An Outcome Evaluation of the Health Information Management and Applied Epidemiology (HIMAE) course in South Africa

Alexa M Juarez

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Public Health

University of Washington

2019

Committee:

Nancy Puttkammer

Mary Lou Thompson

Program Authorized to Offer Degree:

Global Health

University of Washington

Abstract

An Outcome Evaluation of the Health Information Management and Applied Epidemiology (HIMAE) course in South Africa

Alexa M Juarez

Chair of the Supervisory Committee:
Nancy Puttkammer
Department of Global Health

Background:

Health Information Systems (HIS) can provide accurate and reliable data for patient monitoring, disease surveillance, program evaluation, and resource allocation at varying levels of a national health system. South Africa has been making poor progress on meeting targets for HIV epidemic control due to a lack of data capturers at the facility-level and poor data quality. Blended learning has been identified as one feasible option to address this gap. This paper evaluates the Health Information Management and Applied Epidemiology (HIMAE) course designed to train frontline healthcare workers in South Africa.

Methods:

The ten module HIMAE course was disseminated through a Demonstration Programme (DP) and Open Choice Programme (OCP). Participants had access to e-learning, USB, and workbook modalities. Pre- and post-test scores were used to evaluate change in knowledge over the first five modules of the course which was defined as course completion. Completer and non-completer interviews were conducted to assess participant attitudes towards course delivery and uptake and measure translation into daily practice. Descriptive statistics, t-tests, and linear regressions were calculated to determine differences within and across dissemination groups. Interviews were coded and analyzed for common themes based on findings from the quantitative analysis.

Results:

A total of 331 participants took part in the study. 39 participants were enrolled in the DP and 292 were enrolled in the OCP. DP participants were 3.9 (95% CI: 2.2-6.7, p <0.001) times more likely to complete the first five modules than those in the OCP. On average, DP participants' test-scores improved by 24.2 percentage-points and the OCP improved by 21.0 percentage-points. There were no significant differences in mean post-test scores across dissemination groups or mean score differences among cadres. The DP model increased motivation to complete the course but did not change the types of uptake barriers faced by participants.

Conclusion:

Overall, the course significantly improved participant knowledge in data literacy. Interview participants also indicated several ways they applied knowledge gained from the course into their daily practice. Based on these findings, the HIMAE course can provide an approach to remove barriers of poor data quality and large frontline health worker knowledge gaps as well as assist South Africa to make improved progress to meeting targets for HIV control.

List of Abbreviations

ART	Antiretroviral therapy
CDC	Centers for Disease Control and Prevention
CHC	Community Health Care
DHIS	District Health Information System
DP	Demonstration Programme
HIMAE	Health Information Management and Applied Epidemiology
HIS	Health Information System
HIV	Human Immunodeficiency Virus
IDI	In-depth Interviews
IRB	Institutional Review Board
I-TECH	International Training and Education Center for Health
KZN	KwaZulu-Natal
LMIC	Low- and Middle-Income Countries
LMS	Learning Management System
NDOH	National Department of Health
NGO	Non-governmental Organization
OCP	Open Choice Programme
PHC	Primary Health Care
RTC	Regional Training Center
USB	Universal Serial Bus

Introduction

Health Information Systems (HIS) can provide accurate and reliable data for patient monitoring, disease surveillance, routine reporting, program evaluation, and resource allocation at varying levels of a national health system.[1,2] HIS can also provide valuable information to assist countries in monitoring antiretroviral therapy (ART) adherence, human immunodeficiency virus (HIV) prevention intervention uptake, and 90-90-90 target achievements.[1] The 90-90-90 goals aim to diagnose 90% of all HIV-positive persons, provide antiretroviral therapy (ART) for 90% of those diagnosed, and achieve viral suppression for 90% of those treated by 2020.[1] However, many low- and middle-income countries (LMICs) lack HIS strong enough to provide reliable and accurate data for these purposes.[1,2] A major limitation for HIS in LMICs is the small number of trained health care workers to collect and use data.[2] In many cases, front-line health care workers, such as nurses and community health care workers, fill the role of data capturers or health information officers while lacking the appropriate data literacy skills.[1,2] This situation creates gaps and errors as data are collected, aggregated and reported from community and facility levels, to district, province, and national levels.[2]

In South Africa, the District Health Information System (DHIS) was introduced in 1996 to collect monthly facility-based data from primary health services in the public sector.[3] The monthly facility-based data from the DHIS was intended to promote the use of data for decision-making to help improve facility health programs and management.[3] Since then, DHIS has had relatively poor performance due to the lack of adequately trained front-line health care workers and the gap in understanding the importance of accurate data collection.[3,4] In a 2008 study, 10 rural clinics in the KwaZulu-Natal (KZN) province were assessed on data quality and literacy. Only 30% of facility health workers showed competency in HIS related tasks and 86% of workers could not demonstrate accurate interpretation of data findings.[3] South Africa has also been making poor progress on meeting targets for HIV epidemic control, specifically retaining patients on ART.[5] In 2019, Minister of Health Motsoledi called for an urgent response to this challenge, citing the need for initiatives to increase the number of data capturers at the facility-level and improve data quality to accurately track HIV patients.[5] Based on these findings, there is a clear knowledge gap in data literacy among facility health workers and a vital need to improve the quality and use of data collected in the public sector. Improvements in data quality and accuracy are major pathways forward in achieving the 90-90-90 targets and strengthening South Africa's HIS.

E-learning and other forms of self-paced blended learning strategies are being used as a cost-effective way to supplement health care worker education and fill knowledge gaps in the health care work force.[6] E-learning has been shown as an effective method to bridge the observed knowledge gap and to increase capacities for collecting quality data for HIS.[6] Blended learning, an approach mixing face-to-face classroom methods with e-learning, has been used to address limited technical demands in rural areas and low-resource settings.[6] Many investments in strengthening HIS in LMICs have been funded through HIV programs. As a response to this need, the International Training and Education Center for Health (I-TECH) developed a training course on Health Information Management and Applied Epidemiology (HIMAE) in conjunction with the U.S. Centers for Disease Control and Prevention (CDC), the South African National

Department of Health (NDOH), and the University of Witwatersrand. The course is comprised of a series of ten modules that are completed at the user's own pace. The course is disseminated through online e-learning, Universal Serial Bus (USB), and paper workbook modalities. The topics of the modules range from basic numeracy to basic applied epidemiology to data reporting. The HIMAE course was designed to increase data skills to improve data collection performance and decrease the gap of data literacy in front-line health care workers.

In 2017, a pilot study was completed for the HIMAE training course in the Free State, KZN, and Mpumalanga provinces of South Africa. The purpose of the pilot study was to assess and improve the content of the workbook and e-learning resources. Results from the pilot study showed on average a 15% increase in health worker knowledge across HIMAE course modules and that course participants believed that module content was relevant and applicable to their daily practice.[7] Participants also expressed strong preference for the e-learning modules over the paper workbook modality.[7] In 2018, the course expanded across South Africa and was disseminated through two different program approaches, with both approaches again using the e-learning, Universal Serial Bus (USB), and paper workbook modalities. The Demonstration Programme (DP), was designed as a blended learning approach. This model provided more one-on-one support to participants by deploying "champions" who would have regular in-person engagement with course participants in addition to the standard course self-guided learning material. District health authorities identified champions including district personnel and NGO partner staff members. I-TECH provided a 2-day orientation to champions, and each champion was assigned 3-4 health facilities where front-line health care workers registered to complete the course. Participants were provided 1 set of workbooks per health facility, individual USB drives with the course e-learning modules to allow for offline use, logins to the learning management system (LMS) for the course, and access to a WhatsApp group chat. This DP was only disseminated in one Free State district, Lejweleputswa District. The Open Choice Programme (OCP) was a less intensive dissemination model implemented in conjunction with Regional Training Centers (RTCs). RTCs held brief orientation sessions to recruit and register participants into the course. I-TECH and RTCs provided participants with access to course modules via the LMS and basic technical assistance on login and use of the LMS. This model did not include a champion for participants and was offered in Gauteng, KZN, and Limpopo provinces as well as the remaining Free State districts.

The main objective for the evaluation of the HIMAE course expansion, and the focus of this thesis, was to assess knowledge gains in each dissemination model and participant attitudes towards course uptake, modalities, and translation to daily practice. Facility-level data quality evaluations and knowledge retention assessments were also completed 4 to 6 months after course completion. Facility data quality evaluations were intended to assess the effects of the course on completion and accuracy of patient charting and indicator reporting. Course post-tests were repeated 3-6 months after participation in the course, to determine if knowledge and skills were retained following training. Results of the facility data quality evaluations and repeat post-tests are not reported in this thesis.

Methods

Evaluation Design

Comparisons within and across dissemination groups (DP and OCP) were undertaken through a mixed-methods evaluation. A pre-post evaluation design was used to assess the knowledge gained from the course within and across each group. Module completion was defined as completion of both the module pre- and post-tests. Completion of the first five modules was considered to constitute course completion, as these modules represent the core foundation for practical HIMAE skills. In-person qualitative interviews were conducted for completers and non-completers in the DP and OCP to assess participant experience, barriers in course delivery and uptake, and translation of skills into daily practice. WhatsApp group discussions were also evaluated as additional information on facilitators and barriers in course uptake in the DP group.

Population and Sampling

The reference population from which DP and OCP dissemination groups were sampled was the population of front-line healthcare workers and district health office personnel in the four selected South African provinces. The DP was selected solely from one district in the Free State province. The Lejweleputswa district for the DP was identified by the Free State Provincial Health Department based upon interest in participating in the DP, convenience and logistics, and availability of "champions" to collaborate in the carrying out of the DP. District health authorities then purposefully selected 15 target health facilities reflecting a range of levels and location: 85% were Primary Health Care (PHC) or Community Health Care (CHC) facilities and 15% were hospitals. This stratified sample of DP facilities was intended to reflect the distribution of different types of facilities seen nationally. Within these health facilities, all health workers involved in HIV care and treatment were invited by champions to participate in the course.

The second sample of OCP participants was selected from KZN, Limpopo, and Gauteng provinces as well as other districts in Free State. Course orientations were held at RTCs to recruit and register OCP participants. RTC trainers invited health workers from various health facilities and district health personnel to participate in the course orientation based upon a fit between course content and known learning needs. Course participants in both samples represented a variety of health worker cadres including: enrolled and professional nurses, data capturers, facility managers, pharmacists, business administrators, and district health data managers.

Data Collection

Quantitative Data

Demographic information for all participants was collected using the registration and uptake forms at participant orientations. The Moodle-based learning management system (LMS) also collected demographic data from participants as well as a series of optional open-ended questions on participant's reasons for taking the course. Data from these forms and the LMS included participant's name, province, and cadre. Participants completed pre-tests and post-tests online and scores were recorded on the LMS. LMS data for pre- and post-test scores were

logged on a weekly basis to monitor participant progress. Test score data was collected for all registered participants in both dissemination groups.

Qualitative Data

In-depth interviews (IDIs) were conducted to assess course uptake, facilitators and barriers in the course delivery, and the translation of course content to practice. IDIs were held at health facilities for participants in the DP group who had completed at least 5 modules. Participants were purposively sampled based on the range of cadres, facilities, and total number of completed modules. Interviews lasted 30 to 45 minutes for each participant and followed semi-structured interview guides (Appendix)**.** Interviews were recorded and later transcribed for data analysis.

Non-completer telephone interviews were conducted to assess participant experience as well as facilitators and barriers for course uptake. A sample of purposively selected individuals from DP and OCP groups was selected from participants who registered for the course and completed fewer than 3 modules. The sample was selected to have a balance of types of cadres, modules completed, and DP and OCP representation. Interviews lasted 10 to 15 minutes and followed a short interview questionnaire. Interviews were recorded and later transcribed for data analysis.

There was a total of 23 IDIs, 15 completer and 8 non-completer interviews, conducted in this evaluation. All completer interviews and 1 of the 8 non-completer interviews were conducted with participants in the DP. The research team reached out to 20 non-completers to request an interview, but only 8 participants agreed to be interviewed. In total, there were 10 data capturers, 8 nurses, and 2 participants from other cadres interviewed.

A WhatsApp group discussion was also collected to assess common challenges faced by participants in the DP group. The discussion was open to all DP participants and weekly discussion starters were given to start discussions on challenges with the course. The group discussion was exported for analysis and all identifying factors were removed to maintain participant confidentiality.

Data Analysis

Quantitative Analysis

Summaries of the proportion of participants who completed the first five modules of the course were calculated by dissemination group, cadre, and province. Cadres were separated into three main groups: Data Capturer, Nurse, and Other. The Other category was comprised of business administrators, medical students, and other healthcare professions. Module pre- and post-test scores for the DP and OCP completers were summarized by descriptive statistics and graphical dot plots. Combined scores for pre- were generated by averaging the scores from the pre-tests of the first five modules. Combined post-test scores were generated using the same method.

Assessment of knowledge gained from the course was made within and across dissemination groups. The primary outcome for within group comparisons was the average pre- and post-test score of each individual over the first five modules. Paired t-tests were conducted within DP and OCP groups to test the null hypothesis of no difference between mean pre- and post-test

scores. In a secondary analysis, with the difference in pre- vs post- combined score as outcome, linear regression with robust standard errors was used to compare mean pre- versus post- scores within groups by cadre. These analyses treated each participant as an independent observation, with an assumption of no correlation in test scores between participants from the same health facility. A sensitivity analysis was conducted in the DP to explore the effect of treating participants from the same facilities as independent rather than clustered observations within our data. In this analysis, 4 participants from clustered facilities were randomly selected and removed for the sensitivity analysis.

For across-group comparisons, the primary outcome was the combined post-test score over the first five modules. Linear regression with robust standard errors was used to model post-test score as a function of group (DP or OCP), adjusting for pre-test score. In a secondary analysis, cadre was included in the linear regression to adjust for possible confounding by cadre.

The quantitative analysis for DP and OCP dissemination groups was conducted using R Studio. All formal hypothesis tests were conducted at the 5% alpha level.

Qualitative Analysis

The main objective with the qualitative analysis was to explore questions and results generated from the quantitative analysis. Questions were developed to explain the results observed in the quantitative analysis of this evaluation and to explore participant experiences with the course. An a priori thematic codebook was created by the coding team and applied to all transcripts. Coding of the transcripts was an iterative process, and new codes were created as new concepts emerged from the transcripts. Two secondary coders from the I-TECH South Africa research team assisted in coding transcripts to apply context to the coding and analysis. A system of memos was used to identify emerging themes and to suggest new codes to include in the codebook. Data were organized according to main themes that emerged and were used to describe participant experience during the course and translation of knowledge into practice.

Qualitative and quantitative results were incorporated and triangulated to identify synergies in the overall evaluation.

Ethical Considerations

Participation in this evaluation was voluntary and all participants provided either written or verbal informed consent. It was explained to each participant that data collected for the study would remain confidential and names or other identifying factors would be removed from the analysis and dissemination of results. Participants were also able to withdraw from the study at any point without adverse consequences. Routine data on course participation as captured by the LMS was exempted from informed consent for OCP. Institutional Review Boards (IRBs) for the University of Washington, CDC, and Human Sciences Research Council of South Africa reviewed and approved the protocol as a non-research program evaluation.

Results

Participant Characteristics

Table 1 presents the characteristics of participants and their completion of modules 1 through 5. A table of the number and percentage of completion for modules 1 through 10 is located in the Appendix. A total of 331 participants took part in the HIMAE course and 41 completed a minimum of 5 modules, which we define as course completion. Among the total participants, 39 were in the DP and 292 were in the OCP. Participants had on average 5 months to complete the course. Of the 39 participants in the DP, 14 (35.9%) participants completed at least 5 modules while 27 (9.3%) of the 292 participants achieved completion in the OCP. DP participants had a 3.9 (95% CI: 2.2-6.7, p <0.001) times greater chance of completing five modules compared to OCP participants. Limpopo and Gauteng had the highest number of OCP participants (121). Free State and Limpopo had the highest percentage, 27.1% and 14.0%, of participants complete the course. In both the DP and OCP, nurses were the largest cadre enrolling in the course. However, the Other cadres and Data Capturers were more likely to complete the course.

Table 1. Characteristics of HIMAE Course Participants

Characteristic	Total		Completed*			
	N		N		(%)	
All Participants	331		41		12.3	
Group						
DP	39		14		35.9	
OCP	292		27		9.3	
Province						
Free State	59		16		27.1	
KZN	30		1		3.3	
Limpopo	121		17		14.0	
Gauteng	121		7		5.8	
Cadre	OCP	DP	OCP	DP	OCP	DP
Nurse	133	18	11	6	8.3	33.3
Data Capturer	26	14	4	6	15.4	42.9
Other	121	4	12	2	9.9	50.0
Missing	12	3	0	0	0.0	0.0

*Completed is considered completing at least 5 modules

Participant Knowledge Gains Among Participants Completing at least 5 Modules

Within Groups

Mean pre- and post-test scores and score difference over the first five modules for the DP and OCP are presented in Table 2. The results in this table show the average (percentage) scores for the pre- and post-test as well as the difference in percentage points comparing the post- to the pre-test. In the DP, 4 facilities had 2 participants taking the course. Results from the sensitivity analysis were consistent with the primary analysis.

Table 2. Descriptive Statistics of Pre- and Post-test Scores for DP and OCP

Group	N	Scores	Mean	SD	Min	Max
DP	14	Pre-Test	56.9%	19.5%	19.2%	97.1%
		Post-Test	81.1%	10.9%	46.7%	100.0%
		Difference	24.2%	19.5%	-9.4%	65.0%
OCP	27	Pre-Test	61.8%	16.5%	24.8%	95.0%
		Post-Test	82.8%	12.0%	38.8%	100.0%
		Difference	21.0%	15.1%	-9.0%	64.6%
DP Sensitivity Analysis*	10	Pre-Test	54.8%	10.1%	35.6%	66.7%
		Post-Test	81.1%	9.2%	61.4%	93.2%
		Difference	26.2%	9.2%	14.8%	42.4%

* 4 DP participants from clustered facilities were randomly selected and removed for the sensitivity analysis. A random number generator assigned numbers to both participants in a clustered facility and the 4 participants with the highest within facility values were removed. Descriptive statistics and a paired t-test were conducted with the remaining participants.

There were statistically significant differences between mean pre- and post-test scores in both the DP and OCP groups. The DP participants had an average 24.2-percentage point gain ($p<0.001$, 95% CI: 18%-30%) and a 26.2-percentage point gain ($p<0.001$, 95%CI: 19.7%-32.8%) in the sensitivity analysis. The OCP was slightly lower at an average 21.0-percentage point gain ($p<0.001$, 95% CI:15%-26%).

Within the DP, there were no statistically significant score difference between participant cadres as seen in Table 3. DP data capturers had an average post- vs pre- score difference of 26.3-percentage points. Nurses' average score difference was 6.8-percentage points below data capturers while the other cadres had an average difference that was 5.2-percentage higher. Other cadres in the DP had an average 12.0-percentage point increase in their score difference when compared to nurses. he OCP also did not show any significant score differences between cadres within the group. The OCP data capturers had an average post- vs pre- score difference of 27.3-percentage points. Nurses' and other cadres' average score differences were respectively 9.2 and 5.9-percentage points less than data capturers. Nurses had an average 3.3-percentage point decrease in their score difference when compared to other cadres.

Table 3. Regression results of pre- to post-test percent-point differences, adjusting for cadre

Group	Cadre	Percent-Point Mean Difference Estimate	95% CI	P-value
DP	Data Capturer*	*Reference*		
	Nurse	-6.8	[-17.8, 4.2]	0.26
	Other	5.2	[-5.4, 15.8]	0.35
OCP	Data Capturer**	*Reference*		
	Nurse	-9.2	[-21.1, 2.9]	0.15
	Other	-5.9	[-16.5, 4.7]	0.29

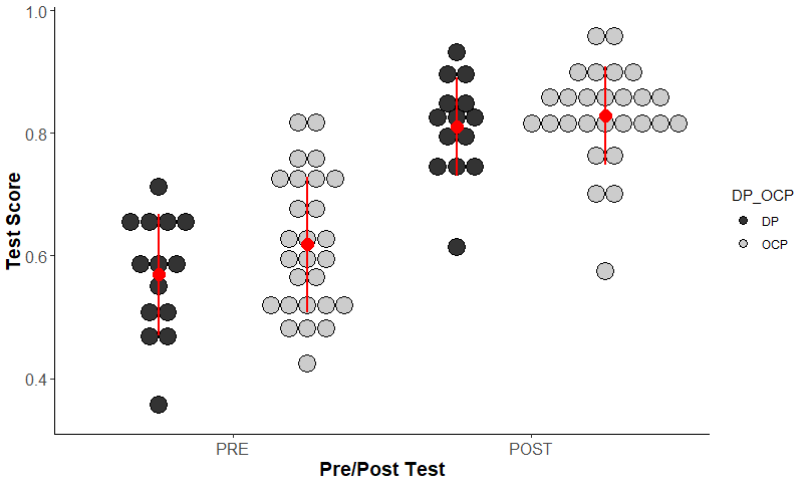
**Estimated mean difference for DP Data Capturer (reference group) was 26.3 (95% CI: 18.7, 33.9). Compared to Other, mean difference for Nurses was 12.0 percentage-points lower (95% CI: -24.0, 0.12, p=0.07).*

***Estimated mean difference for OCP Data Capturer (reference group) was 27.3 (95% CI: 17.5, 37.1). Compared to Other, mean difference for Nurses was 3.3 percentage-points lower (95% CI: -11.5, 5.0, p=0.44).*

Across Groups

In Figure 1, the graphical dot plot displays the mean pre- and post-test scores for each individual participant in the DP and OCP. The red dot highlights the overall mean for each group and the red line represents one standard deviation on either side of the mean. This graph illustrates that the distributions of pre-test scores are comparable between the DP and OCP and that a linear regression comparison of post-test scores across dissemination groups, adjusted for pre-test scores, was reasonable.

Figure 1. Dot Plot of Pre- and Post-Test Scores in DP and OCP



Results from the linear regression showed that average post-test scores were not significantly different across dissemination groups, after adjustment for pre-test score and cadre. As seen in Table 4, on average the OCP scored 0.96-percentage points lower than the DP. Also, there was no significant difference in mean post-test scores by cadre in the adjusted model. Controlling for dissemination group and pre-test score, nurses and other cadres' average post-test scores were 1.7 and 4.4-percentage points higher than data capturers. Nurses scored on average 2.6-percentage points lower on post-test score percentages when compared to other cadres with the same pre-test score.

Table 4. Regression results of post-test scores, adjusting for pre-test scores and cadres

	Post-Score Percentage Estimate	**95% CI**	**P-value**
Dissemination Group			
DP*	*Reference*		
OCP	-0.96	[-6.1, 4.1]	0.72
Cadre			
Data Capturer*	*Reference*		
Nurse**	1.7	[-5.4, 8.8]	0.64
Other	4.4	[-2.6, 11.2]	0.23

*Estimated mean post-score percentage for the reference group (DP Data capturers) was 64.0 (95% CI: 52.6, 75.4).
**Estimated mean post-score percentage for Nurses was 2.6 percentage-points (95% CI: -6.8, 1.5, p=0.23) lower when compared to Other in the same dissemination group.

Completer and Non-Completer In-Depth Interviews

Motivation and Challenges to Taking the HIMAE Course

Participants were eager to register for the course to help with work-related activities. Most data capturers were motivated to take the HIMAE course to refresh their data quality and collection skills. Other cadres, like facility managers and nurses, were motivated to learn more about data management and improving data quality at their facility.

> *"[The course] will make you a better manager so that you are accountable and responsible for your data and then you will be able to use it for the correct purpose moving forward whether it's a program or it's a health project or whatever you will be able to do informed decision if you have that information"*

The non-completer interviews illuminated some of the barriers participants in both OCP and DP faced in completing the course modules, which explained the low completion rates observed in the quantitative analysis. A main barrier for participants in the DP and OCP was logging into the course website. Participants frequently relied on the I-TECH support team to resolve login issues, primarily resetting passwords for participants. Difficulties logging in decreased motivation to complete the course and deterred participants from using the e-learning modality.

Due to limited bandwidth at work, most course participants studied at home or outside the facility. Many were too busy or overburdened while at work to study or work on pre- or post-tests. It was common for participants to study during a lunch break or would wait until the weekend to complete a course module. Very few participants studied during work hours or used work computers and internet to work on course materials.

> *"Most of the time it was after work and I think it was somewhere also during the weekends. Because to come and do it here won't help me, because there's nothing that can help me, except… The only time that you also get to learn and to go into your books is at home, because there's no time here."*

Computer access was also limited for most participants because they either did not own a home computer or did not have access to a work computer. Some of the other challenges included limited computer access or participants not having internet access to take a pre- and post-test. Some participants used their phones to access the course material and to take pre- and post-tests. However, phone data was not consistently available. Most participants would use the workbook at home and then log on to the course website to take the pre- and post-test when a computer, Wi-Fi, or internet data was accessible.

Some participants were intimidated by the e-learning modality of the course and were worried about not having enough computer literacy to complete the modules. Participants mentioned that they would prefer a more traditional classroom format that was not dependent on computer literacy. For completing course modules, these participants primarily depended on the workbooks to access course modules.

Demonstration Program Model

While DP participants were all in the WhatsApp group chat, they seemed to engage in smaller chat groups that they created with peers. The main WhatsApp group chat was primarily used for resolving login issues. Since many participants were at different places in the course, many were nervous about bringing up questions that seemed to hold the class back.

> *"So, when people were discussion in there that they were like at module this and that, and I hadn't done those ones. So, I was feeling, feared that if I asked them, I would be taking them back, you see, because they were already ahead."*

Contact sessions within the DP were conducted specifically where participants were located close to each other. Most participants were comfortable sending personal WhatsApp messages to their champion or asking for help one-on-one. Having an active small group of peers or an active-supportive champion played a large role in motivating participants to seek help and complete course modules.

> *"She even went to the extent of printing the questionnaire for me, because I told her I was having trouble going in, because I had a problem with my cell phone. She even brought me to questionnaire so that I could do the answers of the questionnaire."*

However, some participants in the DP were unaware that they had access to a champion and were hesitant to ask for help on the WhatsApp group. These participants commented that they felt alone and were unmotivated to complete modules when the content was more difficult. The experience from these DP participants were similar to experiences seen in the OCP.

Gains in Knowledge and Skills

Modules 1 through 5 seemed to be easy for participants but most participants regardless of cadre had challenges with the basic numeracy modules. Basic numeracy modules were modules 2 and 4, which could also explain for low completion rates among participants. Many participants enjoyed learning about generating graphs and using Excel. For most participants, this was a new skill.

Regardless of how many modules completed, participants did report learning about data quality on a general level and the HIMAE course was a good refresher to maintain skills related to data literacy. For example, data capturers primarily used this course as a refresher on basic numeracy concepts that had been forgotten. For non-completers, despite not having ample time to complete the course, there was felt to be value in doing the course even if it was only a few modules. Most participants considered the data quality skills gained as a starting point to improve their work.

Many participants also learned skills related to professional development. Participants noted that the course modules provided insight on improving teamwork as well as providing guidance on report and graph generation. Overall, many participants noted that a team's prioritization of data quality is vital for data reporting and accuracy to improve at the facility-level.

"Team work is important, they even asked what changed me because I talk a lot and at meetings I am now focused. You see? It sharpened me, I forgot which module it is but sharpened me a lot."

Translation into Practice

Participants who completed the course gained confidence in their data knowledge and skills. They were motivated to implement the data quality skills they had learned from the course in their own work through reports and collecting data. They were also more motivated to create changes on their own facility teams to improve data collection practices. Data capturers self-reported that they were more scrupulous in recording patient information and frequently verified data that seemed implausible with nurses in the facility. Data capturers also used the knowledge gained in the numeracy modules to verify denominators and numerators reported in the monthly district facility reports.

"I feel proud…being part of the HIMAE course because it has taught us a lot of things, especially understanding of data, errors and discrepancies. As time goes on you have a clear picture and will be able to point out mistakes before you capture that thing then you able to go back and talk to the nurse for further explanation."

"It was just after finishing this course and I was sitting with my manager saying hey I am done with that course and stuff and he was saying but then what do we do with the knowledge…I started paying more attention to the folders of the patients, more especially for HIV and TB."

Nurses in general had an easier time implementing course knowledge in their daily practices. They had more support to make data quality improvements within their team and within their daily work. There was also a sense of improving team work and having a motivating goal to improve data quality as a team. Many nurses relayed information from the course to the facility teams to help improve data collection practices as well as improving the data presented on reports, like adding more descriptive graphs.

"Where I'm working normally, we let the data clerks to do it. I just compile it and they do everything. So, some of the things I was not really clear about them, but after doing the course I realized that what I give to the data clerk is what she's going to record. So, it's very important for me to ensure that the information, everything that I record, is correct and accurate."

Work Place Environment and Support for Data Quality

Computer access was also limited for most participants because they either did not own a home computer or did not have access to a work computer. Some of the other work-related challenges included limited computer access or participants not having internet access at work to take a pre- and post-test. Some participants used their phones to access the course material and to take pre- and post-tests. However, phone data was not consistently available. Most participants would use the workbook at home and then log on to the course website to take the pre- and post-test when a computer, Wi-Fi, or internet data was accessible.

In terms of support, the data capturers reported being undermined in their work. Most facilities allocate more clinical roles to data capturers where there are staff shortages. Data capturers and nurses also reported that facility managers do not understand the importance of data management. In most facilities, managers would not prioritize facility reports or accurate data collection at the facility. In general, data capturers had obstacles applying newly acquired skills to their practice due to managers or other staff members not taking the course and not prioritizing data quality.

> *"I'm pressing them to happen. And I'm trying my very best at my facility to be the best when it comes to the importance of data, like the quality: giving the good quality data. I'm trying my very best, even though I'm alone. Like, I'm doing this alone."*

Application of knowledge in general was also hindered when the facility manager has not taken course as well as poor staff attitude about data quality. Facility managers and other leadership were seen as crucial in motivating translation of knowledge into practice and improving data quality.

> *"We are always told about quality…So let's try and make sure that we improve ourselves because we improve on daily basis although there will be challenges along the road but improvement I think its… even my manager wants correct things she does not like shoddy work…She is the one who taught me that quality comes first before anything else."*

Discussion

The HIMAE training course was implemented as a tool to improve data literacy in frontline healthcare workers. This study quantitatively and qualitatively assessed knowledge gains and participant experiences within and across dissemination groups. Within the DP and OCP, participants demonstrated meaningful and statistically significant percentage point gains from pre- to post-tests about basic numeracy, health information and data systems, and using data for quality improvement. While not all participants completed the course, participants in completer interviews expressed motivation to incorporate data quality knowledge and skills into their daily practice.

The evaluation also assessed participants' intentions to translate knowledge into their daily practice from the completer and non-completer interviews. Results demonstrated that the course addressed the high demand for a data quality-related course. Interviewed participants were eager to take the course to improve work practices related to data quality. They were also motivated to apply their knowledge in their daily practice and share skills they learned with their team to improve team data quality practices. The course was highly accepted by participants and there was a high motivation among completers and non-completers to register in the course. After completing the course, participants were also eager share the course with their peers and superiors within their facility or passed along course-related information. Results from the evaluation illustrate a demand satisfied from frontline healthcare workers taking the course and also a high need for data-related courses.

Another strength of this evaluation was the assessment of a blended learning tool to improve data literacy. One systematic review identified 49 studies of blending learning courses in LMICs.

However, a majority (70%) of these studies were focused on blended learning mechanisms to enhance physician's clinical training.[6] Few studies evaluated blended learning courses for the purpose of improving data quality at a facility or using a blended-learning intervention to fill gaps in frontline health worker data literacy.[6,10] From this evaluation, results illustrated that, even with a blending learning approach, the DP participants faced similar barriers seen in the OCP. The lack of active champions and WhatsApp group participation could have contributed to this as DP participants had varied exposure to these resources. Participants also felt intimidated by the e-learning components due to low computer literacy and emphasized a need for more face-to-face contact sessions. Results from this study suggest that a higher level of champion and peer support is needed in a blended learning model. For example, more effort and leadership engagement may be required to form learning groups at the facility-level to reinforce in-person group learning. Overall, the evaluation results will contribute to the growing research done on blended learning and increase the importance of blending learning initiatives as a method to address data literacy gaps in frontline healthcare workers.

There were implementation challenges with fidelity to the intended DP model. Contrary to the model, some participants in the DP had limited champion interactions or peer involvement. From participant interviews, a strong motivating factor for module completion was having support from either peers or champions. This was especially seen when completing modules related to numeracy which were perceived as the most challenging content. The low-fidelity to the DP model and lack of champion activity could be another reason for not seeing practically meaningful differences between post-test scores across the DP and OCP. Lack of peer or champion interaction has also been viewed as a e-learning challenge from other evaluations on e-learning programs for health care workers and the HIMAE course pilot study.[7,8,9] In future disseminations, a larger emphasis on champion engagement and peer support should be implemented to improve motivation and overall completion of the HIMAE course.

This study had a few limitations. A high non-completer rate resulted in a small sample size in the DP and OCP for the evaluation, which was due to low course uptake after orientation. The results found in this evaluation may not be generalizable due to the high non-completer rate. The low percentage of completers also decreased the likelihood of detecting score differences across dissemination groups and cadres within groups. However, not all differences in test scores were practically meaningful, while the mean pre- vs post- score difference within groups was large, the mean across group difference was small. It is possible that the positive knowledge gains observed reflected that course completers were highly motivated learners, and that these results would have been diluted had more participants completed the course. Alternatively, it is possible that non-completers had more to learn and might have had greater knowledge gains. It is impossible to definitely know the effect of non-completion on the estimate of effectiveness of the course in building knowledge and data literacy. A second important limitation was that an I-TECH researcher conducted all interviews for course completers and non-completers. This could have created social desirability bias within their responses as participants may have associated the interviewer with the course. However, at the beginning of each interview, interviewers strongly advised participants to be honest and share their feedback of the course in an attempt to minimize this bias. The evaluation lacked data collection to track champion activities in support of learners and their mentorship experience. This information

would have been helpful in making improvements to blending learning implementation and champion training.

Despite these limitations, findings from this evaluation suggest that the HIMAE course can increase data literacy in healthcare frontline workers. Further analysis is being conducted to assess whether knowledge gains are retained over time following completion of the course, and to assess the effectiveness of the course on improved data quality at the facility-level. The course is also being repackaged to improve participant uptake and increase course completion. The course will be separated into basic, intermediate, and advanced levels, with each segment broken into approximately 10 hours of learning material. Blended learning implementation modalities which provide more peer and champion support for participants, are also being explored.

The HIMAE course integrated with peer or champion support can assist in improving data literacy as well as increasing data quality in South Africa's DHIS. Data from the DHIS is a vital component for South Africa to track healthcare planning and evaluating healthcare targets like the 90-90-90 goals.[10] This evaluation demonstrated that the HIMAE course can remove barriers of poor data quality, reduce large knowledge gaps among frontline health worker, as well as assist South Africa to make improved progress to meeting targets for HIV epidemic control.

References

1. AbouZahr C, Boerma T. Health information systems: the foundations of public health. Policy and Practice: Bulletin of the World Health Organization. 2005; 83(8).

2. World Health Organization (WHO). Framework and standards for country health information systems. Geneva: Health Metrics Network; 2008.

3. Garrib A, Stoops N, McKenzie A, et al. An Evaluation of the District Health Information System in Rural South Africa. South African Medical Journal. 2008; 98:7.

4. Nicol E, Bradshaw D, Phillips T, Dudley L. Human Factors Affecting the Quality of Routinely Collected Data in South Africa. Medinfo. 2013; 192: 788-792. doi:10.3233/978-1-61499-289-9-788.

5. Motsoaledi A, Scaling Up South Africa's Response to Meet the 90-90-90 Targets. Ministry of Health Republic of South Africa Memo. 2019; 1:3.

6. Frehywot S, Vovides Y, Zohray T, et al. E-learning in medical education in resource constrained low- and middle-income countries. Human Resources for Health. 2013; 11:4.

7. Cowman A. Evaluation of a Distance-Based Training Program on Health Information Management and Applied Epidemiology (HIMAE) in South Africa. University of Washington ResearchWorks Archive, 2017.

8. George PP, Papachristou N, Belisario JM, et al. Online eLearning for undergraduates in health professions: A systematic review of the impact on knowledge, skills, attitudes and satisfaction. Journal of Global Health. 2014; 4:1.

9. Reisach U and Weilemann M. Organisational aspects and benchmarking of e-learning initiatives: a case study with South African community health workers. Global Health Promotion. 2016; 23(2): 57-66.

10. Mphatswe W, Mate KS, Bennett B, et al. Improving public health information: a data quality intervention in KwaZulu-Natal, South Africa. *Bull World Health Organ*. 2012;90(3):176–182. doi:10.2471/BLT.11.092759

Appendix I – Module-by-Module Completion

Table 5. Module-by-Module Completion Rate for DP and OCP

	Participants Completed (%)	
	DP (n=39)	**OCP (n=292)**
Module 1	22 (56.4)	64 (21.9)
Module 2	21 (53.8)	52 (17.8)
Module 3	20 (51.3)	40 (13.7)
Module 4	18 (46.2)	36 (12.3)
Module 5*	16 (41.0)	30 (10.3)
Module 6	17 (43.6)	23 (7.9)
Module 7	14 (35.9)	20 (6.8)
Module 8	15 (38.5)	24 (8.2)
Module 9	13 (33.3)	18 (6.2)
Module 10	12 (30.8)	17 (5.8)

*2 participants in the DP and 3 in the OCP completed module 5 did not complete the entire first 5 modules and therefore, were not counted as a completer.

Appendix II - Knowledge Translation Interview Topic Guide

Semi-Structured Interview Questions

Participants with Strong Completion

1) Completion of HIMAE Course

Some health workers have signed up for the HIMAE course but not managed to complete most or all of the modules. You were able to do this. What enabled you to do this? Possible probes:

- *Feedback on orientation sessions*
- *Can you describe what motivated you to complete the course modules? Probe about:*
    - *Personal circumstances*
    - *Time availability*
    - *Access to computers or tablets. Which mode they used most? How did they use the USBs, workbooks, eLearning, pre and post post-tests?*
    - *Work set-up, including management support*

2) Role of Blended Learning and Support

- *What was the role of HIMAE champions (what did they do, how were they helpful or unhelpful?)*
- *What was the role of peer learners in your site?*
- *Role of WhatsApp?*

3) Overall Opinion of Training

- *What modules or parts of the HIMAE course did you enjoy **most**? Why?*
- *What modules or parts of the HIMAE course did you enjoy **least**? Why?*
- *How do you feel about the knowledge and skills covered in the HIMAE course?*
    - *What are the main areas in which you believe you have gained knowledge and/or skills?*

    - *What are areas where you would have liked to gain knowledge or skills, but did not?*
- *To what extent was the training effective or ineffective in teaching you things that you can use in your job?*
    - *Which topics do you think were well-taught in terms of applying the knowledge in your work?*
        - *Which were not?*
    - *Which topics do you think were easy to learn?*
        - *Which were too complex to fully learn during the training?*

4) Ability to Apply Knowledge

- *To what extent are you able to put your new knowledge and skills into practice?*
    - *If not, why not?*
    - *What barriers did you face in applying your knowledge and skills at your home facility?*
- *In what ways are you applying your knowledge and skills from this course in practice at your home site?*
    - *Please share specific examples of times when you used learning from the course in your daily work.*
    - *Can you explain how your reporting of indicators has changed?*

- *Did you do anything in your home facility to improve data quality? Can you explain what you have done?*
    - *Data collection?*
    - *Data management?*
    - *Reporting?*
    - *Analysis?*
    - *Dissemination?*

- *Are there any other practices at work that you have changed after having completed this course? Please explain.*
    - *Have you attempted to work on changing group practices at your site since having completed this course?*
- *Continuous Quality Improvement*
    - *What does continuous quality improvement mean to you?*
    - *Are you involved in continuous quality improvement at your site? If yes, how?*
    - *Can you share any examples of how the course support your role in quality improvement?*
    - *How could the course have been more effective in supporting your role in CQI?*

5) Work Environment and Next Steps

- *Do you feel that you had the support of your facility management in applying the HIMAE course to practice?*
    - *Were/Are there any specific barriers to your applying this at your post?*

- *Have you had the opportunity to work with others who have completed this course?*
    - *Could you describe that experience?*

- *What could be changed to make it easier to apply this knowledge at your home facility?*

    - *What would motivate you to apply this training at your home facility?*
    - *What is acting as a barrier to your application of this material?*
    - *Do you feel that your supervisor supports your application of this training at your home facility?*
- *What actions would help support the application of this material at your home facility?*

Thank you very much for your time and for all the helpful information you have provided!