Lower in BS for 8 years
BS—PEP	0.012	− 2.521	Lower in BS for 8 Years

PMTCT prevention of mother to child transmission, *HCT* HIV counselling and testing, *PEP* post-exposure prophylaxis, *BS* blood saftey, *CrT* care and treatment, *VMMC* voluntary medical male circumcision

Table 5 Results for Wilcoxon signed rank test for distribution of facility records in situation D (X00)

Situation D (X00)			
Pairwise comparison by programmatic area	**Wilcoxon signed ranks test (*P* value)**	**Wilcoxon signed ranks test (Z value)**	**Distribution of records in situation D based on pairwise comparison by programmatic area**
PMTCT—HCT	0.012	− 2.521	Higher in PMTCT for 8 years
CrT—HCT	0.012	− 2.521	Higher in CrT for 8 years
BS—HCT	0.012	− 2.524	Lower in BS for 8 years
VMMC—PMTCT	0.012	− 2.521	Lower in VMMC for 8 years
PEP—PMTCT	0.012	− 2.521	Lower in PEP for 8 years
BS—PMTCT	0.012	− 2.521	Lower in BS for 8 years
VMMC—CrT	0.012	− 2.524	Lower in VMMC for 8 years
PEP—CrT	0.012	− 2.527	Lower in PEP for 8 years
BS—CrT	0.012	− 2.524	Lower in BS for 8 years
BS—VMMC	0.018	− 2.375	Lower in BS for 8 years
BS—PEP	0.012	− 2.524	Lower in BS for 8 years

PMTCT prevention of mother to child transmission, *HCT* HIV counselling and testing, *CrT* care and treatment, *PEP* post-exposure prophylaxis, *BS* blood safety, *VMMC* Voluntary Medical Male Circumcision

reports were missing [48]. Whereas some studies retain these facility records, others opt to make adjustments for incomplete reporting. Maiga et al. posit that these adjustments need to be made in a transparent manner when creating the new data set with no modifications made on the underlying reported data [13].

In this study, all facility records were included (situation G and H) irrespective of incomplete reporting, which was similar to the approach taken by Thawer et al. [19]. On the other hand, Maina et al. opted to adjust for incomplete reporting, apart from where missing reports were considered an indication that no services were provided [48]. Furthermore, a number of studies in DHIS2 have identified duplicate records [16, 18, 19], with removal or exclusion as the common action undertaken to prepare the data set for analyses. These findings thus demonstrate duplication as a prevalent issue within DHIS2 [16, 18, 19, 49].

Whereas studies using DHIS2 data have found it necessary to clean the extracted data prior to analyses [16, 18, 19], transparent and systematic approaches are still lacking in literature [20]. Given that contexts were data is being used vary, there is no one-size fits all solution to data cleaning, considering the many existing approaches as well as the subjective component of data quality [25, 26]. As such, transparent and systematic documentation of procedures is valuable as it also increases the validity in research [21]. Moreover, existing literature advocates the need for clear and transparent description of data set creation and data cleaning methods [9, 21, 22]. Therefore, the generic five-step approach developed in this study is a step toward the right direction as it provides a systematic sequence that can be adopted for cleaning data extracted from DHIS2.

In addition, the statistical analysis employed such as non-parametric tests provide an overview of distribution of facility records containing quality issues within the various programmatic areas, hence necessitating need for further investigations where necessary. These statistics also provided a picture of the most reported programmatic areas, which contain data within their reports.

Moreover, as revealed in the screening, diagnosis and treatment phases presented in this paper, data cleaning process can be time consuming. Real-world data such as the DHIS2 data and merging of real-world data sets as shown in this paper may be noisy, inconsistent and incomplete. In the treatment stage, we present the actions taken to ensure that only meaningful data is included for subsequent analysis. Data cleaning also resulted to a smaller data set than the original as demonstrated in the results [29]. As such, the final clean data set obtained in this study is more suitable for its intended use than in its original form.

A limitation in this study was inability to determine the causality of some of the issues encountered. Whereas quality issues are in part attributed to insufficient skills or data entry errors committed at the facility level [14], some of the issues encountered from our findings (such as duplication, situation E and F) are assumed to be stemming from within the system. Nonetheless, there is need for further investigation on causality. In addition, given that situation D was identified as a result of merging two data sets extracted from DHIS2, it was expected that if reports contain indicator data, then their respective Reporting Rate and Reporting Rate on Time should be recorded. Nonetheless, it was also not possible within the confines of this study to identify the causality for situation D. As such, further investigations are also required.

In addition, there are also limitations with human augmented procedures as human is to error especially when dealing with extremely large data sets as posited by other studies [24]. Moreover, data cleaning for large data sets can also be time consuming. Nonetheless, identifying and understanding issues within the data using a human-driven approach provides better perspective prior to developing automatic procedures, which can then detect the identified issues. Therefore, there is need for developing automated procedures or tools for purposes of detecting and handling the different situation types in Table 1.

DHIS2 incorporated a quality tool, which used a similar concept as that used in calculating Cumulative Percent Completion in this study, to flag facilities with more than 10 percent zero or missing values in the annual report [12]. Based on this, we recommend that facilities with 100 percent zero or missing values also be flagged in the annual report in order to identify empty reports, as well situation where Reporting Rate on Time is zero in the annual report. Further still automated statistical procedures can be developed within the system to perform various analyses such as calculating the number of empty reports submitted by a facility for a sought period of time, per programmatic area. This could provide beneficial practical implications such as enabling decision-makers to understand the frequency of provision of certain services among the six programmatic areas within a particular period among health facilities. We also recommend for measures to be established within DHIS2 implementations to ensure that cases reported as zero appear in DHIS2.

Such findings could be used to improve the quality of reporting. Automatic procedures should also be accompanied by data visualizations, and analyses, integrated within the iterative process in order to provide insights [35]. In addition, user engagement in development of automatic procedures and actively training users in identifying and discovering various issues within the data may contribute to better quality of data [35, 37].

Conclusion

Comprehensive, transparent and systematic reporting of cleaning process is important for validity of the research studies [21]. The data cleaning included in this article was semi-automatic. It complemented the automatic procedures and resulted in improved data quality for data use in secondary analyses, which could not be secured by the automated procedures solemnly. In addition, based on our knowledge, this was the first systematic attempt to transparently report on the developed and applied data cleaning procedures for HIV-indicator data reporting in DHIS2 in Kenya. Furthermore, more robust and systematic data cleaning processes should be integrated to current inbuilt DHIS2 data quality mechanisms to ensure highest quality data.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12911-020-01315-7.

Additional file 1. Programmatic areas (reports) with respective indicators as per MOH 731- Comprehensive HIV/AIDS Facility Reporting Form extracted from DHIS2.

Additional file 2. Facility report submission data extracted from DHIS2.

Abbreviations

BS: Blood safety; CPC: Cumulative percent completion; CrT: Care and treatment; DHIS2: District Health Information System Version 2; EMR: Electronic medical record; HIV: Human immunodeficiency virus; HCT: HIV counselling and testing; KeHMS: Kenya Health Management System; KMFL: Kenya Master Facility List; LMICs: Low-and middle-income countries; MOH: Ministry of Health; NGO: Non-Governmental Organization; PEP: Post-exposure prophylaxis; PMTCT: Prevention of mother to child transmission; RHIS: Routine health information systems; RR: Reporting rate; RRT: Reporting rate on time; VMMC: Voluntary Medical Male Circumcision.

Acknowledgements

Not applicable.

Disclaimer

The findings and conclusions in this report are those of the authors and do not represent the official position of the Ministry of Health in Kenya.

Authors' contributions

MG, AB, and MW designed the study. AB and MW supervised the study. MG and AB analyzed the data. MG wrote the final manuscript. All authors discussed the results and reviewed the final manuscript. All authors read and approved the final manuscript.

Funding

This work was supported in part by the NORHED program (Norad: Project QZA-0484). The content is solely the responsibility of the authors and does not represent the official views of the Norwegian Agency for Development Cooperation.

Availability of data and materials

The data sets generated during the current study are available in the national District Health Information Software 2 online database, https://hiskenya.org/.

Ethics approval

Ethical approval for this study was obtained from the Institutional Review and Ethics Committee (IREC) Moi University/Moi Teaching and Referral Hospital (Reference: IREC/2019/78).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1] Department of Information Science and Media Studies, University of Bergen, Bergen, Norway. [2] Vanderbilt University Medical Center, Nashville, USA. [3] Department of Biomedical Engineering, Linköping University, Linköping, Sweden. [4] Institute of Biomedical Informatics, Moi University, Eldoret, Kenya.

Received: 7 April 2020 Accepted: 4 November 2020
Published online: 13 November 2020

References

1. Hotchkiss DR, Diana ML, Foreit KGF. How can routine health information systems improve health systems functioning in lowand middle-income countries? Assessing the evidence base. Adv Health Care Manag. 2012;12:25–58.
2. De Lay PR. Nicole Massoud DLR, Carae KAS and M. Strategic information for HIV programmes. In: The HIV pandemic: local and Global Implications. Oxford Scholarship Online; 2007. p. 146.
3. Beck EJ, Mays N, Whiteside A, Zuniga JM. The HIV Pandemic: Local and Global Implications. Oxford: Oxford University Press; 2009. p. 1–840.
4. Granich R, Gupta S, Hall I, Aberle-Grasse J, Hader S, Mermin J. Status and methodology of publicly available national HIV care continua and 90–90-90 targets: a systematic review. PLoS Med. 2017;14:e1002253.
5. Peersman G, Rugg D, Erkkola T, Kirwango E, Yang J. Are the investments in monitoring and evaluation systems paying off? Jaids. 2009;52(Suppl 2):8796.
6. Kariuki JM, Manders E-J, Richards J, Oluoch T, Kimanga D, Wanyee S, et al. Automating indicator data reporting from health facility EMR to a national aggregate data system in Kenya: an Interoperability field-test using OpenMRS and DHIS2. Online J Public Health Inform. 2016;8:e188.
7. Karuri J, Waiganjo P, Orwa D, Manya A. DHIS2: the tool to improve health data demand and use in Kenya. J Health Inform Dev Ctries. 2014;8:38–60.
8. Dehnavieh R, Haghdoost AA, 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. 2019;48:62–75.
9. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLOS Med. 2015;12:e1001885.
10. Dziadkowiec O, Callahan T, Ozkaynak M, Reeder B, Welton J. Using a data quality framework to clean data extracted from the electronic health record: a case study. eGEMs. 2016;4(1):11.
11. Dhis2 Documentation Team. Control data quality. DHIS2 user manual. 2020 https://docs.dhis2.org/2.31/en/user/html/dhis2_user_manual_en_full.html#control_data_quality. Accessed 10 Oct 2020.
12. Haugen JÅ, Hjemås G, Poppe O. Manual for the DHIS2 quality tool. Understanding the basics of improving data quality. 2017. https://ssb.brage.unit.no/ssb-xmlui/handle/11250/2460843. Accessed 30 Jan 2020.
13. Maïga A, Jiwani SS, Mutua MK, Porth TA, Taylor CM, Asiki G, et al. Generating statistics from health facility data: the state of routine health information systems in Eastern and Southern Africa. BMJ Global Health. 2019;4:e001849.
14. 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(Suppl 4):1–6.
15. Fan W, Geerts F. Foundations of data quality management. Synth Lect Data Manag. 2012;4:1–217.
16. Githinji S, Oyando R, Malinga J, Ejersa W, Soti D, Rono J, et al. Completeness of malaria indicator data reporting via the District Health Information Software 2 in Kenya, 2011–2015. BMC Malar J. 2017;16:1–11.
17. Wilhelm JA, Qiu M, Paina L, Colantuoni E, Mukuru M, Ssengooba F, et al. The impact of PEPFAR transition on HIV service delivery at health facilities in Uganda. PLoS ONE. 2019;14:e0223426.
18. Maina JK, Macharia PM, Ouma PO, Snow RW, Okiro EA. Coverage of routine reporting on malaria parasitological testing in Kenya, 2015–2016. Glob Health Action. 2017;10:1413266.
19. Thawer SG, Chacky F, Runge M, Reaves E, Mandike R, Lazaro S, et al. Subnational stratification of malaria risk in mainland Tanzania: a simplified assembly of survey and routine data. Malar J. 2020;19:177.
20. Shikuku DN, Muganda M, Amunga SO, Obwanda EO, Muga A, Matete T, et al. Door-to-door immunization strategy for improving access and utilization of immunization services in hard-to-reach areas: a case of Migori County, Kenya. BMC Public Health. 2019;19:1–11.
21. Van Den Broeck J, Cunningham SA, Eeckels R, Herbst K. Data cleaning: detecting, diagnosing, and editing data abnormalities. PLoS Med. 2005;2:966–70.
22. Leahey E, Entwisle B, Einaudi P. Diversity in everyday research practice: the case of data editing. Sociol Methods Res. 2003;32:64–89.
23. Wang RY, Strong DM. Beyond accuracy: what data quality means to data consumers. J Manag Inf Syst. 1996;12:5–33.

24. Langouri MA, Zheng Z, Chiang F, Golab L, Szlichta J. Contextual data cleaning. In 2018 IEEE 34th INTERNATIONAL CONFERENCE DATA ENGINEERING Work. 2018. p. 21–4.
25. Strong DM, Lee YW, Wang RY. Data quality in context. Commun ACM. 1997;40:103–10.
26. Bertossi L, Rizzolo F, Jiang L. Data quality is context dependent. In Lecture notes in business information processing. 2011. p. 52–67.
27. Bolchini C, Curino CA, Orsi G, Quintarelli E, Rossato R, Schreiber FA, et al. And what can context do for data? Commun ACM. 2009;52:136–40.
28. Chapman AD. Principles and methods of data cleaning primary species data, 1st ed. Report for the Global Biodiversity Information Facility. GBIF; 2005.
29. Zhang S, Zhang C, Yang Q. Data preparation for data mining. Appl Artif Intell. 2003;17:375–81.
30. Fayyad U, Piatetsky-Shapiro G, Smyth P. Knowledge discovery and data mining: towards a unifying framework. 1996. 31.
31. Oliveira P, Rodrigues F, Galhardas H. A taxonomy of data quality problems. In: 2nd International work data information quality. 2005. p. 219
32. Li L, Peng T, Kennedy J. A rule based taxonomy of dirty data. GSTF Int J Comput. 2011. https://doi.org/10.5176/2010-2283_1.2.52.
33. Müller H, Freytag J-C. Problems, methods, and challenges in comprehensive data cleansing challenges. Technical Report HUB-IB-164, Humboldt University, Berlin. 2003. p. 1–23.
34. Seheult AH, Green PJ, Rousseeuw PJ, Leroy AM. Robust regression and outlier detection. J R Stat Soc Ser A Stat Soc. 1989;152:133.
35. Hellerstein JM. Quantitative data cleaning for large databases. United Nations Economics Committee Europe. 2008. 42.
36. Kang H. The prevention and handling of the missing data. Korean J Anesthesiol. 2013;64:402–6.
37. Chu X, Ilyas IF, Krishnan S, Wang J. Data cleaning: overview and emerging challenges. In: Proceedings of the ACM SIGMOD international conference on management of data. New York: ACM Press; 2016. p. 2201–6.
38. Vassiliadis P, Vagena Z, Skiadopoulos S, Karayannidis N, Sellis T. Arktos: a tool for data cleaning and transformation in data warehouse environments. IEEE Data Eng Bull. 2000;23:2000.1.109.2911
39. WHO. Data Quality Review (DQR) Toolkit . WHO. World Health Organization; 2019: who.int/healthinfo/tools_data_analysis/en/. Accessed 5 Mar 2020.
40. Measure Evaluation. User Manual Routine Data Quality Assessment RDQA User Manual. 2015. https://www.measureevaluation.org/resources/tools/data-quality/rdqa-guidelines-2015. Accessed 23 Nov 2018.
41. World Health Organization. The immunization data quaity self-assessment (DQS) tool. World Health Organization. 2005 . www.who.int/vaccines-documents/. Accessed 6 Aug 2020.
42. Shanks G, Corbitt B. Understanding data quality: social and cultural aspects. In: 10th Australasian conference on information systems. 1999. p. 785–97.
43. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. J Am Med Inform Assoc. 2013;20:144–51.
44. Savik K, Fan Q, Bliss D, Harms S. Preparing a large data set for analysis: using the minimum data set to study perineal dermatitis. J Adv Nurs. 2005;52(4):399–409.
45. Miao Z, Sathyanarayanan S, Fong E, Paiva W, Delen D. An assessment and cleaning framework for electronic health records data. In: Industrial and systems engineering research conference. 2018.
46. Kulkarni DK. Interpretation and display of research results. Indian J Anaesth. 2016;60:657–61.
47. Luo W, Gallagher M, Loveday B, Ballantyne S, Connor JP, Wiles J. Detecting contaminated birthdates using generalized additive models. BMC Bioinform. 2014;12(15):1–9.
48. Maina I, Wanjal P, Soti D, Kipruto H, Droti B, Boerma T. Using health-facility data to assess subnational coverage of maternal and child health indicators, Kenya. Bull World Health Organ. 2017;95(10):683–94.
49. Bhattacharya AA, Umar N, Audu A, Allen E, Schellenberg JRM, Marchant T. Quality of routine facility data for monitoring priority maternal and newborn indicators in DHIS2: a case study from Gombe State, Nigeria. PLoS ONE. 2019;14:e0211265.

Publisher's Note

Gesicho *et al. BMC Med Inform Decis Mak (2021) 21:6*
https://doi.org/10.1186/s12911-020-01367-9

BMC Medical Informatics and Decision Making

RESEARCH ARTICLE

Open Access

Evaluating performance of health care facilities at meeting HIV-indicator reporting requirements in Kenya: an application of K-means clustering algorithm

Milka Bochere Gesicho[1,4*], Martin Chieng Were[2,4] and Ankica Babic[1,3]

Abstract

Background: The ability to report complete, accurate and timely data by HIV care providers and other entities is a key aspect in monitoring trends in HIV prevention, treatment and care, hence contributing to its eradication. In many low-middle-income-countries (LMICs), aggregate HIV data reporting is done through the District Health Information Software 2 (DHIS2). Nevertheless, despite a long-standing requirement to report HIV-indicator data to DHIS2 in LMICs, few rigorous evaluations exist to evaluate adequacy of health facility reporting at meeting completeness and timeliness requirements over time. The aim of this study is to conduct a comprehensive assessment of the reporting status for HIV-indicators, from the time of DHIS2 implementation, using Kenya as a case study.

Methods: A retrospective observational study was conducted to assess reporting performance of health facilities providing any of the HIV services in all 47 counties in Kenya between 2011 and 2018. Using data extracted from DHIS2, K-means clustering algorithm was used to identify homogeneous groups of health facilities based on their performance in meeting timeliness and completeness facility reporting requirements for each of the six programmatic areas. Average silhouette coefficient was used in measuring the quality of the selected clusters.

Results: Based on percentage average facility reporting completeness and timeliness, four homogeneous groups of facilities were identified namely: best performers, average performers, poor performers and outlier performers. Apart from blood safety reports, a distinct pattern was observed in five of the remaining reports, with the proportion of best performing facilities increasing and the proportion of poor performing facilities decreasing over time. However, between 2016 and 2018, the proportion of best performers declined in some of the programmatic areas. Over the study period, no distinct pattern or trend in proportion changes was observed among facilities in the average and outlier groups.

Conclusions: The identified clusters revealed general improvements in reporting performance in the various reporting areas over time, but with noticeable decrease in some areas between 2016 and 2018. This signifies the need for continuous performance monitoring with possible integration of machine learning and visualization approaches into national HIV reporting systems.

Keywords: K-means clustering, Completeness, Timeliness, Performance, DHIS2

*Correspondence: milcagesicho@gmail.com
[1] Department of Information Science and Media Studies, University of Bergen, Bergen, Norway
Full list of author information is available at the end of the article

Background

The Human Immunodeficiency Virus (HIV) epidemic remains a challenge globally with highest infected numbers found in countries in East and Southern Africa, which accounted for an estimated 20.7 million infected individuals in 2019 [1]. Efforts to eradicate the HIV epidemic have seen affected countries in low-middle-income-countries (LMICs) receive substantial support from donors and multilateral global organizations in order to scale-up HIV services such as antiretroviral therapy (ART), prevention of mother-to-child transmission (PMTCT) of HIV, and HIV testing and counselling (HTC) [2]. This has brought about the need to strengthen strategic information on HIV. Health Management Information Systems (HMIS), through better data quality, improves decision-making such as informing policy, measuring program effectiveness, advocacy and resource allocation [3]. Ministries of Health (MoH) and donor organizations require facilities providing HIV services to report several aggregated HIV-indicators as part of Monitoring and Evaluation (M&E) program [4, 5].

The scale-up of HIV services has contributed to strengthening of HMIS in many low-middle-income-countries, resulting in improved availability of routinely generated HIV aggregate indicator data from health facilities to the national level [6]. HIV indicator data typically comes from aggregation of monthly reports generated by various facilities that are collated in summary forms and submitted to an aggregate-level HMIS or reporting system [6]. One such national-level data aggregation system is the District Health Information Software Version 2 (DHIS2), which has been adopted by many LMICs [7].

Aggregate data stored in systems such as DHIS2 are only as good as their quality [8]. Therefore, the ability to report complete, accurate and timely data by HIV care providers and other entities is a key aspect in monitoring trends in HIV care. Various approaches to evaluating data quality have been proposed such as desk reviews, data verification or system assessments across the following data quality dimensions; completeness, timeliness, internal consistency of reported data, external comparisons and external consistency of population data [9]. Evaluations on quality of indicator reporting leveraging some of these approaches have previously been conducted within DHIS2 based on various data quality dimensions [10–14].Nonetheless, despite a long-standing requirement to report HIV indicator data to DHIS2 in LMICs, few rigorous evaluations exist to evaluate adequacy of health facility reporting at meeting completeness and timeliness requirements over time.

Rigorous reporting by facilities into DHIS2 over time is imperative to identify changes in trends and implement timely interventions [14]. In this study, we aim to leverage on machine learning algorithms as well as data visualization approaches to conduct a comprehensive assessment of the reporting performance for HIV-indicators at the national-level by facilities using completeness and timeliness indicators, with Kenya as a case study.

Methods

Related works

Table 1 illustrates some of the related studies that have extracted data from DHIS2 in order to evaluate performance at meeting the various dimensions of data quality. In addition, data from these studies was gathered from various time periods as well as various areas within health care such as malaria.

Whereas our study focused on facility reporting completeness and timeliness of HIV-indicators for the period of 2011 to 2018, the difference compared with the other studies is leveraging of the k-means clustering algorithm.

Table 1 Summary of some of the related works evaluating various dimensions of data quality

Studies	Dimensions evaluated					
	Facility reporting completeness	Indicator data completeness	Timeliness	Internal consistency	External consistency	Summary
Bhattacharya et al. [10]	X	X	X	X	X	Extracted priority maternal and neonatal health indicators Data gathered from July 2016 to June 2017
Githinji et al. [11]	X	X	–	–	–	Extracted malaria indicator data Data gathered from 2011–2015
Adokiya et al. [12]	X	–	X	–	–	Extracted disease surveillance and response reports Data gathered from 2012 and 2013
Nisingizwe et al. [14]	X	X	–	X	–	Extracted health management information systems data for selected indicators Data gathered from 2008–2012
Kiberu et al. [13]	X	–	X	–		Extracted inpatient and outpatient data Data gathered from 2011/12 and after 2012/13

Study setting
This study was conducted in Kenya, a sub-Saharan country made up of 47 counties. Administratively, the health care service delivery system has six levels, namely: community, dispensary, health center, district hospital, provincial hospital, and national referral hospital [15]. Kenya adopted the DHIS2 in 2011 at the national level for aggregation of health data across different levels of the health system [16, 17].

Study design
A retrospective observational study was conducted in order to identify reporting performance over time by health facilities in meeting completeness and timeliness reporting requirements.

Data source
Data for facilities reporting completeness and timeliness between the years 2011 and 2018 were extracted from the DHIS2 in Kenya. DHIS2 is a web-based open-source health management information system developed for purposes of collecting aggregate level data routinely generated across health facilities in various countries [7, 16]. DHIS2 also supports various activities and contains modules for processes such as data management and analytics, which contain features for data visualization, charts, pivot tables and dashboards [18]. It is also currently in use by ministries of health in over 70 countries [19]. In Kenya, DHIS2 was rolled out nationally in the year 2011 [16]. Reporting completeness and timeless data were extracted from Kenya's DHIS2 for all facilities in all the 47 counties in Kenya. Systematic procedures were used in cleaning the data using a generic five-step approach as outlined in Gesicho et al. [20]. Data used were only for facilities that offered one or more of the outlined HIV services that required reporting, namely: (1) HIV testing and counselling (HTC), (2) Prevention of Mother to Child Transmission (PMTCT), (3) Care and Treatment (CRT), (4) Voluntary Medical Male Circumcision (VMMC), (5) Post-Exposure Prophylaxis (PEP) and (6) Blood Safety (BS). These data were derived based on the MOH 731 Comprehensive HIV/AIDS facility-reporting form, which is the major monthly HIV summary report required by the MOH in Kenya and used by health facilities for reporting of HIV-indicators into DHIS2. It is worth noting that health facilities are not required to report on indicators for all the six programmatic areas, but only those for which they provide services. As such, there are variations in number of facilities (n) in the various programmatic reporting areas.

Measures
Facility reporting completeness and timeliness
Percentage completeness in facility reporting is calculated automatically within Kenya's DHIS2 and is defined as the number of actual monthly reports received divided by the expected number of reports in a given year. Percentage timeliness in facility reporting is also calculated automatically within Kenya's DHIS2 and is defined as the number of actual monthly reports received on time (by the 15th of every month) divided by the expected number of reports in a given year. Facility reporting completeness and timeliness were selected as indicators for assessing reporting performance as they were readily available within DHIS2 for the eight year period covered by the study.

Outcome measures
The primary outcome of interest consisted of identifying the performance in reporting by health facilities over time (2011–2018), with facilities put into various performance clusters and performance evaluated in the various programmatic areas.

Data analysis
K-means algorithm was preferred due to its efficiency and suitability in pattern recognition, its simplicity, ease of implementation as well as its empirical success [21]. K-means algorithm is a non-hierarchical procedure where k represents the number of clusters, which need to be specified prior to any clustering [22]. Given that K-means algorithm uses unsupervised learning, the idea was to group the health facilities into k homogeneous groups based on their performance in completeness and timeliness, in each of the six programmatic areas for each of the study years. Based on the data set and purpose of this study, we used the average silhouette coefficient, which is an intrinsic method of measuring the quality of a cluster [23]. The average value of the silhouette coefficient ranges between -1 (least preferable value indicating poor structure) and $+1$ (most preferable value indicating good structure). According to Kaufman and Rousseeuw, average silhouette measure that is greater than $+0.5$ indicates reasonable partitioning of data, whereas greater than $+0.7$ indicates a strong partitioning [24]. On the other hand average silhouette measures lower than $+0.5$ indicate a weak or artificial partitioning, whereas below $+0.2$ indicates no clusters can be exhibited from the data [24].

In order to determine the number of clusters (k) to be generated, the Euclidean distance measure was applied and k was specified within a set of values [21, 25]. The range of *k* values was then iteratively re-run with two values of k (k = 3 and k = 4) and inspecting the average corresponding silhouette values [26].

The proportion of facilities in the various cluster groups was then determined by calculating the percentage number of facilities in a particular cluster group out of the total facilities in that particular year. To illustrate the average performance of facilities within the various cluster groups, we developed a scatter chart visualization using Tableau [27]. In addition, HTC programmatic area was used as an illustrative example for the visualization, given that it is one of the most reported programmatic areas. Figures and tables were developed using Microsoft Word and Excel (Microsoft Office Version 18.2008.12711.0). All analyses were performed using SPSS [28]. A summary of the methods is illustrated in Fig. 1.

Results

Results from the silhouette coefficient average measures for each reporting area are presented in Table 2. The results ascertain that the average silhouette values for both k = 3 and k = 4 produce reasonable to strong partitioning except for 2011 under CRT where the values for k = 3 where below 0.5, hence k = 4 was used in this case. Therefore, based on method criteria and interpretability of the data set, either k = 3 and k = 4 were used where reasonable to strong partitions were identified in the average silhouette measures. As such, k = 4 was used when more variation could be provided in the data from four clusters, and k = 3 was used when three clusters provided more variation than four clusters. For VMMC and PEP programmatic areas, the number of health facilities was not enough to conduct cluster analysis in the year 2011.

The four clusters were characterized based on health facility performance as follows:

> *Best performers* This cluster consisted of health facilities that had the highest percentage in reporting completeness and timeliness in a particular reporting year.
> *Average performers* This cluster consisted of health facilities that had lower percentage in reporting completeness and timeliness compared to best performers in a particular year.
> *Poor performers* This cluster consisted of health facilities with lowest percentage in reporting completeness and timeliness in a particular year.
> *Outlier performers* This cluster consisted of health facilities with high percentage in completeness compared to average performers, but with low percentage in timeliness in that particular year.

Performance was therefore categorized per year by cluster. As such, the average percentage reporting completeness and timeliness for a particular cluster group may vary by year. It is worth noting that there were no clusters with low completeness and high timeliness as reports cannot be on time if they were not submitted in the first place. Detailed results by cluster for each reporting programmatic area are outlined below.

In Table 3 and Fig. 2, we present the segmentation of facilities based on performance cluster groups according to the HTC programmatic area. As such, Table 3 includes the average percentage for facility reporting completeness and timeliness for each cluster group in HTC for the number of facilities (n) in a particular year.

Figure 2 consists of a graphical presentation of the proportion of facilities in each cluster group per year for HTC. Based on performance trends presented in Fig. 2, the proportion of best performing facilities accounted for 72.55% in 2016, which was a progressive increase from 31.50% in 2012. Nonetheless, in 2017 and 2018 the proportion of best performing facilities accounted for 58.30% and 51.08% respectively, which was a progressive decrease from 72.55% in 2016. On the other hand, the proportion of poor performing facilities accounted for 3.40% in 2016, which was a progressive decrease from 74.93% in 2011. However, the proportion of poor performing facilities accounted for 13.49% in 2018, which was a progressive increase from 3.40% in 2016.

The proportion of average and outlier performing facilities varied in the different years with no steady trend. Nonetheless, in the latter years, the proportion of average performing facilities accounted for 20.02% in 2018, which was a progressive increase from 6.00% in 2016. On the other hand, proportion of outlier performers accounted for 15.40% in 2018, which was a decrease from 18.02% in 2017.

In Table 4 and Fig. 3, we present the segmentation of facilities based on performance cluster groups according to the PMTCT programmatic area. As such, Table 4 includes the average percentage for facility reporting completeness and timeliness for each cluster group in PMTCT for the number of facilities (n) in a particular year.

Figure 3 consists of a graphical presentation of the proportion of facilities in each cluster group per year for PMTCT. Based on performance trends presented in Fig. 3, the proportion of best performing facilities accounted for 74.01% in 2015, which was a progressive increase from 18.80% in 2011. Nonetheless, in 2018 the proportion of best performing facilities accounted for 47.15%, which was a progressive decrease from 74.01% in 2015. On the other hand, the proportion of poor

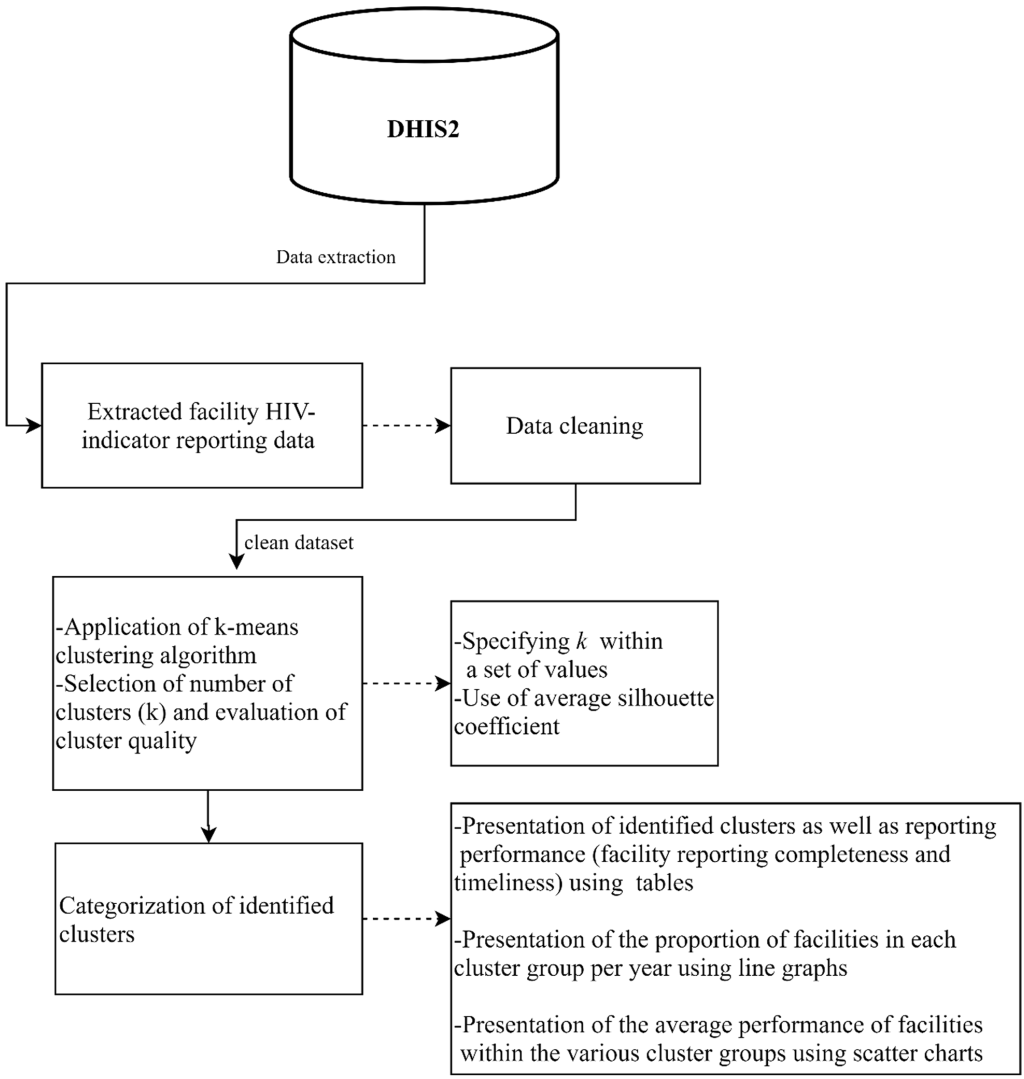

Fig. 1 Summary of methods

performing facilities accounted for 3.66% in 2015, which was a progressive decrease from 77.07% in 2011. However, in 2018 the proportion of poor performing facilities accounted for 14.61%, which was a progressive increase from 3.66% in 2015.

The proportion of average and outlier performing facilities varied in the different years with no steady trend. Nonetheless, for the latter years, proportion of average performing facilities accounted for 20.34% in 2018, which was an increase from 17.19% in 2017. On the other hand, proportion of outlier performers accounted for 17.90% in 2018, which was an increase from 3.65% in 2016.

In Table 5 and Fig. 4, we present the segmentation of facilities based on performance cluster groups according

Table 2 Average of the Silhouette of a k-means clustering when k = 3 and k = 4

Average silhouette measures								
HTC			**PMTCT**			**CRT**		
Year	**K = 3**	**K = 4**	**Year**	**K = 3**	**K = 4**	**Year**	**K = 3**	**K = 4**
2011	0.800	0.775	2011	0.674	0.706	2011	**0.368**	0.582
2012	0.526	0.563	2012	0.585	0.588	2012	0.556	0.599
2013	0.659	0.648	2013	0.654	0.632	2013	0.637	0.618
2014	0.669	0.669	2014	0.676	0.666	2014	0.692	0.663
2015	0.737	0.709	2015	0.649	0.711	2015	0.710	0.705
2016	0.749	0.754	2016	0.791	0.774	2016	0.708	0.710
2017	0.685	0.673	2017	0.699	0.677	2017	0.696	0.700
2018	0.593	0.714	2018	0.689	0.707	2018	0.654	0.701
VMMC			**PEP**			**BS**		
Year	**K = 3**	**K = 4**	**Year**	**K = 3**	**K = 4**	**Year**	**K = 3**	**K = 4**
2011	[a]	[a]	2011	0.704	0.679	2011	[a]	[a]
2012	1.00	[b]	2012	0.593	0.605	2012	0.734	0.730
2013	0.64	0.669	2013	0.639	0.629	2013	0.732	0.687
2014	0.634	0.661	2014	0.675	0.667	2014	0.712	0.650
2015	0.733	0.681	2015	0.682	0.673	2015	0.617	0.641
2016	0.708	0.699	2016	0.696	0.665	2016	0.719	0.680
2017	0.765	0.733	2017	0.621	0.611	2017	0.577	0.637
2018	0.657	0.636	2018	0.650	0.673	2018	0.610	0.607

[a] There are not enough valid cases to conduct the specified cluster analysis

[b] In the data, there is insufficient variation to honor the four clusters specified. The number of clusters is reduced to 3

to the CRT programmatic area. As such, Table 5 includes the average percentage for facility reporting completeness and timeliness for each cluster group in CRT for the number of facilities (n) in a particular year.

Figure 4 consists of a graphical presentation of the proportion of facilities in each cluster group per year for CRT. Based on performance trends presented in Fig. 4, the proportion of best performing facilities accounted for 75.49% in 2016, which was a progressive increase from 5.65% in 2011. Nonetheless, in 2018 the proportion of best performing facilities accounted for 53.24%, which was a progressive decrease from 75.49% in 2016. On the other hand, the proportion of poor performing facilities accounted for 2.99% in 2016, which was a progressive decrease from 71.75% in 2011. However, in 2018 the proportion of poor performing facilities accounted for 17.47%, which was a progressive increase from 2.99% in 2016.

The proportion of average and outlier performing facilities varied in the different years with no steady trend. Nonetheless, for the latter years the proportion of average performing facilities accounted for 24.81% in 2018, which was an increase from 7.06% in 2016. On the other hand, proportion of outlier performers accounted for 4.48% in 2018, which was a progressive decrease from 14.46% in 2016.

In Table 6 and Fig. 5, we present the segmentation of facilities based on performance cluster groups according to the VMMC programmatic area. As such, Table 6 includes the average percentage for facility reporting completeness and timeliness for each cluster group in VMMC for the number of facilities (n) in a particular year.

Figure 5 consists of a graphical presentation of the proportion of facilities in each cluster group per year for VMMC. Based on performance trends presented in Fig. 5, the proportion of best performing facilities accounted for 54.35% in 2016, which was a progressive increase from 8.70% in 2013. Nonetheless, in 2018 the proportion of best performing facilities accounted for 17.31%, which was a progressive decrease from 54.35% in 2016. On the other hand, the proportion of poor performing facilities accounted for 13.04% in 2016, which was a progressive decrease from 39.13%% in 2013. However, in 2017 and 2018 the proportion of poor performing facilities accounted for 21.88% and 21.15%, which was a progressive increase from 13.04% in 2016.

The proportion of average and outlier performing facilities varied in the different years with no steady trend.

Table 3 HIV testing and counselling (HTC)-health facility (n) segmentation based on performance clusters

Year	2011				2012			
Cluster group	**Best n = 0**	**Average n = 177**	**Poor n = 556**	**Outlier n = 9**	**Best n = 1206**	**Average n = 1301**	**Poor n = 794**	**Outlier n = 528**
MOH 731-1 HTC completeness	0.00	24.49	13.07	91.67	90.08	55.30	25.68	86.75
MOH 731-1 HTC timeliness	0.00	16.63	2.91	21.30	80.47	45.65	16.11	46.17
Year	**2013**				**2014**			
Cluster group	**Best n = 3219**	**Average n = 806**	**Poor n = 437**	**Outlier n = 427**	**Best n = 3837**	**Average n = 568**	**Poor n = 297**	**Outlier n = 615**
MOH 731-1 HTC completeness	96.73	68.77	32.96	89.86	98.18	73.07	33.75	95.94
MOH 731-1 HTC timeliness	89.55	57.33	21.63	43.00	92.96	62.42	23.02	54.06
Year	**2015**				**2016**			
Cluster group	**Best n = 3916**	**Average n = 1172**	**Poor n = 296**	**Outlier n = 282**	**Best n = 4376**	**Average n = 362**	**Poor n = 205**	**Outlier n = 1089**
MOH 731-1 HTC completeness	99.40	88.30	34.57	93.09	99.34	69.15	31.47	91.29
MOH 731-1 HTC timeliness	96.33	71.71	27.45	33.45	95.89	51.07	20.29	74.04
Year	**2017**				**2018**			
Cluster group	**Best n = 3698**	**Average n = 1164**	**Poor n = 338**	**Outlier n = 1143**	**Best n = 3403**	**Average n = 1334**	**Poor n = 899**	**Outlier n = 1026**
MOH 731-1 HTC completeness	97.98	64.47	32.69	94.20	88.48	52.68	26.87	77.35
MOH 731-1 HTC timeliness	93.92	57.04	23.59	64.33	86.93	48.84	22.98	64.65

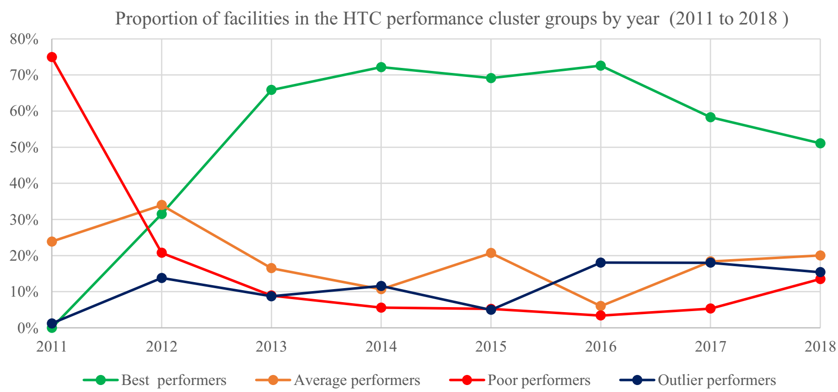

Fig. 2 HTC performance trend based on proportion of facilities by year

Table 4 Prevention of Mother to Child Transmission (PMTCT)—health facility (n) segmentation based on performance clusters

Year	2011				2012			
Cluster group	Best n = 132	Average n = 20	Poor n = 541	Outlier n = 9	Best n = 1052	Average n = 1230	Poor n = 782	Outlier n = 508
MOH 731-2 PMTCT completeness	21.67	38.32	12.91	91.67	90.03	55.51	26.09	85.65
MOH 731-2 PMTCT timeliness	18.64	4.58	2.81	18.52	80.87	45.55	16.20	47.33
Year	2013				2014			
Cluster group	Best n = 2277	Average n = 1188	Poor n = 527	Outlier n = 444	Best n = 2737	Average n = 1210	Poor n = 277	Outlier n = 586
MOH 731-2 PMTCT completeness	97.73	84.02	37.19	85.98	98.61	89.43	37.03	96.26
MOH 731-2 PMTCT timeliness	92.11	63.53	26.11	29.70	92.31	59.29	24.02	14.54
Year	2015				2016			
Cluster group	Best n = 3785	Average n = 517	Poor n = 187	Outlier n = 625	Best n = 2732	Average n = 1156	Poor n = 237	Outlier n = 194
MOH 731-2 PMTCT completeness	98.84	75.61	30.34	98.13	99.43	90.03	37.95	89.32
MOH 731-2 PMTCT timeliness	91.22	61.72	21.97	38.76	95.42	72.36	25.98	38.46
Year	2017				2018			
Cluster group	Best n = 3456	Average n = 944	Poor n = 348	Outlier n = 744	Best n = 2685	Average n = 1259	Poor n = 832	Outlier n = 1018
MOH 731-2 PMTCT completeness	97.58	64.96	38.51	93.73	88.48	53.03	27.69	79.02
MOH 731-2 PMTCT timeliness	91.54	58.59	26.55	54.52	86.72	48.22	22.65	63.20

Nonetheless, for the latter years, the proportion of average performing facilities accounted for 25.00% in 2018, which was an increase from 15.63% in 2017. On the other hand, proportion of outlier performers accounted for 36.54% in 2018, which was a progressive increase from 10.87% in 2016.

In Table 7 and Fig. 6, we present the segmentation of facilities based on performance cluster groups according to the PEP programmatic area. As such, Table 7 includes the average percentage for facility reporting completeness and timeliness for each cluster group in PEP for the number of facilities (n) in a particular year.

Figure 6 consists of a graphical presentation of the proportion of facilities in each cluster group per year for PEP. Based on performance trends presented in Fig. 6, the proportion of best performing facilities accounted for 66.76% in 2015, which was a progressive increase from 2.99% in 2011. Nonetheless, in 2018 the proportion of best performing facilities accounted for 51.24%, which was a decrease from 66.01% in 2017. On the other hand, the proportion of poor performing facilities accounted for 3.91% in 2016, which was a progressive decrease from 17.76% in 2013. However, in 2018 the proportion of poor performing facilities accounted for 18.59%, which was a progressive increase from 3.91% in 2016.

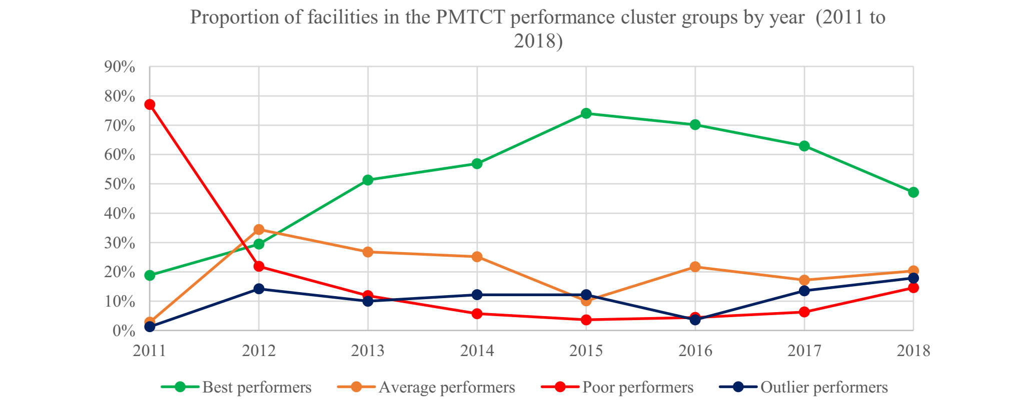

Fig. 3 PMTCT performance trend based on proportion of facilities by year

The proportion of average and outlier performing facilities varied in the different years with no steady trend. Nonetheless, for the latter years the proportion of average performing facilities accounted for 28.76% in 2018, which was an increase from 17.09% in 2017. On the other hand, proportion of outlier performers accounted for 1.41% in 2018, which was a progressive decrease from 24.78% in 2016.

In Table 8 and Fig. 7, we present the segmentation of facilities based on performance cluster groups according to the BS programmatic area. As such, Table 8 includes the average percentage for facility reporting completeness and timeliness for each cluster group in BS for the number of facilities (n) in a particular year.

Figure 7 consists of a graphical presentation of the proportion of facilities in each cluster group per year for BS. Based on performance trends presented in Fig. 7, the proportion of best performing facilities accounted for 26.67% in 2015 and 2016, which was a decrease from 33.33% in 2014. Nonetheless, in 2018 the proportion of best performing facilities accounted for 15.38%, which was a decrease from 32.00% in 2017. On the other hand, the proportion of poor performing facilities accounted for 20.00% in 2015 and 2016, which was a progressive decrease from 43.48% in 2011. However, in 2017 the proportion of poor performing facilities accounted for 24.00%, which was an increase from 2016. For the latter years, the proportion of average performing facilities accounted for 28.00% in 2017 and 38.46% in 2018. On the other hand, proportion of outlier performers accounted for 16.00% in 2017 and 23.08% 2018. Nonetheless, there have been a general progressive decrease in facilities submitting BS indicators from 2013 to 2018.

Scatter chart visualization of HTC performance clusters

In this section, we present an interactive visual representation of performance cluster groups using scatter charts. As an illustrative example using performance reporting of the HTC programmatic area, Fig. 8 demonstrates the visualization of the average performance of facilities by county for the period 2011 to 2018. Each of the four performance cluster groups are represented using a similar color approach in Figs. 2, 3, 4, 5, 6 and 7. Each point contains the following attributes: name of county, number of facilities represented in that county, and the average completeness and timeliness for the facilities, which are displayed upon hovering the mouse on a point. For example, a green point may represent the average completeness and timeliness for the number of facilities in Nairobi county, which were in the best performing cluster in a particular year. This scenario is replicated for other counties and performance clusters. It is worth noting that facilities represented in each point are of varying characteristics such as type (hospital, health center), and ownership (private, public), hence are clustered based on performance. As such, the points in the scatter chart visualization provide a clear illustration of the four performance cluster groups and their behavior over time. For instance, the initial year of reporting shows only few clusters. Nonetheless, as reporting increases with time, more clusters develop.

Moreover, the outlier performance cluster has shown some improvement in performance as demonstrated with the left movement in the chart over time. The best performing cluster (green) also demonstrates a similar observation with the most improvement in 2016. The illustration in Fig. 2 further shows the proportion of best performing facilities being higher in 2016.

Table 5 Care and Treatment (CRT)—health facility (n) segmentation based on performance clusters

Year	2011				2012			
Cluster group	Best n = 20	Average n = 76	Poor n = 254	Outlier n = 4	Best n = 634	Average n = 662	Poor n = 430	Outlier n = 98
MOH 731-3 care and treatment completeness	42.50	21.61	12.49	93.75	90.00	57.90	24.69	84.61
MOH 731-3 care and treatment timeliness	2.09	17.79	2.70	22.93	76.54	46.74	15.29	22.81
Year	**2013**				**2014**			
Cluster group	**Best n = 1063**	**Average n = 587**	**Poor n = 217**	**Outlier n = 219**	**Best n = 1407**	**Average n = 554**	**Poor n = 204**	**Outlier n = 236**
MOH 731-3 care and treatment completeness	97.67	81.29	31.24	90.05	98.67	87.86	34.51	94.53
MOH 731-3 care and treatment timeliness	90.81	59.82	19.79	24.03	92.01	62.55	27.03	24.65
Year	**2015**				**2016**			
Cluster group	**Best n = 1647**	**Average n = 607**	**Poor n = 132**	**Outlier n = 227**	**Best n = 2171**	**Average n = 203**	**Poor n = 86**	**Outlier n = 416**
MOH 731-3 care and treatment completeness	99.00	93.63	35.73	93.96	99.09	76.65	27.22	97.21
MOH 731-3 care and treatment timeliness	94.71	66.06	23.13	25.27	91.13	59.43	16.15	38.87
Year	**2017**				**2018**			
Cluster group	**Best n = 1837**	**Average n = 750**	**Poor n = 264**	**Outlier n = 241**	**Best n = 1676**	**Average n = 781**	**Poor n = 550**	**Outlier n = 141**
MOH 731-3 care and treatment completeness	98.82	92.41	43.10	95.74	86.94	55.13	26.68	71.65
MOH 731-3 care and treatment timeliness	94.22	65.41	32.70	27.61	81.75	50.71	23.26	21.37

Further still, the average facility reporting completeness and timeliness among the average performance cluster group (orange), seemed to have improved in 2015 compared with previous and subsequent years, based on the upward shift in the chart.

Discussion

The results of our study demonstrate how k-means clustering and interactive cluster-based visualization can be used in identifying patterns and categories within national-level HIV reporting systems, uncovering previously unrecognized patterns. The four categories

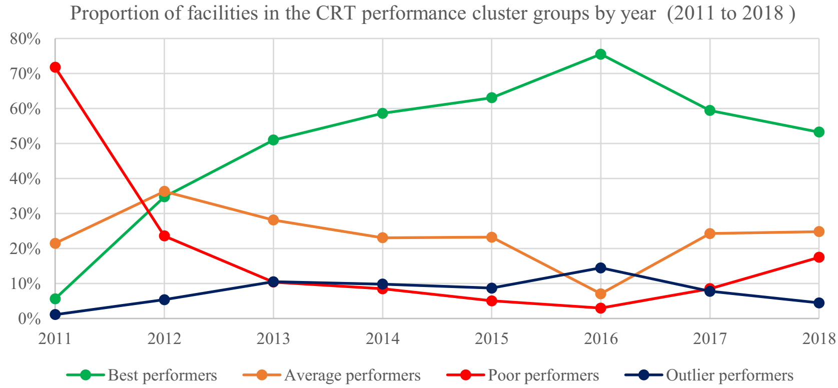

Fig. 4 CRT performance trend based on proportion of facilities by year

Table 6 Voluntary Medical Male Circumcision (VMMC)-health facility (n) segmentation based on performance clusters

Year	2012				2013			
Cluster group	Best n = 0	Average n = 2	Poor n = 2	Outlier n = 2	Best n = 2	Average n = 7	Poor n = 4	Outlier n = 5
MOH 731-4 VMMC completeness	0.00	17.00	8.00	8.00	54.50	35.57	13.89	51.80
MOH 731-4 VMMC timeliness	0.00	17.00	8.00	0.00	50.00	19.00	7.33	23.40
Year	**2014**				**2015**			
Cluster group	**Best n = 7**	**Average n = 14**	**Poor n = 16**	**Outlier n = 5**	**Best n = 15**	**Average n = 7**	**Poor n = 7**	**Outlier n = 15**
MOH 731-4 VMMC completeness	85.86	51.14	20.38	81.80	95.07	50.00	15.57	86.67
MOH 731-4 VMMC timeliness	81.14	39.36	13.00	36.60	88.38	42.86	14.43	62.20
Year	**2016**				**2017**			
Cluster group	**Best n = 25**	**Average n = 10**	**Poor n = 7**	**Outlier n = 4**	**Best n = 28**	**Average n = 10**	**Poor n = 14**	**Outlier n = 12**
MOH 731-4 VMMC completeness	97.12	67.60	17.86	70.75	92.61	52.40	17.79	86.83
MOH 731-4 VMMC timeliness	90.00	62.60	13.14	16.75	86.88	37.40	10.57	58.31
Year	**2018**							
Cluster group	**Best n = 9**	**Average n = 13**	**Poor n = 11**	**Outlier n = 19**				
MOH 731-4 VMMC completeness	85.73	43.94	19.09	61.58				
MOH 731-4 VMMC timeliness	81.11	36.15	16.36	55.26				

identified (best performers, average performers, poor performers, and outlier performers) reveal the variation in reporting performance among facilities with respect to year and programmatic area. Moreover, apart from the BS programmatic area, a distinct pattern observed in five of the other programmatic areas was that as the proportion of best performing facilities increased, the proportion of poor performing facilities decreased. In addition, the proportion of facilities in the best performing cluster was higher over time, compared to the proportion of

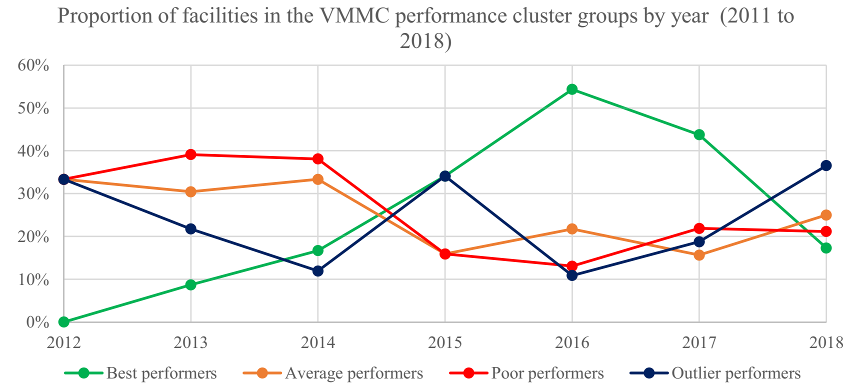

Fig. 5 VMMC performance trend based on proportion of facilities by year

facilities in the other performance clusters. These observations denote improvements in reporting over time within Kenya.

Factors that could explain these improvements in part include data quality improvement procedures done through progressive trainings of those collecting primary data and of health records information officers, provision of technical reporting support to facilities [16]. Other factors such as automation of indicator reporting by electronic medical records (EMRs) to the DHIS2, have the potential to improve routine reporting based on evidence from feasibility studies conducted [29]. With future prospects on automating indicator data reporting, cohort studies can be conducted to establish their impact based on facility reporting completeness and timeliness performance in DHIS2. Further, concerted efforts in improving routine performance of HMIS, touching on technical, behavioral and organizational domains can improve reporting in Kenya [30].

However, despite the observed improvements in performance, there was a decline in proportion of best performing facilities in different years (between 2016 and 2018), depending on the programmatic area. It is worth noting that Kenya experienced one of the longest health worker strike in the public-sector from 5 December 2016 to November 2017, lasting a total of 250 days [31]. The first phase (5 December to 14 March 2017), involved a doctors strike lasting 100 days [31]. Whereas the second phase (5 June to 1 November 2017) involved a nurses strike lasting 150 days [31]. As such, although there may have been other factors that contributed to the decline in proportion of best performing facilities, we suspect that these strikes might have also affected the reporting process. In addition, the decline in 2018 may be attributed to the introduction of new MOH731 summary reporting tools revised in 2018. As such, some facilities were still using the old tool while others had already began using the new tool, signifying the need to improve approaches during transition of reported data.

In overall, we observed that average percentage timeliness tended to be lower compared to average percentage completeness in all the four performance groups. This observation is reflected in other similar studies [12, 32]. Nonetheless, as much as this observation was common among the four performance groups, the outlier performance group specifically brings to light larger disparities between average completeness and timeliness. For instance, as presented in Table 3 for the year 2011, we see that average completeness is 91.67% and timeliness 21.30%. Similar observations can be made for subsequent tables in the various programmatic areas.

Given that timeliness plays an important role in decision-making, there is a cause for concern when there is good effort in submitting of reports, with limitations on timeliness especially in the outlier performance group. As such, there is need for qualitative enquiries to investigate the large disparities in average percentage completeness and timeliness. This is because various factors could act as barriers or facilitators to health facilities ability to attaining and maintaining good completeness and timeliness reporting performance. These factors could be targeted by ministries of health in developing strategies to improve reporting performance of health facilities.

A limitation observed in the scatter chart was that the data points become densely packed in cases where they are many in a small area, hence making it difficult to identify the various points within a cluster. An example is best performers (Fig. 8), more so in 2016. Nonetheless,

Table 7 Post-Exposure Prophylaxis (PEP)-health facility (n) segmentation based on performance clusters

Year	2011				2012			
Cluster groups	**Best n = 0**	**Average n = 2**	**Poor n = 63**	**Outlier n = 2**	**Best n = 173**	**Average n = 256**	**Poor n = 328**	**Outlier n = 34**
MOH 731-5 post-exposure prophylaxis completeness	0.00	54.20	13.48	95.85	84.98	54.24	23.56	89.71
MOH 731-5 post-exposure prophylaxis timeliness	0.00	4.15	6.18	8.35	73.74	44.72	15.65	34.07
Year	**2013**				**2014**			
Cluster groups	**Best n = 583**	**Average n = 281**	**Poor n = 205**	**Outlier n = 85**	**Best n = 677**	**Average n = 221**	**Poor n = 124**	**Outlier n = 281**
MOH 731-5 post-exposure prophylaxis completeness	94.44	61.18	29.01	87.45	97.04	56.66	23.39	83.53
MOH 731-5 post-exposure prophylaxis timeliness	88.01	51.75	20.00	41.74	93.14	40.20	17.24	63.91
Year	**2015**				**2016**			
Cluster groups	**Best n = 954**	**Average n = 305**	**Poor n = 103**	**Outlier n = 67**	**Best n = 953**	**Average n = 161**	**Poor n = 61**	**Outlier n = 387**
MOH 731-5 post-exposure prophylaxis completeness	97.14	76.33	27.25	78.37	98.15	59.58	27.85	83.22
MOH 731-5 post-exposure prophylaxis timeliness	93.05	62.86	22.34	29.24	95.37	46.37	22.83	70.99
Year	**2017**				**2018**			
Cluster groups	**Best n = 1031**	**Average n = 267**	**Poor n = 137**	**Outlier n = 127**	**Best n = 725**	**Average n = 407**	**Poor n = 263**	**Outlier n = 20**
MOH 731-5 post-exposure prophylaxis completeness	95.73	66.51	38.29	90.02	85.04	54.06	24.07	80.50
MOH 731-5 post-exposure prophylaxis timeliness	91.35	59.21	28.23	54.50	82.38	49.99	20.32	36.46

interactive components (mouse hovering and filtering) incorporated within the scatter chart facilitate access to detailed information. As such, this allows for closer examination of various elements within the data set such as performance in individual counties and number of facilities within a county for a particular performance cluster. This also enables identifying areas that warrant further investigation in their performance, which contributes to informed decision-making. The interactive approach was also used based on the need to visualize various facets of data simultaneously, which can be a challenge [33].

Incorporation of these analyses as well as visualizations to run in real time within aggregate-level HMIS, have the potential to allow monitoring and timely responsiveness to performance changes. Moreover, off shelf software

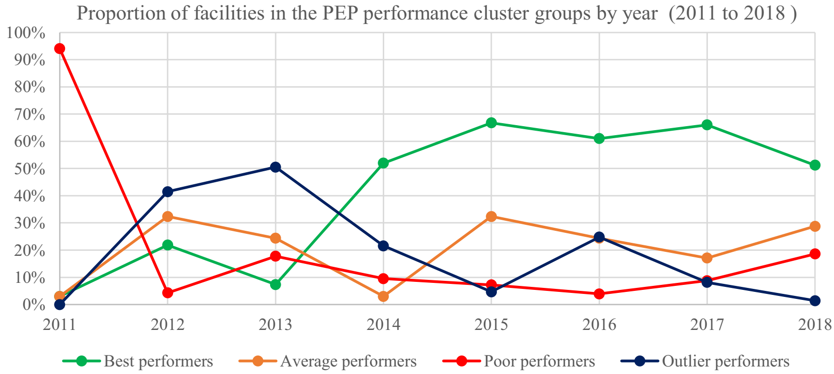

Fig. 6 PEP performance trend based on proportion of facilities by year

Table 8 Blood safety (BS)—health facility segmentation based on performance clusters

Year	2012				2013			
Cluster groups	Best n = 3	Average n = 8	Poor n = 10	Outlier n = 2	Best n = 8	Average n = 8	Poor n = 11	Outlier n = 12
MOH 731-6 blood safety completeness	69.67	43.75	18.30	100.00	94.88	37.50	15.82	75.75
MOH 731-6 blood safety timeliness	67.00	35.25	14.23	54.00	91.50	30.13	8.18	57.00
Year	**2014**				**2015**			
Cluster groups	**Best n = 11**	**Average n = 10**	**Poor n = 9**	**Outlier n = 3**	**Best n = 8**	**Average n = 14**	**Poor n = 6**	**Outlier n = 2**
MOH 731-6 blood safety completeness	95.55	67.60	47.33	97.33	87.38	62.43	22.17	58.00
MOH 731-6 blood safety timeliness	87.95	62.50	40.33	22.33	81.25	45.86	15.17	8.50
Year	**2016**				**2017**			
Cluster groups	**Best n = 8**	**Average n = 9**	**Poor n = 6**	**Outlier n = 7**	**Best n = 8**	**Average n = 7**	**Poor n = 6**	**Outlier n = 4**
MOH 731-6 blood safety completeness	94.88	69.56	47.33	27.14	83.25	56.00	26.33	79.25
MOH 731-6 blood safety timeliness	92.79	62.00	40.33	17.86	78.13	54.86	22.33	41.50
Year	**2018**							
Cluster groups	**Best n = 2**	**Average n = 5**	**Poor n = 3**	**Outlier n = 3**				
MOH 731-6 blood safety completeness	85.00	54.00	26.67	66.67				
MOH 731-6 blood safety timeliness	75.00	34.00	26.67	53.33				

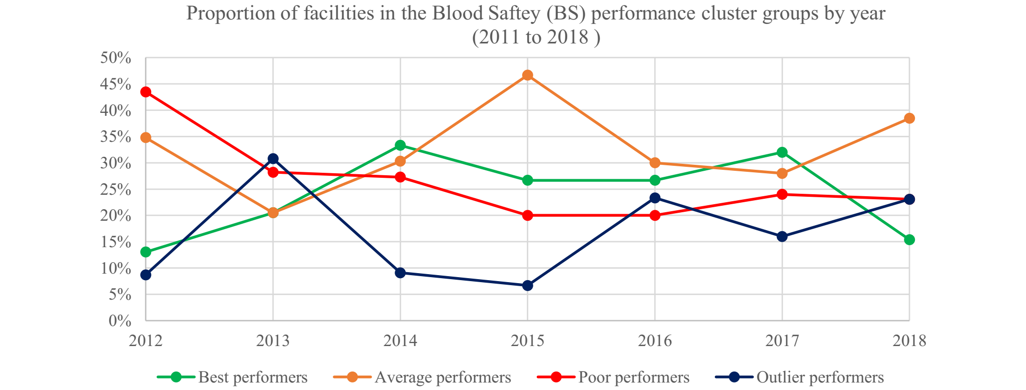

Fig. 7 BS performance trend based on proportion of facilities by year

such as Tableau [27], which provide basic modules for free usage can be leveraged as a cost effective alternative for representing and sharing analysis for routinely collected data that has been extracted from large data systems.

The scope of the study can be relevant for many countries dealing with HIV reporting in aggregate-level HMIS. However, the limitation in this study is that data have been collected and analyzed for one country only. Nonetheless, the indicators used (completeness and timeliness) could also be relevant in other contexts. Further, the findings only reflect trends and associations, and do not explain causality. Investigations, including use of qualitative approaches, are needed to definitively determine causes of the observed trends and variations. While we only looked at clustering based on performance, we recognize that performance can be associated with several other factors including facility ownership (private vs public), facility type and level, (for example hospital, dispensary), presence or absence of electronic reporting systems, geographical location and infrastructure availability, among others.

One of the future aims will be to determine factors influencing movement of facilities between clusters with special attention to factors associated with decrease in performance.

Conclusions

K-means clustering and interactive cluster-based visualization was applied to identify patterns of performance in terms of completeness and timeliness of facility reporting in six HIV programmatic areas. This resulted to four clusters: best performers, average performers, poor performers, and outlier performers, depending on average percentage of completeness and timeliness. The identified clusters revealed general improvements in reporting performance in the various reporting areas over time, but with most noticeable decrease in some programmatic areas between 2016 and 2018. This signifies the need for continuous performance monitoring with possible integration of machine learning and visualization approaches into national HIV reporting systems.

As future work, we will also work with the relevant decision-makers in the study country to incorporate the demonstrated machine learning and visualization approaches for use in automatic and continuous assessment of reporting performance within Kenya.

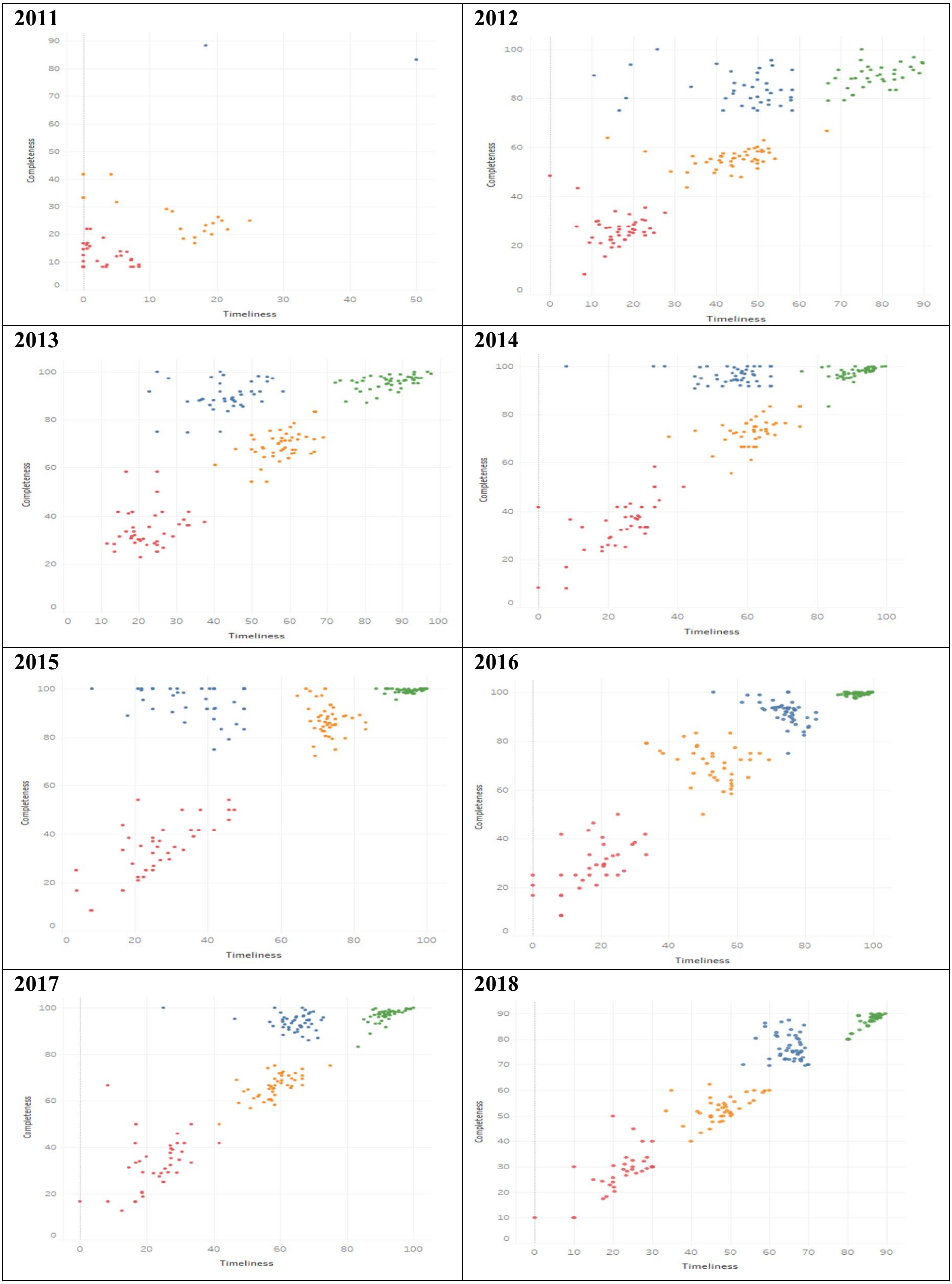

Fig. 8 Cluster visualization of facility performance by county illustration for HIV Testing and Counselling

Abbreviations

ART: Antiretroviral therapy; BS: Blood Safety; CRT: Care and Treatment; DHIS2: District Health Information System Version 2; EMRs: Electronic Medical Record System; HTC: HIV Testing and Counselling; HIV: Human Immunodeficiency Virus; HMIS: Health Information Management Systems; LMICs: Low-middle-income countries; M&E: Monitoring and Evaluation; MoH: Ministry of Health; PEP: Post-Exposure Prophylaxis; PMTCT: Prevention of Mother to Child Transmission; VMMC: Voluntary Medical Male Circumcision.

Acknowledgements

Not applicable.

Authors' contributions

MG, AB, and MW designed the study. AB and MW supervised the study. MG and AB analyzed the data. All authors discussed the results, reviewed, and approved the final manuscript. MG wrote the final manuscript. All authors read and approved the final manuscript.

Funding

This work was supported in part by the NORHED program (Norad: Project QZA-0484). The content is solely the responsibility of the authors and does not represent the official views of the Norwegian Agency for Development Cooperation.

Availability of data and materials

The data sets generated during the current study are available in the national District Health Information Software 2 online database, https://hiskenya.org/.

Ethics approval and consent to participate

Ethical approval for this study was obtained from the Institutional Review and Ethics Committee (IREC) Moi University/Moi Teaching and Referral Hospital (Reference: IREC/2019/78).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Disclaimer

The findings and conclusions in this report are those of the authors and do not represent the official position of the Ministry of Health in Kenya.

Author details

[1] Department of Information Science and Media Studies, University of Bergen, Bergen, Norway. [2] Vanderbilt University Medical Center, Nashville, USA. [3] Department of Biomedical Engineering, Linköping University, Linköping, Sweden. [4] Institute of Biomedical Informatics, Moi University, Eldoret, Kenya.

Received: 5 August 2020 Accepted: 8 December 2020
Published online: 06 January 2021

References

1. Global HIV and AIDS statistics—2020 fact sheet | UNAIDS. https://www.unaids.org/en/resources/fact-sheet. Accessed 14 July 2020.
2. UNAIDS. Towards universal access. In: UNAIDS Annual Report. 2009. http://www.unaids.org/en/KnowledgeCentre/Resources/Publications/default.asp. Accessed 14 July 2020.
3. Mbondo M, Scherer J, Aluoch GO, Sundsmo A, Mwaura N. Organizational HIV monitoring and evaluation capacity rapid needs assessment: the case of Kenya. Pan Afr Med J. 2013;14:1–7.
4. Porter LE, Bouey PD, Curtis S, Hochgesang M, Idele P, Jefferson B, et al. Beyond indicators. JAIDS J Acquir Immune Defic Syndr. 2012;60:S120–6.
5. Ekouevi DK, Karcher S, Coffie PA. Strengthening health systems through HIV monitoring and evaluation in Sub-Saharan Africa. Curr Opin HIV AIDS. 2011;6:245–50.
6. Saito S, Howard AA, Chege D, Ellman TM, Ahoua L, Elul B, et al. Monitoring quality at scale: implementing quality assurance in a diverse, multicountry HIV program. AIDS. 2015;29:S129–36.
7. Dehnavieh R, Haghdoost AA, 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. 2019;48:62–75.
8. Manya A, Nielsen P. Reporting practices and data quality in health information systems in developing countries: an exploratory case study in Kenya. J Health Inform Dev Ctries. 2016;10:114–26.
9. WHO. Data Quality Review (DQR) Toolkit. WHO. In :World Health Organization; 2019. http://who.int/healthinfo/tools_data_analysis/en/. Accessed 5 Mar 2020.
10. Bhattacharya AA, Umar N, Audu A, Allen E, Schellenberg JRM, Marchant T. Quality of routine facility data for monitoring priority maternal and newborn indicators in DHIS2: a case study from Gombe State, Nigeria. PLoS ONE J. 2019;14:e0211265.
11. Githinji S, Oyando R, Malinga J, Ejersa W, Soti D, Rono J, et al. Completeness of malaria indicator data reporting via the District Health Information Software 2 in Kenya, 2011–2015. BMC Malar J. 2017;16:1–11.
12. Adokiya MN, Awoonor-Williams JK, Beiersmann C, Müller O. Evaluation of the reporting completeness and timeliness of the integrated disease surveillance and response system in northern Ghana. Ghana Med J. 2016;50:3–8.
13. Kiberu VM, Matovu JK, Makumbi F, Kyozira C, Mukooyo E, Wanyenze RK. Strengthening district-based health reporting through the district health management information software system: the Ugandan experience. BMC Med Inform Decis Mak. 2014;14:40.
14. 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.
15. Muga R, Kizito P, Mbayah MM, Gakuruh T. Overview of the health system in Kenya. In: Kenya Service provision assessment survey 2004; 1999, p. 13–24.
16. Manya A, Braa J, Øverland L, Titlestad O, Mumo J, Nzioka C. National roll out of District Health Information Software (DHIS 2) in Kenya, 2011–central server and cloud based infrastructure. IST-Africa. 2012;2012:1–9.
17. Karuri J, Waiganjo P, Orwa D, Manya A. DHIS2: the tool to improve health data demand and use in Kenya. J Health Inform Dev Ctries. 2014;8:38–60.
18. DHIS2: DHIS2 overview. https://www.dhis2.org/overview. Accessed 28 Sep 2020.
19. DHIS2: SDHIS2 in action. https://www.dhis2.org/in-action. Accessed 28 Sep 2020.
20. Gesicho MB, Were MC, Babic A. Data cleaning process for HIV-indicator data extracted from DHIS2 national reporting system: a case study of Kenya. BMC Med Inform Decis Mak. 2020;20:293.
21. Jain AK. Data clustering: 50 years beyond K-means. Pattern Recognit Lett. 2010;31:651–66.
22. Everitt BS, Landau S, Leese M, Stahl D. Cluster analysis. 5th ed. Chichester: Wiley; 2011.
23. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20:53–65.
24. Kaufman L, Rousseeuw PJ. Finding groups in data, an introduction to cluster analysis. Hoboken: Wiley; 1990.
25. Pham DT, Dimov SS, Nguyen CD. Selection of K in K-means clustering. Proc Inst Mech Eng Part C J Mech Eng Sci. 2005;219:103–19.
26. Thinsungnoen T, Kaoungku N, Durongdumronchai P, Kerdprasop K, Kerdprasop N. The clustering validity with silhouette and sum of squared errors. In: International conference on industrial application engineering; 2015, p. 44–51.
27. Murray D, Chabot C. Tableau your data!: fast and easy visual analysis with tableau software. Hoboken: Wiley; 2013. p. 528.
28. Corp IBM. IBM SPSS statistics for windows version 25. Armonk: IBM Corp; 2017.
29. Kariuki JM, Manders E-J, Richards J, Oluoch T, Kimanga D, Wanyee S, et al. Automating indicator data reporting from health facility EMR to a

national aggregate data system in Kenya: an interoperability field-test using OpenMRS and DHIS2. Online J Public Health Inform. 2016;8:e188.
30. 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.
31. Irimu G, Ogero M, Mbevi G, Kariuki C, Gathara D, Akech S, et al. Tackling health professionals' strikes: an essential part of health system strengthening in Kenya. BMJ Global Health. 2018;3:1136.
32. Joseph Wu T-S, Kagoli M, Kaasbøll JJ, Bjune GA. Integrated Disease Surveillance and Response (IDSR) in Malawi: implementation gaps and challenges for timely alert. PLoS ONE. 2018;13:e0200858.
33. Ola O, Sedig K. Beyond simple charts: design of visualizations for big health data. Online J Public Health Inform. 2016;8:e195.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

J. Mantas et al. (Eds.)

doi:10.3233/SHTI200514

K-Means Clustering in Monitoring Facility Reporting of HIV Indicator Data: Case of Kenya

Milka B. GESICHO [a,d,1], Ankica BABIC [a,b] and Martin C. WERE [c,d]
[a] *Department of Information Science and Media Studies, University of Bergen, Norway*
[b] *Department of Biomedical Engineering, Linköping University, Sweden*
[c] *Vanderbilt University Medical Center, US*
[d] *Institute of Biomedical Informatics, Moi University, Kenya*

Abstract. Health management information systems (HMISs) in low- and middle-income countries have been used to collect large amounts of data after years of implementation, especially in support of HIV care services. National-level aggregate reporting data derived from HMISs are essential for informed decision-making. However, the optimal statistical approaches and algorithms for deriving key insights from these data are yet to be fully and adequately utilized. This paper demonstrates use of the k-means clustering algorithm as an approach in supporting monitoring of facility reporting and data-informed decision-making, using the case example of Kenya HIV national reporting data. Results reveal four homogeneous cluster categories that can be used in assessing overall facility performance and rating of that performance.

Keywords. HIV-indicator, dhis2, k-means clustering, monitoring, data-use

1. Introduction

Implementation of Health Management Information Systems (HMISs) for purposes of improving monitoring and evaluation efforts toward eradication of HIV in low- and middle-income countries has resulted in large amounts of data. Facilities using HMISs are required to submit various reports to aggregate-level HMISs[1], such as the District Health Information Software Version 2 (DHIS2) used in many countries [2]. These aggregate data are essential for program monitoring and evaluation (M&E) and for data-informed decision making (DIDM). DIDM is essential in informing policy and advocacy, and in program design, improvement, operations and management [2]. The ultimate aim of DIDM is achievement of improved health outcomes. For the submitted reports to be of best use to monitoring and evaluation (M&E) efforts, they must be complete, accurate and submitted in a timely manner. For the case of HIV, a weakness in understanding HIV information use infrequently addressed in previous studies is how M & E teams at the national level can utilize various approaches to derive insights from HIV facility reporting data aggregated in HMISs. In this study, we demonstrate use of the k-means

[1] Corresponding Author, Milka Gesicho, Department of Information Science and Media Studies, University of Bergen, Norway; E-mail: milcagesicho@gmail.com.

clustering algorithm as an approach in supporting monitoring of facility reporting and DIDM, using the case example of Kenya HIV national reporting data.

2. Methods

A retrospective observational study was used in monitoring performance trends in HIV reporting from health facilities in Kenya. DHIS2, the national aggregate reporting system, was used in extracting facility HIV reporting completeness and timeliness data for all health facilities in all 47 counties in Kenya, for the year 2011 to 2018. Systematic procedures were used in cleaning the data prior to analysis. Facilities in this study included only those offering HIV care and treatment services. This study explored an automated approach of grouping facilities based on their reporting completeness and timeliness, as a way of determining overall facility performance in reporting. Facility reporting completeness was defined as the extent to which facilities submit the expected number of reports, and timeliness as reporting submission within the defined reporting deadline. The actual number of reports submitted by facilities are automatically calculated within DHIS2, against the expected number of reports. The k-means clustering algorithm was used in identifying homogeneous groups within the data. The average silhouette coefficient was used in measuring the quality of the selected clusters [3]. All analyses were conducted in SPSS.

3. Results

A total of 18,394 HIV care and treatment reports from a total 3,242 facilities for the period 2011-2018 were evaluated. Based on the average silhouette measures for each year (ranging from 0.58 to 0.70); the k value used was four (k=4), with the four homogeneous groups of facilities identified as: best performers, average performers, poor performers, and outlier performers. Figure 1 to Figure 4 illustrate the exact performance (report timeliness and completeness) over time by facilities in each of these clusters. Figure 1 illustrates results for facilities in the best performers cluster, where average percentage completeness and timeliness was high (80% and above) in the various years (2012 to 2018).

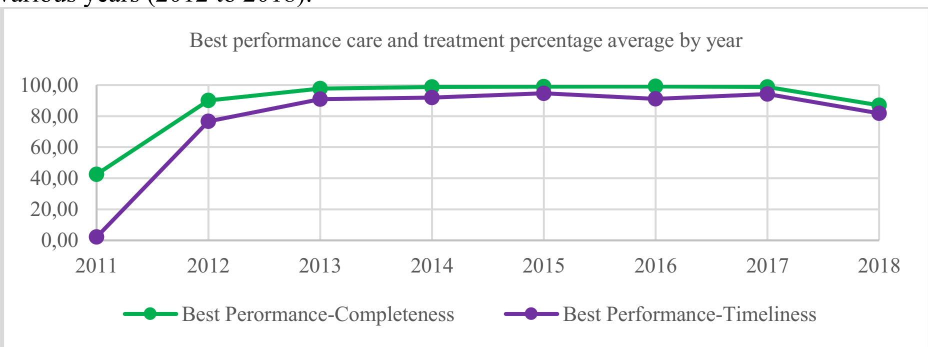

Figure 1. Care and treatment facility reporting best performance.

Figure 2 illustrates results for facilities in the average performance cluster, where percentage completeness and timeliness were lower in comparison to best performance facilities in the various years respectively. For instance, performance in 2015 for timeliness and completeness is lower by 28.65% and 5.37% respectively compared to performance in 2015 for best performance (Figure 1.).

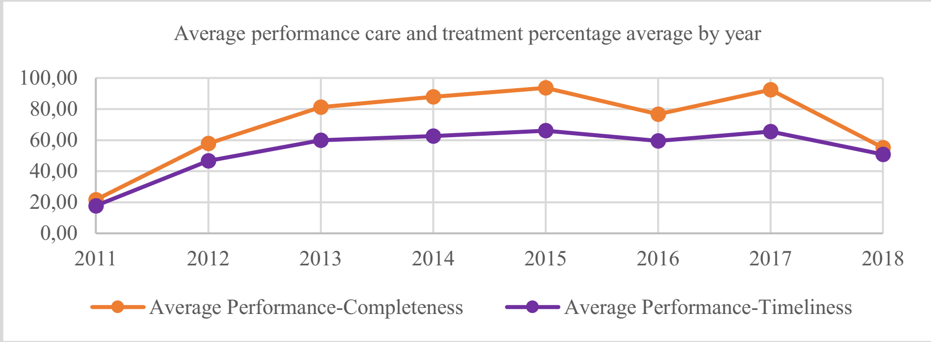

Figure 2. Care and treatment facility reporting average performance

Figure 3 illustrates results for facilities in the poor performance cluster, where percentage completeness and timeliness was low (below 50%) in the various years.

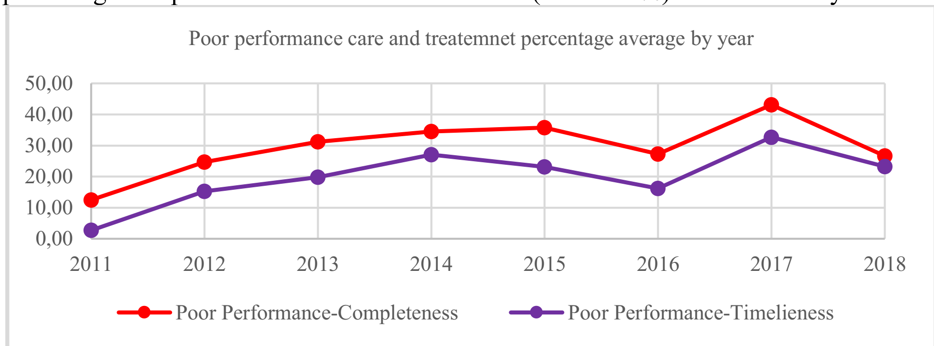

Figure 3. Care and treatment facility reporting poor performance.

Figure 4 illustrates results for facilities in the outlier performance cluster, where there was an evidently big gap between percentage completeness and timeliness in the various years. This depicts scenarios where timeliness was a problem despite good performances in completeness.

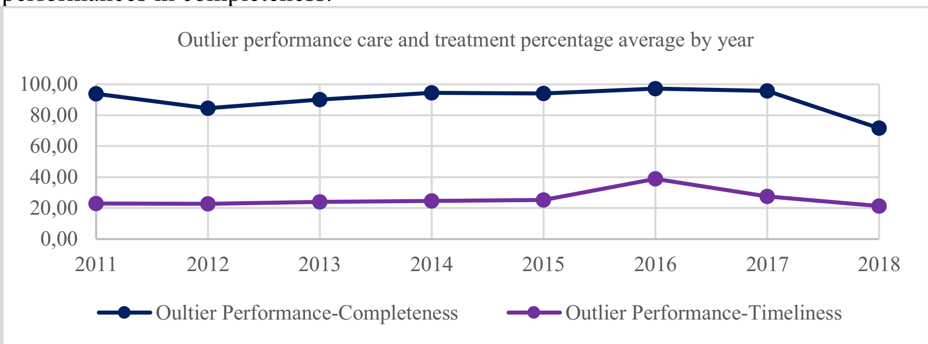

Figure 4. Care and treatment facility reporting outlier performance.

4. Discussion

In this paper we illustrate use of the k-means clustering algorithm as an approach in assessing retrospective HIV facility reporting to determine facility performance, as a way of informing reporting improvement mechanisms. These analyses provide at a glance view of various categories that emerge based on performance of facilities in meeting completeness and timeliness requirements in reporting. These results can be used by M&E teams to identify facilities whose performance is satisfactory or not, therefore providing a baseline for further evaluations and development of sustainable solutions.

Furthermore, due to the volume, velocity and veracity of health data consolidated from various sources, representing various facets of data in a way that makes sense is a challenge. In this study, we used line graphs, which are simple visualizations that can be used to represent data in a way that promotes development of insights at a glance. Figure 1 portrays an ideal situation of good facility reporting. If success is to be achieved in terms of meeting reporting requirements, then the ultimate goal for these evaluations should be to enable all facilities to attain and maintain similar results as illustrated in the best performing category. A common attribute among average, poor and outlier performance is the discrepancy between completeness and timeliness as represented by the gaps observed between them in the respective performance categories. There is therefore need for investigating issues that bring about delays more so in the outlier performance group, which has the largest gap in the completeness and timeliness measures. As the next step, we will further disaggregate the results by facility characteristics and geographic region, and also look at additional reporting domains.

5. Conclusions

The k-means clustering algorithm is essential in automatically finding homogenous groups within aggregate reporting data. This serves as a good baseline for monitoring the progression of health facility reporting performance by management and M&E teams that use large amounts of data collected from integrated data sources.

Ethical approval and Acknowledgements

Ethical approval was obtained from the IREC in Moi University-No.0003362. This work was supported in part by the NORHED program (Norad: Project QZA-0484). The content is solely the responsibility of the authors.

References

[1] Manya A, Nielsen P, Reporting practices and data quality in health information systems in developing countries: an exploratory case study in Kenya, J. Health Inform. Dev. Ctries 10 (2016), 114–126.

[2] Karuri J, Waiganjo P, Orwa D, Manya A, DHIS2: The Tool to Improve Health Data Demand and Use in Kenya, J. Health Inform. Dev. Ctrie 8 (2014), 38–60.

[3] Kaufman L, Rousseeuw PJ, Finding Groups in Data, An Introduction to Cluster Analysis. John Wiley & Sons, Inc, New Jersey, 1990.

PLOS ONE

RESEARCH ARTICLE

Identifying barriers and facilitators in HIV-indicator reporting for different health facility performances: A qualitative case study

Milka B. Gesicho[1,2☯]*, **Ankica Babic**[1,3☯]

1 Department of Information Science and Media Studies, University of Bergen, Bergen, Norway, **2** Institute of Biomedical Informatics, Moi University, Kesses, Kenya, **3** Department of Biomedical Engineering, Linköping University, Linköping, Sweden

☯ These authors contributed equally to this work.
* milcagesicho@gmail.com

Citation: Gesicho MB, Babic A (2021) Identifying barriers and facilitators in HIV-indicator reporting for different health facility performances: A qualitative case study. PLoS ONE 16(2): e0247525. https://doi.org/10.1371/journal.pone.0247525

Editor: Vijayaprasad Gopichandran, ESIC Medical College & PGIMSR, INDIA

Received: December 18, 2020

Accepted: February 9, 2021

Published: February 25, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0247525

Data Availability Statement: All relevant data are within the paper.

Funding: This work was supported in part by the NORHED program (Norad: Project QZA-0484). link: (https://www.norad.no/en/front/funding/norhed/

Abstract

Identifying barriers and facilitators in HIV-indicator reporting contributes to strengthening HIV monitoring and evaluation efforts by acknowledging contributors to success, as well as identifying weaknesses within the system that require improvement. Nonetheless, there is paucity in identifying and comparing barriers and facilitators in HIV-indicator data reporting among facilities that perform well and those that perform poorly at meeting reporting completeness and timeliness requirements. Therefore, this study aims to use a qualitative approach in identifying and comparing the current state of barriers and facilitators in routine reporting of HIV-indicators by facilities performing well, and those performing poorly in meeting facility reporting completeness and timeliness requirements to District Health Information Software2 (DHIS2). A multiple qualitative case study design was employed. The criteria for case selection was based on performance in HIV-indicator facility reporting completeness and timeliness. Areas of interest revolved around reporting procedures, organizational, behavioral, and technical factors. Purposive sampling was used to identify key informants in the study. Data was collected using semi-structured in-depth interviews with 13 participants, and included archival records on facility reporting performance, looking into documentation, and informal direct observation at 13 facilities in Kenya. Findings revealed that facilitators and barriers in reporting emerged from the following factors: interrelationship between workload, teamwork and skilled personnel, role of an EMRs system in reporting, time constraints, availability and access-rights to DHIS2, complexity of reports, staff rotation, availability of trainings and mentorship, motivation, availability of standard operating procedures and resources. There was less variation in barriers and facilitators faced by facilities performing well and those performing poorly. Continuous evaluations have been advocated within health information systems literature. Therefore, continuous qualitative assessments are also necessary in order to determine improvements and recurring of similar issues. These assessments have also complemented other quantitative analyses related to this study.

projects/health-informatics-training-and-research-in-east-africa-for-improved-health-care-hi-train/) The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

In a bid to eradicate the HIV epidemic, enormous strides have been made toward achieving the UNAIDS "90 90 90" ambitious target [1]. The goal of these targets was that by 2020, "90% of all people living with HIV will know their HIV status; 90% of all people with diagnosed HIV infection will receive sustained antiretroviral therapy; and 90% of all people receiving antiretroviral therapy will have viral suppression" in order to end the epidemic by 2030 [1]. Therefore, tracking the national response of HIV is salient in determining progress as well as recognizing whether efforts to scale up HIV services are of value. As such, Monitoring and Evaluation (M&E) systems, which are regarded as the cornerstone of HIV services, have been established in low and middle-income countries (LMICs) to provide high quality strategic information for decision-making [2,3].

To measure HIV program effectiveness and patient outcomes, Ministries of Health (MoH), as well as international donor organizations, such as Presidential Emergency Plan for AIDs Relief (PEPFAR) require the various health facilities to report several aggregated indicators as part of M&E program [4]. The scale up of HIV services in LMICs has resulted to strengthening of Health Management Information Systems (HMIS) to facilitate with collection, management, and availability of timely, complete and accurate data. Therefore, HMISs such as the District Health Information Software Version 2 (DHIS2) have been implemented in over 70 countries to promote availability of routine aggregated indicator data within health care [5]. As such, routine reporting of HIV-indicators in many LMICs is performed using the District Health Information Software Version 2 (DHIS2) [6].

Therefore, aggregate indicator data from various HIV services are collected using paper-based summary forms at the facility level, which are then entered into the DHIS2. These data are required to be submitted monthly to the national level by health facilities within stipulated timelines. As such, the HIV-indicator data are used for M&E by converting the raw data collected, to information for utilization in decision-making. Nevertheless, M&E systems in LMICs experience significant challenges in completeness, timeliness and accuracy in reporting data [2,7]. Some of these challenges are often brought about by various issues revolving around organizational, technical and behavioral factors [8].

Moreover, identifying barriers and facilitators in HIV reporting contributes to strengthening HIV M&E efforts by acknowledging contributors to success, and identifying weaknesses within the system that require improvement [3,9–12]. As such, various assessments have been conducted in a bid to improve performance of facilities at meeting data quality reporting requirements such as accuracy, completeness and timeliness, which are essential for M&E [9,13–15]. Nonetheless, qualitative assessments that follow these evaluations are meagre. This will facilitate understanding similarities or differences across various issues within facilities with varying performances, given that some facilities perform better in meeting data quality reporting requirements than others. As such, there is paucity in identifying and comparing barriers and facilitators in HIV-indicator reporting among facilities that perform well and those that perform poorly at meeting reporting requirements such as completeness and timeliness.

This study aimed at conducting qualitative case study using Kenya as an example, to identify and compare the current state of barriers and facilitators in routine reporting of HIV-indicators, based on facility performance. This was conducted among facilities performing well, and those performing poorly in meeting DHIS2 reporting requirements (facility reporting completeness and timeliness). The findings of this study aimed at contributing towards strengthening HIV-M&E efforts, which are of interest to various stakeholders including ministries of health.

Method

Study setting

This study was conducted in Kenya's capital Nairobi. Kenya is a sub-Saharan country in East Africa with 47 administrative counties, and it's health system is categorized in six levels, which include (i) community services, (ii) dispensary/clinics, (iii) health centers, (iv) sub-county hospitals, (v) county referral hospitals, and (vi) national referral hospitals [16,17].

Study design

A qualitative case study approach was used. Miles and Huberman define a case as "a phenomenon of some sort occurring in a bounded context" [18]. Creswell on the other hand defines a case as "a program, organization, event, activity, process, or one or more individuals" [19]. As such, the "case" in this study is defined as a health care facility offering HIV services. The cases in this study were bounded by context, which includes only HIV healthcare facilities that meet the following criteria (i) located in Nairobi (ii) either use EMRs system or paper in reporting, (iii) reporting performance (facility reporting completeness and timeliness).

Data collection

Purposive sampling was used to identify the cases (health facilities) from constituencies in Nairobi [20]. In addition, the cases were also drawn from level two and level three of the health system, which comprises of clinics and health centers. Hence, the type of purposive sampling used was stratified purposeful sampling, whereby health facilities from the two levels were selected based on reporting performance [21].

Thus, these cases (health facilities) were selected based on performance in facility reporting completeness and timeliness of Care and Treatment reports for the years 2017 and 2018 with more details outlined in Gesicho et al. [21,22]. Facility reporting completeness was defined as the percentage of actual reports submitted to DHIS2 against the expected reports, whereas facility reporting timeliness was defined as the percentage of actual reports submitted to DHIS2 on time, against the expected reports.

In Kenya, HIV-indictor reports are submitted in DHIS2 on a monthly basis by facilities offering HIV services using the MOH-mandated form called "*MOH731- Comprehensive HIV/AIDS Facility Reporting Form*" (MOH-731). Care and Treatment is among the most reported and salient HIV services offered by health facilities in Kenya. The criteria for selection was based on HIV reporting performance by facilities, which was categorized as best performers, average performers, poor performers and outlier performers. This grouping was based on a cluster analysis conducted in order to evaluate the reporting performance by facilities using completeness and timeliness as performance indicators as outlined in Gesicho et al. [21,22]. In this study, reporting performance was categorized into two main groups, facilities performing well (best performers = 3 facilities, average performers = 3 facilities) and facilities performing poorly (outlier performers = 4 facilities and poor performers = 2 facilities). One health facility where a sub-county office is located was also included to provide more information on reporting by facilities as all facilities are required to submit paper-based reports to their respective sub-county office.

Purposive sampling was also used to identify key informants in the study in order to conduct in-depth interviews [20]. Therefore, the key informants who in this study are the units of analysis included personnel in charge of reporting as they serve as the focal point around which all reporting activities take place.

Data was collected using semi-structured in-depth interviews with 13 participants, in 13 health facilities, with one facility visit being specifically to the sub-county office. These were drawn from six of the 17 constituencies in Nairobi, which include Kasarani, Embakasi North, Embakasi South, Embakasi East, Embakasi Central, Kamukunji, Dagoretti South, and Dagoretti North. Archival records on facility reporting performance, documentation and informal direct observation were also used to collect the data. These sources of evidence aimed to ensure improved credibility through data triangulation. Archival records such as retrospective quantitative data on facility reporting timeliness and completeness were retrieved from DHIS2 in order to identify the reporting performance of a facility. Documentation such as hard copy standard operating procedures, and data quality assessments (DQA) reports were sought and perused. Informal observations were also carried out during the assessments, which involved observing the documents put on office walls, office space, interaction with colleagues and environment where the facility was located.

The in-depth interviews were conducted using an interview guide. The interview guide explored the technical, behavioral, and organizational factors, which facilitated or hindered the reporting process in the health facilities based on participants' experiences and perceptions. Interviews took place in enclosed workstations of the participants and lasted approximately between one to one and a half hours. Data saturation was achieved in the sampled cases given that no new data emerged during the final interviews. All data were collected between September 2019 and November 2019.

Interview guide

The interview guide used in this qualitative assessment was based on the conceptual framework for performance of the routine information system management (PRISM), which was developed to strengthen routine health information system (RHIS) performance management [8]. A routine health information system (RHIS) is comprised of inputs (RHIS determinants), processes (RHIS process) and outputs (improved RHIS performance), which are components of a routine health information system [8]. We explored the RHIS determinants of performance and RHIS process specified within the PRISM framework. The RHIS determinants include technical (complexity of reporting form, procedures), behavioral (competence, motivation) and organizational determinants (availability of resources, training). The RHIS process elements explored include: data collection process, data processing, data transmission and data quality checking and feedback mechanisms. This study aimed to identify and compare the barriers and facilitators linking to RHIS determinants and process, among facilities in the four performance cluster categories (best performer, average performers, poor performers, and outlier performers). For detailed information on the interview guide (See S1 Appendix: Interview guide).

Ethical considerations

Ethical approval was obtained from the Institutional Research and Ethics Committee in Moi University-No.0003362. Other approvals were obtained from the Ministry of Health, Nairobi County and from affiliated constituencies where the facilities were sampled from. Research license was obtained from the National Commission for Science, Technology & Innovation in Kenya. Privacy and confidentiality were ensured by not revealing the identities of the participants nor the facilities that took part in this study. Written informed consent was obtained from all participants who were interviewed.

Data analysis

Data from interviews and data sources were analyzed together. This followed the analysis framework developed by Morse, which outlines four key cognitive processes used in developing theory from data [23]. These processes include comprehension, synthesis, theorizing and recontextualizing. The process of comprehending begins during data collection. Comprehension also involves coding, which enables sorting data, and uncovering underlying meanings in the text [23]. The synthesis process involves aggregating several stories or cases to describe typical, composite patterns [23]. Content analysis was used in the comprehension and synthesis processes. A provisional 'start list' of codes was created based on the research questions in order to make the coding process manageable. Theorizing involves a systematic process that entails finding alternative explanations until an explanation that best fits the data is sought [23]. Within-case and cross-case comparisons were used in the theorizing process by using cross case displays presented by Miles and Huberman [18]. Only analyzed cross-case data were presented in order to ensure that the confidentiality of sites, which may be identifiable from the within-case analysis. Re-contextualization involved comparing the findings with previous research in order to enhance trustworthiness. QSR NVivo was chosen as the Computer-Assisted Qualitative Analysis Software (CAQDAS), were all data was managed [24].

Trustworthiness

Various approaches were used in order to ensure trustworthiness of the study. To ensure credibility of the study, prolonged engagement with participants was conducted long enough to gain trust and establish rapport [25]. Triangulation of data sources was conducted using aforementioned multiple sources. Peer debriefing was also carried out in sessions after conducting two to three interviews and during analysis [25]. To ensure dependability and confirmability, an audit trail using QSR NVivo (Version 12), was used to manage and store data. Reflexivity was achieved on the basis that the researcher was a stranger in the facility settings hence had no established familiarity with participants prior to the study, which may be a threat to validity (bias) [25]. Transferability was attained through provision of thick descriptions hence enabling applicability of findings in other similar settings using DHIS2 to submit monthly reports.

Result

A summary of the findings is presented in Table 1.

It is worth noting that the key respondents interviewed identified their positions as either data officers or M&E assistants but performed the same tasks. Hence, the position title depended on the term used by facility or supporting partner in charge of a facility. Thus, the term 'data officers' is used as a general term in this study to refer to the people mandated with data reporting process in health facilities. Detailed findings revolving around technical, organizational, and behavioral factors are outlined below.

Interrelation between workload, teamwork, and capacity in reporting

Emerging interrelated factors that influenced the reporting process in the various facilities include workload, teamwork, and human resource capacity. As such, the presence or absence of a combination of these factors either positively or negatively influenced the reporting process. Workload referred to the amount of work present in each health facility and this varied from facility to facility with some having more workload than others. Teamwork in this case means that the people involved in reporting assist each other in collection, aggregation, and verification of data. For instance, in some cases nurses at the various service points are required to update correct aggregate numbers for indicators and submit them on time during the reporting period to the data officer for verification. An informal observation made when conducting interviews was the interaction between the nurses, clinicians and the data officer, which revealed a sense of putting in effort to provide the data required by the data officers. For

Table 1. Summary of findings on barriers and facilitators in HIV-indicator reporting.

Factors	Category	Summary of key findings
1. Interrelationship between workload, teamwork, and capacity in reporting	Organizational	Presence of teamwork as well as sufficient number of skilled personnel in facilities with a lot of workload emerged as a facilitator in the reporting process regardless of facility performance group.
2. Role of an EMRs system in easing the reporting process	Reporting process, Technical	Presence of an EMRs system facilitated the ease of data collection in the reporting process. Nonetheless, it did not equate to good performance as two of the facilities in the best performance group were not using an EMRs system whereas all facilities in the outlier performance group had an EMRs system.
3. Reporting timeframe and adherence to reporting deadlines	Reporting process	Reporting days falling on weekends emerged as a major time constraint in the reporting process among most of the facilities visited. Facilities in the various performance groups echoed the same issue.
4. Access rights and availability of national aggregate reporting system	Technical	Lack of access rights to DHIS2 by facilities was a contributor to late submission of reports. This is especially when reports submitted by hand to the sub-county are entered late in DHIS2. Lack of availability of DHIS2 due to issues such as system down times and lagging internet contributed to slowing down the reporting process.
5. Complexity of reports, staff rotations, and role of mentorship in reporting	Technical	Documentation errors were among the main issues resulting from these factors regardless of facility performance group.
6. Fit between individual, task and technology in reporting	Technical	Fit between individual task and technology is a facilitator in reporting. Facility that lacked fit between individual and task reported lack of motivation.
7. Motivation and awareness of reporting performance	Behavioral	Data officers and M&E assistants interviewed used ad hoc approaches to determine their individual performance in submitting reports. For instance, some of them mentioned that once they have submitted a report on time and with no questioning of the data, then they have performed well. Nonetheless, facilities generally depended on key administrators such as in-charges to provide feedback on performance. Good feedback was a motivating factor.
8. Availability of Standard Operating Procedures, Training, and Supervision	Organizational	Regardless of reporting performance, facilities funded by partners had SOPs and on job trainings whereas those not funded did not. Supportive supervisions were also reported to be present but not frequent.

https://doi.org/10.1371/journal.pone.0247525.t001

instance, in some facilities, an observation was made on nurses bringing updates on aggregated data written on paper to the data officer On enquiry, the interviewee responded that regular updates promoted data quality as it enabled accountability when performing weekly and monthly verification of patient numbers per indicator.

If this teamwork is not present, the data officer may have a difficult time in collecting the aggregated indicator data. On the other hand, human resource capacity referred to sufficient number of skilled personnel involved in the reporting process. Hence, the interrelationship comes about in the sense that more workload required skilled human resource as well as teamwork in order to get the job done.

As such, respondents attributed time-consuming data collection process to more workload. The amount of workload was determined by factors such as services offered as well as the number of patients visiting the facilities. Health facilities with well-established Comprehensive Care Centers (CCCs) offered more HIV services and therefore received more patients. Hence, such facilities required to use more registers due to many service points as well as patient files. As such, human resource capacity as well as teamwork is important as stated by some of the respondents:

*"You know in other facilities, you find the clinician is one, he is the one who is the pmtct, he is the one seeing all other CCC patients, sometimes they have only one or two counselors depending on the workload. Workload in facility X CCC alone has 3500 patients, we have 3 clinicians, two nurses, 8 counselors. So those are so many registers."- **Data officer facility K (outlier performer)***

*"Basically I should be doing the reports but since the testing points are so many, they do compile their report for the partner and give me a copy"- **Data officer facility C (best performer)***

As such, having to deal with many registers can also affect the quality of data since it is time consuming for the data officer to perform assessments given that there was only one data officer in most facilities visited. As stated by one of the respondents:

*"Ideally, I am supposed to go register by register counting, that is a quality assessment. Am supposed to be counting. At times, I may not be having that time sincerely. But the supervisor is also thorough."—**Data officer facility A (best performer)***

Teamwork then plays a big role in facilities that have a lot of workload. For teamwork to be effective, there needs to be trained personnel involved, who have a good knowledge and understand the indicators. This limits the documentation issues hence making the process faster as there will be minimal corrections in the data as reported indictors will be tallying. In cases where teamwork is present, but the personnel involved in reporting process do not have a good knowledge and understanding of indicator data, issues of documentation frequently arise. Moreover, in some cases the data officer has to do the collection and verification of indicator data as there is no one to provide a summarized report, which then consumes time taken to collect data. This then slows down the reporting process as data needs to be verified and amended.

Moreover, respondents in facilities that used an EMRs system for retrospective data entry cited backlogs due to accumulation of patient files, which need to be input to the system before the reporting period. The data in the systems need to be up to date in order to generate reports and queries needed for populating the MOH-731. In most cases, there was only one person who was assigned with the retrospective data entry to the EMRs system, which then contributed to backlog when they were absent from duty. Some of the data officers interviewed mentioned that they felt overwhelmed especially during the reporting period. As stated by one of the participants:

*"Like this week, I will get a huge back log because I have to generate the reports so by the time I am generating the reports I will not do data entry. You know I am the same person doing data entry of the daily activities and then I am the one who generates the reports. So, the reporting week I am always getting overwhelmed"–**Data officer facility J (average performer)***

Hence, teamwork, as well as sufficient number of skilled human resource is salient especially in facilities with a lot of workload as this can either slow down or speed up the reporting process.

Role of an EMRs system in easing the reporting process

Some of the facilities visited, had EMRs systems implemented as point of care (one facility) or retrospective data entry (four facilities). In other facilities, the EMRs systems were non-functional meaning that they were present but not being utilized in reporting or for other tasks in the facility (three facilities). In the facility where EMRs system was used as point of care, data entry was done by clinicians. Nevertheless, the data officers were required to still verify the data in the systems with that in the patient files before submitting reports. For those that were not point of care, data was entered retrospectively from the patient files.

EMRs systems were particularly useful in obtaining HIV-indicators for Care and Treatment (CRT) using queries. Hence, EMRs systems played a role in reporting by contributing to faster data collection, thus reducing the time taken by data officers in collecting data for various indicators from registers. The ease of reporting attributed by EMRs systems can be described using the following responses from respondents:

> *"It is very easy compared to the registers. If I was to use the registers when doing this work, I would have a lot of calculators and tally sheets. If you see the registers that were there then, there was a lot of paperwork, a lot of time, a lot of documentation. When you have something that actually sums it up and all you have to do is little."–**Data officer facility I (outlier performer)***

> *"It makes it easy for me to get data because I just run the queries for the data that I want. Instead of going back to the files to count the data one by one how many have been started on treatment, and looking for the ages, those enrolled etc."–**Data officer facility A (best performer)***

In contrast, facilities where EMRs systems were not implemented had to do the reporting process manually, which was time consuming. Nonetheless, some of the facilities had EMRs system but were not being utilized due to factors such as data migration of files to the EMRs system, lack of human resource designated to use the ERMs system such as clinicians, and system challenges. As stated by the respondents:

> *"Until now they brought in a new clinician, whose now three months in. But if we had this clinician and the system was put in place, it would be running."–**M&E assistant facility C (best performer)***

> *"It is there but we did not have a clinician. For it to be fully integrated has been a problem. A clinician just came in the other day, now is when she has started to use the EMR system"–**M&E assistant M&E assistant facility B (best performer)***

Further still, there seemed to be a demarcation regarding acceptability of roles and responsibilities between clinicians hired by the county and those hired by the partners. As one respondent stated:

> *"By the way there is an issue between the clinicians, the staff from the county who are being told to work for a program. So, they will just work. You know a partner needs more than the government when it comes to CCC"–**M&E assistant facility C (best performer)***

Moreover, in the facilities visited, clinicians interacting with the EMRs system for HIV services were hired by the partners. This enabled clinicians to only focus on CCC as it emerged that clinicians hired by the government had other tasks and were hesitant to perform HIV related tasks, which also included use of EMRs systems in HIV services.

Reporting timeframe and adherence to reporting deadlines

All the health facilities visited were aware of the reporting deadline and repercussions for late submission of reports. Given that repercussions were imposed for late report submission, facilities were keen on ensuring that both hardcopy and softcopy reports were submitted within the set deadlines regardless of the facility location. The aspect of location came about as an informal observation was made whereby some of the facilities were located inside densely populated slums, which gathered mud during rainy seasons further adding to the challenge of accessibility in and out of the facility.

Health facilities that had no access to DHIS2 were required to submit their reports by 5th of every month to their sub-county. On the other hand, deadlines for facilities with access to DHIS2 varied with some facilities having strict deadlines of 5th, while others having a grace period up to the 10th of every month. Nevertheless, time constraint challenges in a bid to beat the reporting deadline were stated among the respondents, as some viewed the reporting period as being too tight. Some of the factors that contributed to this notion include time taken to revise the documentation issues, reporting days falling on weekends, assignment of data officers to more than one reporting site, and parallel reporting.

Documentation issues identified in the reports when conducting data quality checks needed to be corrected before the reports are submitted. In some cases, identifying some of the errors can be time consuming. For example, issues such as identifying "true missing", which results from cases where the reported numbers are too low thus raising an alarm. Hence, finding the correct number of people who missed their appointments can be time consuming and may contribute to delays in reporting. As stated by one respondent who was correcting missed appointments at the time of the visit:

> *"My biggest challenge is timelines. Like you see just know they are asking me to send the corrections, which I have not sent and I am actually working on. So, timelines is the issue."–**Data officer facility I (outlier performer)***

Another issue is reporting dates falling on weekends. This means that the weekend takes up the reporting days, which then reduces the number of days for reporting. This was an issue of concern in most facilities visited, as the deadlines are not extended regardless of the circumstance. As some of the respondents stated:

> *"But there is this time when weekend takes the whole of your date 1, 2, 3 and then date 4 is Monday, that is when you are starting reporting and you are supposed to submit by date 5. Which is tomorrow, that is a challenge"–**Data officer facility K (outlier performer)***

> "*Sometimes when the weekend eats up the reporting time, it's a challenge.*"*–**Data officer facility A (best performer)***

Further still, there are data officers who have been assigned more than one facility to support in HIV-indicator reporting. Given the few reporting days, they are often required to organize their schedules to ensure that reports are submitted on time. In some cases, this may lead to delay as some facilities may be given more priority than others. As stated by one of the respondents:

> Sometimes I may focus on another facility. Because this is a small facility, it may just take 3 hours to finish my report. So by the time I come here, they have submitted the other reports and only remain with the MOH-731. So I will just do it and take it to sub county or give them to submit to sub county. So, I don't know which day it goes. But when it is me who takes it, I normally take it late after the deadline.–**M&E assistant facility G (poor performer)**

> *No. I don't put the same day. I also have another facility which I support, so you find first day of reporting I am here then second day of reporting I am there.–**M&E assistant facility B (best performer)***

Furthermore, most of the facilities are funded hence required to do reports required by the MoH (MOH-731) and those required by supporting partners -Data for Accountability

Transparency and Impact (DATIM). As such, these two HIV summary reports are required per facility in cases where funding is involved, which may be tasking for a data officer reporting for more than one facility. One of the respondents termed this as "roving".

> "*. . . but we have people that rove. But I don't think roving gives someone concentration to ensure quality. Because you have to think of two things at a time. I think it should be one individual per facility.*"–***Data officer facility L (outlier performer)***

Access rights and availability of national aggregate reporting system

For facilities to be authorized with credentials to access DHIS2, they have to meet a certain criterion, which the sub-county assesses. As such, not all facilities have access rights to the DHIS2. As such, they are required to submit hard copy MOH-731 reports to the sub-county in order for the reports to be entered in the system. Nevertheless, sub-counties especially those with many constituencies are faced with challenges such as lack of capacity to ensure that all reports are entered on time. This is because, unlike ordinary data entry, MOH-731 reports require numbers to tally otherwise the data will be questionable. DHIS2 has data quality mechanisms that flag these errors within the data upon data entry. This enables facilities to detect and amend questionable data before report submission.

As such, in situations where health facilities with no access-rights to DHIS2 submit hard copy MOH-731 reports for data entry at the sub-countiesMOH-731, errors were more likely to be encountered by the health records information officers (HRIOs) during data entry to DHIS2. As a result, sub-county HRIOs are tasked with contacting facilities through phone calls in order to clarify the errors encountered in data. As stated by one health records information officer at the sub-county:

> *"The goodness with the system, for example if you get somebody has put that they have counselled 20 people, and they tested 19, the system flags out, it will alert that why would this facility counsel 20 people and test 19? Where is the 1? So it is now up to me to take my phone and make a call to that facility. You see now they explain. Maybe if it was a transcription error, we are able to go there, because you know they need to correct from the source documents. They don't just tell me "oh no it was wrong, it is 20/ 20". No, we need to go there and counter check, we sign and they sign against that number, so that we are able to put it in the KHIS.* "- ***sub-county Health records Information Officer***

A major issue that emerged during an interview at one of the sub-county offices revealed that the ratio of sub-county officer's to facilities is low in that, the number of facilities are more than sub-county HRIOs. For instance, there are sub-counties with four officers (including volunteers) working with over 50 facilities, which do not submit data to DHIS2 at facility level.

This leads to late data entry to DHIS2, which then affects timeliness. Facilities categorized with outlier performance seemed to have been plagued with this issue based on the interviews conducted. As one respondent stated:

> "*. . . There was a time they (facility) never had the credentials for DHIS2, so I could not upload. So, it depends when did they (sub county) upload.*"–***Data officer facility I (outlier performer)***

In addition, it is important for systems to be available whenever they are needed for use. Otherwise, this may frustrate efforts of users in trying to accomplish their tasks. Interviewees who used EMRs systems reported that it was always available when needed. This is because the

EMRs systems did not require internet connection. However, there were varied concerns regarding availability of the DHIS2 especially during reporting periods. Given that DHIS2 requires internet connection, some of the issues highlighted include lagging of the internet connection and system down times. As some of the respondents stated:

> "*. . .. When the internet is not working properly and DHIS2 sometimes has an issue, it says bad gateway. So, when DHIS itself has an issue it can delay me."*–***Data officer facility A (best performer)***

> "*Some of these delays are also caused by the system. It goes down, it takes time to load, sometimes we do not have network, and sometimes the system has just gone down"*–***sub-county Health records Information Officer***

> "*Then we have internet issues, at times the internet is slow, so inputting the data to dhis2 you have to go an extra mile do it in the evening or very early in the morning when it is not crowded."*–***Data officer facility J (average performer)***

System availability issues in DHIS2 were also reported at the sub-county office visited. Therefore, absence of system availability contributed to delays in reporting or slowing down the reporting process at both facility and sub-county level.

Complexity of reports, staff rotations and role of mentorship in reporting

The reports in the MOH-731 were either categorized as complex or easy by respondents, based on the number of registers required, number of indicators, and amount of knowledge required to understand the indicators. CRT reports were among those considered by interviewees as complex, which required knowledge and understanding of the indicators. These reports also took more time to prepare compared to the others especially in facilities with more workload. These reports were also likely to have more documentation issues compared to the others. A HRIO at the sub-county whose office receives reports from various facilities stated as follows:

> "*Same to Care and treatment, the indicators are very technical for our health care workers to understand. So you find a lot of transcription errors. It is data that for you to rely on, it needs serious people, people who really understand the data."*–***sub-county Health records Information Officer***

Documentation issues identified include wrong calculations, gaps within the data, legibility of written figures, misinterpretation of the indicators and codes, and failure to update the registers. Wrong calculations arose in situations where the figures aggregated by nurses or clinicians did not match with those of the data officers. Gaps within the data include scenarios where for instance ten people were counselled and only nine were tested, or ten pregnant women were identified as positive and only five were given prophylaxis. These documentation issues need to be corrected by the data officers before submission of the reports. Documentation issues therefore contribute to delays or slowing down the reporting process more so in cases where there is parallel reporting, which entailed submitting similar reports to both the MoH (submission of MOH-731) and supporting partners (submission of DATIM). As such, data reported in DATIM should be similar to that reported in the MOH-731. A challenge posed by interviewees was amending both reporting tools when documentation issues are encountered. Responses from interviewees on issues of documentation are as follows:

"*. . .since I had to submit both datims and 731 by 5th, it was hectic. You know you have to work on the negligible stuff like numbers, they don't tally, and then you find an undocumented register, not well documented, it still delays you."–**Data officer facility I (outlier performer)***

"*. . . or if something was missed in the register and you want to report, they have to fill the gap and you know for instance a HEI was drawn a PCR at 24 months that is 2 years later, if you do not fill in the results and then you are supposed to report on the outcome or IPT register, anyway the challenge would be I the clinician or nurse is not filling in the register. "–**Data officer facility K (outlier performer)***

Gaps were also identified in some of the reports perused during interviews, which had gone unidentified in previous years. Post-Exposure Prophylaxis (PEP), Voluntary Medical Male Circumcision (VMMC) and Blood Safety (BS) were among reports considered to be easy as they contained few indicators. In addition, majority of facilities did not offer VMMC and BS, (replaced with Methadone Assisted Therapy at time of study) hence reporting these indicators was easy as it was a matter of just recording zeros. In addition, for facilities that offered VMMC, it happened occasionally especially when students where on holiday as this facilitated performing of the minor circumcision surgeries.

Another contributing factor to documentation is the issue of rotation of nurses involved in reporting and staff turn overs. For instance, at facility level, nurses that were in one service area may be transferred to another service area hence need to be retrained in order to prevent documentation issues. These staff rotations and staff turnovers can be cumbersome to some of the data officers as expressed in the interviewees' responses:

"*. . .you have completely new people who have never even seen an ANC register, you take them indicator by indicator. Then when they catch up, they leave. Then you start all over again"–**Data officer facility I (outlier performer)***

"*But the blunder in private facilities is high staff turnover, there is this staff who is doing a good job this month the next month that one got greener pastures, there is a new one again. And then you start again mentoring after a couple of weeks that one again disappears."–**sub-county Health records Information Officer***

Data officers therefore conduct mentorships through training of indicators and HIV reporting at the various service points, whereas supportive supervisions are conducted on need be basis by the sub-county health records information officers. This in turn promotes better documentation of registers. This is also intended to decrease delays during the reporting period caused by time taken to amend documentation issues.

Fit between individual, task and technology in reporting

A salient aspect in the reporting process for an individual is the fit between task and fit between technology. This means that the individual has the right competency for the task and technology or tools used to complete the task [26]. All facilities apart from one in this study had employed data officers or M&E assistants, which ensured fit between individual, technology and task. Moreover, all the data officers or M&E assistants interviewed had studied at least health records or both health records and information technology. In addition, those utilizing EMRs systems had received on job training on using various aspects within the EMRs system such as using queries to retrieve data, and trouble shooting. The facility with no data officer happened to be a private facility, and a nurse and a pharmacist were in charge of reporting.

Thus, a misfit between individual and task was identified. As such, this was extra work to the pharmacist who doubled up in doing the reports as stated:

> "*You see this is extra work aside from my main work. I told the people form the ministry to give the nurse to be doing this work. They said nurse cannot order drugs, so I said I will just do it because it is once a month.* ***Pharmacist facility D (average performer)***

Hence, using the existing employees to multitask in reporting rather than employing a data officer has a potentially negative impact on the data quality of reports as well as motivation of the employees.

Motivation and awareness of reporting performance

Factors contributing to motivation in reporting among the interviewees include passion for the work, patient well-being, good performance appraisal feedback, gaining insights from the data and support from supervisors. Moreover, it emerged that the presence of motivation implicitly contributes to data quality in terms of accuracy, completeness, and timeliness in reporting. This was revealed by how the respondents reported a sense of commitment towards the patients, which prompts them to strive for good reporting as stated:

> "*You know a camera guy who always gives good shots but is never seen in the pictures. I am the guy that I will not see the patient, but I will make sure that I tally everyone. I know everyone by their numbers, and I see their impact in my numbers. Probably I have touched a life somewhere by doing what I did. That motivates me to continue. I made an impact to someone's life when we found out they were positive and I reported. I will give a tally of everyone who came in the facility. It's like you have a role, it's not a main role, you are behind the scenes. That keeps me motivated. It's a good feel."–****Data officer facility I (outlier performer)***

> "*First, this profession is about people. When I fail to commit to them, I am killing one or two people there. I become motivated when I see them come again. If I do it in a negative way, there will be a lot of murmurs. It is out of passion as well."–****Data officer facility H (poor performer)***

Awareness of facility reporting performance also provides knowledge as to whether a facility is meeting reporting requirements for completeness and timeliness. Moreover, good reporting performance not only portrays a facility in good light but also for the data officer(s) in charge during a particular reporting period. Nonetheless, performance reports for facility completeness and timelines in DHIS2 were not utilized by respondents for purposes of identifying respective facility reporting performance. As such, as long as the report was submitted within the submission deadline and without questioning of the data, facility reporting completion and timeliness requirement was considered having being met. One of the responses on this is as stated:

> "*As long as I have sent it be 5th in terms of timeliness, I know I have performed well. And if they have not kept on calling me all the time to ask me about the data, then I know I have done well."–****Data officer facility B (best performer)***

As such, data officers often relied on the approaches used by various respective health facilities key administrators to convey feedback on facility and individual reporting performance. Some of the feedback channels include through WhatsApp groups, through in-charges, during data review meetings, and during performance appraisals. One of the responses on this by an interviewee is as stated:

"*Yes we always have a data review whereby you see your performance, they also project the way you keep reporting. If you are in green, it means your reports are always on time, if you are on yellow you try but sometimes you are late, if you are red, you are always late. But no major repercussion, just try to not submit the reports late or pressure will mount on you.*" ***Data officer facility K (outlier performer)***

Hence, awareness of performance also implicitly contributed to motivation in reporting especially when feedback is good. Another factor that could potentially contribute to motivation is increasing space in records offices in some of the facilities. A general informal observation revealed some of the office spaces to be quite small in some of the facilities with some of the respondents eluding that more space would be good for them.

Availability of standard operating procedures, training, and supervision

Among the facilities visited, those funded had existing Standard Operating Procedures (SOPs) that were developed by the supporting partners. An informal observation was made whereby some of the facilities had put copies of their SOPs on the walls, while others had theirs only stored in cabinets. The SOPs regarding data quality were examined in order to understand the procedures put in place. The SOPs for data quality laid down roles and responsibilities for data management, data quality assurance hence providing guidance to data officers. Furthermore, internal monthly data review meetings were carried out in these facilities in order to review the HIV indicator data quality. These meeting mostly comprised of nurses, clinicians, and data officer employed by the partners.

Training is a very salient component in reporting. Nonetheless it was provided mostly by supporting partners for their employees. Training of the MOH-731 tools were provided only once when the tools had been updated. Nonetheless, supportive supervisions from sub-county doubled up as training in some facilities. Some facilities reported having had sub-county supervisions at least once a year, while others were hesitant in admitting to having no visits from the sub-county. Some of the responses are as follows:

"For the county no, but for the program if they organize, they call us."–***Data officer facility L (outlier performer)***

"*That is the challenge. We rarely get trainings. We lack good training, just good training. We go for some data quality assessment after three months but that is not enough. Because things to do with information, they transform and keep changing."*–***Data officer facility H (poor performer)***

Another emerging issue on training was the dependability on availability of funding in order for trainings to take place as stated by some of the respondents:

Yea, now that the is no funding, that (training) is once in our dreams.–***Data officer facility B (average performer)***

It also depends on funding. When there was money we used to go monthly, then it changed to quarterly. That is training for all of us because these tools keep on changing. There is a time we went 1 week for training because of these new tools.–***Data officer facility A (best performer)***

Hence, factors contributing to frequency of training and supervision were attributed to availability of funding, and availability of sufficient human resources to conduct the trainings, especially supportive supervision, which relied on sub-county staff.

Discussion

This qualitative case study identified barriers and facilitators in HIV-indicator reporting that were linked to the RHIS reporting process and determinants of RHIS (technical, behavioral and organizational determinants). Facilities performing well and those performing poorly might have contextual differences that affect their performance, hence influencing the barriers or facilitators faced as posited in other studies [27,28].

This study revealed that whereas facilities may demonstrate differences in reporting performance, they are likely to face similar barriers and facilitators regardless of the contextual differences. For instance, it was assumed that facilities with an EMRs system were likely to have good reporting performance. This is because these facilities are required to have availability of trained personnel, appropriate infrastructure, adequate security, support and maintenance protocols, and accessible management support prior to EMRs system implementation [29]. Nonetheless, it emerged that all facilities that performed poorly in timeliness (outlier performance group) had EMRs system implementations. In addition, among the facilities that had high reporting performance in completeness and timelines (best performing group), only one had a functioning retrospective stand-alone EMRs system, whereas the rest had EMRs systems that were not in use due to lack of human resource. Therefore, EMRs systems implementation in facilities did not translate to good performance in reporting to DHIS2.

This can also be attributed to the lack of interoperability between the EMRs systems and DHIS2 to enable seamless data transmission. This lack of interoperability between EMRs systems and national aggregate systems remains a challenge in LMICs, leading to systems operating in silos [29,30]. As such, given that EMRs systems have been attributed as having the potential to dramatically reduce the data collection burden by automating the reporting process [30,31], this potential has not yet been realized as revealed in our findings. Nonetheless, utilization of EMRs system in reporting still contributed to easing the data collection burden, as data was retrieved from EMRs systems rather than multiple registers, which is time consuming. Furthermore, facilities that utilized EMRs systems reported to have obtained technical support and training when required through their supporting partners. This is an indication of a step towards the right direction, as compared to previous studies, which indicate challenges such as lack of IT support [11,32].

A number of issues were also identified, which contributed to facilities performing poorly in timeliness. These include issues such as, time-consuming efforts to correct data quality issues in reports, and submission of reports by hand to the sub-county office. As such, insufficient human resources, lack of availability of the DHIS2 during reporting, and slow internet connection at the sub-county level resulted to late entries of the reports submitted by hand, which then resulted to poor performance in timeliness among facilities. This then hinders the repercussion measures that had been put in place to deter late reporting. This is because reports maybe submitted on time to the sub-counties by facilities in order to avoid repercussions for late reporting, however, they are entered late in the system at the sub-county level.

Nonetheless, efforts have been made by sub-counties to offload the data entry burden by ensuring facilities that meet a predefined criterion have access rights to DHIS2. As such, despite some of the facilities having access to DHIS2 at the time of study, it emerged that this was not the case in previous years. Hence, it was assumed that the lack of access rights to DHIS2 might have also been a contributor to poor performance in timeliness.

Supportive supervisions and mentorship have been identified as good approaches that contribute to providing quality data for M&E [11,33]. Nonetheless, insufficient human resource at the sub-county limits frequent supportive supervision at the facility level. Moreover, although on job training was provided in facilities funded by supporting partners, facilities that did not receive external funding depended on trainings and supportive supervision provided by sub-county. Nonetheless, supervision was not conducted as frequently as expected based on respondents' perspectives. Only two facilities among those assessed were not funded by supporting partners and therefore lacked frequent on job training. In addition, staff rotations and staff transfer frustrated supportive supervision and mentorship efforts in the sub-county, both in facilities with external funding, and those without. This was because they brought about demand to retrain incoming staff, which proved to be tasking as reiterated by respondents.

Our findings also revealed time constraints echoed by facilities in all performance groups as a major concern during the reporting period. These time constraints were further aggravated by issues such as, reporting period falling on a weekend (which meant that a day or two were deducted from the reporting period) and data officers being assigned to more than one health facility, which led to some facilities being given less priority compared to others in reporting MOH-731. As such, dealing with more than one facility has the potential to heighten the chances of reporting burden and risk of hampering data quality. In addition findings from this study reveal parallel reporting of HIV-indicator by facilities funded by supporting partners, which continues to be a challenge facing M&E systems in LMICs [11,34,35]. Time constraints issues are further aggravated by documentation errors brought about by staff rotation and staff transfer of health workers involved in reporting, lack of understanding the indicators, and DHIS2 availability issues, which further slowed down the reporting process. Ledikwe et al. also reported similar issues related on data gaps as result of changes in staffing [11].

In order to ensure timeliness in reporting, a resource intensive approach would entail strengthening capacity in health facilities through rigorous staff training, strong internet connection in facilities with computers, ensuring EMRs systems lying dormant in facilities are utilized for reporting and ensuring many facilities are able to meet requirements needed to obtain access rights to DHIS2 in order to perform data entry for themselves. These will facilitate timely submission of good quality reports to DHIS2. Nonetheless, a less resource intensive approach will entail ensuring availability of sufficient skilled human resources at the sub-county level to perform data entry tasks and provide supportive supervision at facility level. Continuous training of health workers on reporting will also enhance data quality and fill the gap left through rotations and staff turnovers.

This study has identified facilitators and barriers in HIV reporting among facilities with various performances in completeness and timeliness in reporting HIV-indicators to DHIS2. Findings reveal that similar barriers and facilitators are shared across the different performances. Nonetheless, it emerged that skilled human resources involved in reporting, in combination with access to DHIS2 promote better performance of facilities reporting completeness and timeliness.

A limitation of this assessment was that there had been staff turnovers and rotations among data officers in some of the facilities, which made it difficult to provide conclusive descriptions of happenings of previous years. In addition, though we used cases in one county, we expect that the findings revealed in this assessment are transferrable in other counties and in LMICs in similar contexts. In addition, given that facility performance was not based on indicator data completeness and accuracy, future verification exercises are warranted for the various facility performance categories. Further still, there have been efforts by the MoH in availing training tools such as eLearning portals that provide courses for DHIS2 and HIV M&E, which can provide training for facility health workers. Nonetheless, we did not delve into usage of such portals, hence future studies can be conducted in order to assess the benefit of these courses in relation to performance in data quality.

Conclusion

The study identified barriers and facilitators linked to the RHIS process and determinants of RHIS, which include three interrelated factors, technical, organizational and behavioral. The findings demonstrated that whereas facilities may demonstrate different performances in completeness and timeliness in reporting, the barriers and facilitators that they face may be less different among them. It was expected that EMRs systems would improve reporting, and in order to realize their potential in reporting, there needs to be integration of DHIS2 and EMRs systems as posited in feasibility studies conducted [31]. This study could not attribute best performers to presence of EMRs systems. Nonetheless, future prospects to automate indicator reporting between EMRs systems and DHIS2 will pave way to determine whether best performing facilities is accelerated by use of EMRs system. Continuous evaluations have been advocated within health information systems literature. Therefore, continuous qualitative assessments are also necessary in order to determine improvements, as well as recurring of similar issues based on previous assessments. These assessments have also complemented other quantitative analyses related to this study [21].

Supporting information

S1 Appendix. Interview guide.
(DOCX)

Acknowledgments

We would like to thank all the respondents who took part in this qualitative analysis.

Author Contributions

Conceptualization: Milka B. Gesicho, Ankica Babic.

Formal analysis: Milka B. Gesicho.

Investigation: Milka B. Gesicho.

Methodology: Milka B. Gesicho, Ankica Babic.

Supervision: Ankica Babic.

Validation: Ankica Babic.

Writing – original draft: Milka B. Gesicho.

Writing – review & editing: Milka B. Gesicho.

References

1. UNAIDS. 90-90-90 An ambitious treatment target to help end the AIDS epidemic [Internet]. 2014 [cited 2018 Jan 5]. Available from: http://files.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2014/90-90-90_en.pdf.
2. Ekouevi DK, Karcher S, Coffie PA. Strengthening health systems through HIV monitoring and evaluation in Sub-Saharan Africa. Current Opinion in HIV and AIDS. 2011 p. 245–50. https://doi.org/10.1097/COH.0b013e3283479316 PMID: 21546830
3. Nash D, Elul B, Rabkin M, Tun M, Saito S, Becker M, et al. Strategies for more effective monitoring and evaluation systems in HIV programmatic scale-up in resource-limited settings: Implications for health

systems strengthening. J Acquir Immune Defic Syndr JAIDS. 2009; 52:58–62. https://doi.org/10.1097/QAI.0b013e3181bbcc45 PMID: 19858942

4. Saito S, Howard AA, Chege D, Ellman TM, Ahoua L, Elul B, et al. Monitoring quality at scale: implementing quality assurance in a diverse, multicountry HIV program. AIDS. 2015; 29 Suppl 2:S129–36.
5. Health Information Systems Program (HISP). DHIS 2 [Internet]. Available from: https://www.dhis2.org/.
6. Dehnavieh R, Haghdoost AA, 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. 2019 May 1; 48(2):62–75. https://doi.org/10.1177/1833358318777713 PMID: 29898604
7. Mpofu M, Semo B-W, Grignon J, Lebelonyane R, Ludick S, Matshediso E, et al. Strengthening monitoring and evaluation (M&E) and building sustainable health information systems in resource limited countries: lessons learned from an M&E task-shifting initiative in Botswana. BMC Public Health. 2014; 14.
8. Aqil A, Lippeveld T, Hozumi D. PRISM framework: A paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy Plan. 2009 May 1; 24(3):217–28. https://doi.org/10.1093/heapol/czp010 PMID: 19304786
9. Kiberu VM, Matovu JK, Makumbi F, Kyozira C, Mukooyo E, Wanyenze RK. Strengthening district-based health reporting through the district health management information software system: The Ugandan experience. BMC Med Inform Decis Mak. 2014 Dec 13; 14(1):40. https://doi.org/10.1186/1472-6947-14-40 PMID: 24886567
10. Kawonga M, Blaauw D, Fonn S. Aligning vertical interventions to health systems: a case study of the HIV monitoring and evaluation system in South Africa. Heal Res Policy Syst. 2012 Dec 26; 10(1):2. https://doi.org/10.1186/1478-4505-10-2 PMID: 22280794
11. Ledikwe JH, Grignon J, Lebelonyane R, Ludick S, Matshediso E, Sento BW, et al. Improving the quality of health information: a qualitative assessment of data management and reporting systems in Botswana. Heal Res Policy Syst. 2014 Dec 30; 12(1):7. https://doi.org/10.1186/1478-4505-12-7 PMID: 24479822
12. Glèlè Ahanhanzo Y, Ouedraogo LT, Kpozèhouen A, Coppieters Y, Makoutodé M, Wilmet-Dramaix M. Factors associated with data quality in the routine health information system of Benin. Arch Public Health. 2014; 72(1):25. https://doi.org/10.1186/2049-3258-72-25 PMID: 25114792
13. Makombe SD, Hochgesang M, Jahn A, Tweya H, Hedt B, Chuka S, et al. Assessing the quality of data aggregated by antiretroviral treatment clinics in Malawi. Vol. 86, Bulletin of the World Health Organization. World Health Organization; 2008. p. 310–4. https://doi.org/10.2471/blt.07.044685 PMID: 18438520
14. 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(4):1–14. https://doi.org/10.1371/journal.pone.0195362 PMID: 29668691
15. 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 Inform. 2016 Nov 1; 95:60–70. https://doi.org/10.1016/j.ijmedinf.2016.09.006 PMID: 27697233
16. Muga R., Kizito P., Mbayah M. GT. Overview of the Health System in Kenya. In: Kenya Service Provision Assessment Survey 2004.1999. p. 13–24.
17. World Health Organization (WHO). Primary health care systems (PRIMASYS): case study from Kenya, abridged version. Alliance Heal Policy Syst Res. 2017;1–16.
18. Miles MB, Huberman AM, Saldaña J. Qualitative data analysis: a methods sourcebook. Thousand Oaks, CA: Sage; 2013. 381 p.
19. Creswel JW. The Selection of a Research Approach. Res Des Qual Quant Mix methods approaches. 2008;3–22.
20. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K, et al. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Heal. 2016; 42(5):533–44.
21. Gesicho MB, Babic A, Were MC. K-means Clustering in Monitoring Facility Reporting of HIV Indicator data:Case of Kenya. In: Studies in Health Technology and Informatics. IOS Press; 2020. p. 143–6. https://doi.org/10.3233/SHTI200514 PMID: 32604621
22. Gesicho MB, Were MC, Babic A. Evaluating performance of health care facilities at meeting HIV-indicator reporting requirements in Kenya: an application of K-means clustering algorithm. BMC Med Inform Decis Mak. 2021; 21(1):6. https://doi.org/10.1186/s12911-020-01367-9 PMID: 33407380
23. Morse JM. Critical issues in qualitative research methods. Sage Publications; 1994. 23–46 p.

24. QSR International. NVivo Qualitative Data Analysis Software [Internet]. 2019. Available from: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home.
25. Creswell JW, Miller DL. Determining validity in qualitative inquiry. Theory Pract. 2000; 39(3):124–30.
26. Goodhue DL, Thompson RL. Task-Technology Fit and Individual Performance. MIS Q.1995; 19(2):213.
27. Joseph Wu T-S, Kagoli M, Kaasbøll JJ, Bjune GA. Integrated Disease Surveillance and Response (IDSR) in Malawi: Implementation gaps and challenges for timely alert. Uthman O, editor. PLoS One. 2018; 13(11):e0200858. https://doi.org/10.1371/journal.pone.0200858 PMID: 30496177
28. Bwesigye DA. Factors Associated With Incomplete Reporting Of HIV/AIDS By Uganda's Surveillance System. Heal Syst Policy Res. 2017; 04(01).
29. Kang' S, Puttkammer N, Wanyee S, Kimanga D, Madrano J, Muthee V, et al. A national standards-based assessment on functionality of electronic medical records systems used in Kenyan public-Sector health facilities. Int J Med Inform. 2017; 97:68–75. https://doi.org/10.1016/j.ijmedinf.2016.09.013 PMID: 27919397
30. Kariuki J, Manders E, Richards J, Oluoch T, Mutiso J. Automating Indicator Data Reporting from an EMR to Aggregate Data System Using OpenMRS and DHIS2. J Heal Informatics Africa. 2013; 1(1).
31. Kariuki JM, Manders E-J, Richards J, Oluoch T, Kimanga D, Wanyee S, et al. Automating indicator data reporting from health facility EMR to a national aggregate data system in Kenya: An Interoperability field-test using OpenMRS and DHIS2. Online J Public Health Inform. 2016; 8(2):e188. https://doi.org/10.5210/ojphi.v8i2.6722 PMID: 28149444
32. Mbondo M, Scherer J, Aluoch GO, Sundsmo A, Mwaura N. Organizational HIV monitoring and evaluation capacity rapid needs assessment: The case of Kenya. Pan Afr Med J. 2013; 14:1–7. https://doi.org/10.11604/pamj.2013.14.1.1092 PMID: 23503692
33. Ledikwe JH, Reason LL, Burnett SM, Busang L, Bodika S, Lebelonyane R, et al. Establishing a health information workforce: Innovation for low- and middle-income countries. Hum Resour Health. 2013 Jul 18; 11(1):35. https://doi.org/10.1186/1478-4491-11-35 PMID: 23866794
34. Souteyrand Y, Banda MA, Kaufman J, Perriëns JH, Yu D. Investment in HIV/AIDS programs: Does it help strengthen health systems in developing countries? Global Health. 2008; 4. https://doi.org/10.1186/1744-8603-4-8 PMID: 18796148
35. 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(Suppl 4):1–6. https://doi.org/10.7448/IAS.19.5.20847 PMID: 27443274

UNIVERSITAS BERGENSIS

uib.no

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

ISBN: 9788230858004 (print)
9788230841211 (PDF)