Impact Evaluation of the Kenya Frontline Field Epidemiology Training Program

Zeinab Gura Roka

Division for Human Resource for Health Development

Jane Githuku

Field Epidemiology and Laboratory Training Program

Mark Obonyo

Field Epidemiology and Laboratory Training Program

Evalyne Kanyina

Field Epidemiology and Laboratory Training Program

Elvis Omondi Oyugi

Field Epidemiology and Laboratory Training Program

Waqo Gufu Boru

Field Epidemiology and Laboratory Training Program

JAMES Anthony RANSOM (✉ james.ransom@piretpartners.com)

Ministry of Health <https://orcid.org/0000-0001-6528-3994>

Research article

Keywords: field epidemiology, workforce development, Kenya, training

Posted Date: January 9th, 2020

DOI: <https://doi.org/10.21203/rs.2.20485/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background In 2014, Kenya's field epidemiology and laboratory training program (FELTP) initiated a 3-month-long field-based frontline training (FETP-F) for local public health workers.

Methods Between February and April 2017, FELTP conducted a mixed-methods evaluation to examine outcomes achieved among 2014 and 2015 graduates of the trainings. Data quality assessment (DQA) and data consistency assessment (DCA) scores, on-time-reporting (OTR) percentages, and ratings of the training experience were the quantitative measures tracked from baseline and then at 6-month intervals up to 18 months post-completion of the training. The qualitative component consisted of semi-structured face-to-face interviews and observations. Quantitative data were analyzed using one-way analysis of variance (ANOVA). Qualitative data were transcribed and analyzed to identify key themes and dimensions. **Results** One hundred and three graduates were included. For the qualitative component, we reached saturation after 19 onsite interviews and observation exercises. ANOVA showed that the trainings had small but significant impacts on mean DQA and OTR scores.

Results showed an insignificant increase in mean DCA scores. Qualitative analyses showed that 68% of respondents acquired new skills, 83% applied those skills to their day-to-day work, and 91% improved work methods.

Conclusion The findings show that FETP-F is effective in improving work methods, facilitating behavior change, and improving key public health competencies.

Introduction

Strengthened health systems played a key role in the rise in global life expectancy that occurred throughout the 20th century [Roka et al., 2017]. The health workforce is the backbone of each health system, which facilitates the smooth implementation of health action for sustainable socio-economic development [Jones et al., 2017]. Developing the public health workforce in low-to-middle-income countries is a global priority. Workforce competencies and public health agency quality have important implications for global health preparedness, local disease surveillance and response capacity, health systems infrastructure, and overall population health outcomes [Kostava et al., 2017].

Field epidemiology training – frontline (FETP-F) is a three-month competency based, service-oriented collaborative training program that is anchored within the Kenya Ministry of Health (MoH). The partners of FETP-F include Ministry of Agriculture, Livestock and Fisheries, Centers for Disease Control and Prevention (CDC), Kenya Medical Research Institute (KEMRI), and county and sub-county health departments and hospitals.

Together with its partners, FELTP offers a structured local health worker training process (Table 1). The project focused on integration of frontline training as part of the FELTP pyramid (Fig. 1). The first phase

of frontline training was implemented between September 2014 and December 2016 throughout all 47 counties in Kenya.

The goal of FETP-F was to stimulate improvements in local frontline health workers' ability to detect, report, and respond to unusual health events.

FETP-F was designed to improve overall knowledge, skills, and improvements in day-to-day work methods in the work place. Based on this hypothesis, we focused our evaluation on the expected outcomes and associated impact indicators outlined in Table 2.

Table 1a. Didactic structure of FETP-F				
Item	Topic	Course Number	Learning objectives	Outputs
1	Introduction to epidemiology	1	Understand basic epidemiology concepts of person, place, time, agent, host, environment	Pre-test Quiz Case study Post-test
2	Introduction to surveillance	1	Understand basic concepts of surveillance: active vs passive; case definitions; process for prioritizing diseases and other health conditions for surveillance	Pre-test Quiz Case study Post-test
3	Statistics	1,2	Understand how to calculate descriptive statistics: measures of central tendency & dispersion Understand how to calculate measures of frequency: expressing public health data in terms of proportions, rates, ratios	Pre-test Quiz Incorporation of statistical calculations in each case study exercise Post-test
4	Computer software in public health	1,2	Understand how to enter, clean, organize, manipulate, analyze, and display public health data using MS-Excel and OpenEpi software	Daily computer lab sessions Use to collect field data Use to analyze field project data Use to display field project data
5	Data audits	1,2	Application of standard tools to assess the completeness and accuracy of health facility data Application of standard tool to assess the consistency of reported data between health facility, county reports, state reports, and DHIS	Completed DQA tool Completed DCA tool
6	Outbreak investigations	2	Understand the standard steps to undertaking an investigation into rumors, unusual health events	Pre-test Quiz Case study Post-test
7	Data analytics	1,2,3	How to organize and analyze, develop visual displays, communicate, and present public health data	Presenting field project final results to cohort peers and faculty

Table 1b. FETP-F expected outcomes and key indicators	
Expected outcomes	Indicators
Improved skills in field epidemiology, surveillance, and data analytics	Net score improvement on pre- and post-self-evaluation scale
Improved quality of health facility data	Net improvement in DQA scores measured at baseline, 6 months, 12 months, and 18 months post-graduation
Improved consistency of surveillance data within their programmatic area or field project topic area	Net improvement in DCA scores measured at baseline, 6 months, 12 months, and 18 months post-graduation
More and better interaction with data generated by the participant at their work place and/or within their programmatic practice arear	Improved proportion of quarterly on-time submission of MoH form 753 to the county health department at baseline, 6 months, 12 months, and 18 months post-graduation
Improved skills in communicating public health data	<ul style="list-style-type: none"> Number of abstracts written for health conferences Number of abstracts accepted at the submitted conferences The number of manuscripts written for publication in peer review journals

This evaluation estimates the impact of FETP-F, which targets governmental local (county, sub-county, and health facility) health workers (medical and clinical officers, nurses, laboratory scientists, health information and public health officers, and veterinarians) in Kenya.

Table 2
FETP-F expected outcomes and key indicators

Expected outcomes	Indicators
Improved skills in field epidemiology, surveillance, and data analytics	Net score improvement on pre- and post-self-evaluation scale
Improved quality of health facility data	Net improvement in DQA scores measured at baseline, 6 months, 12 months, and 18 months post-graduation
Improved consistency of surveillance data within their programmatic area or field project topic area	Net improvement in DCA scores measured at baseline, 6 months, 12 months, and 18 months post-graduation
More and better interaction with data generated by the participant at their work place and/or within their programmatic practice area	Improved proportion of quarterly on-time submission of MoH form 753 to the county health department at baseline, 6 months, 12 months, and 18 months post-graduation
Improved skills in communicating public health data	<ul style="list-style-type: none"> • Number of abstracts written for health conferences • Number of abstracts accepted at the submitted conferences • The number of manuscripts written for publication in peer review journals • The number of manuscripts accepted for publication in peer review journals

Methods

Between February and April 2017, FELTP used quantitative, semi-quantitative, and qualitative methods to evaluate all FETP-F activities. Groups 1–6 formed the population for the process evaluation because they graduated ≥ 18 months before the impact evaluation began. We followed-up with graduates of these 6 groups to gather the data for the quantitative portion of the summative evaluation. Field project data provided baseline information for the DQA, DCA, and OTR measures. By February 2017, we had 103 graduates who had provided all pieces of the quantitative information. We randomly selected 35 graduates from the 103 for on-site follow-up for 1) face-to-face interviews, 2) observation of their work environment, 3) interviews with their supervisors, 4) interviews with their colleagues, 5) verification of the DQA and DCA and OTR data reported on the past 18 months.

Quantitative measures. We used interrupted repeated measures on 3 quantitative values. Because all participants had to complete these activities, we used the field project values as the baseline values. Afterward, during the process evaluation at 6 months, we gauged their measures at that time. One year post-graduation, we gauged the same measures again. The final measure was taken at least 18 months post-graduation. All measures were self-reported via online survey. For all quantitative measures, we conducted one-way ANOVAs using MS-Excel's data toolpak. The quantitative measures are described below.

DQA scores. The participants had to complete a DQA for their field project, and we used these scores as the baseline scores. The DQA tool was designed to: (1) verify the quality of health facility data, (2) assess the system that produces that data, and (3) develop action plans to improve both [Cheburet et al., 2016]. We subsequently asked graduates to measure DQA scores from the same data source as the field project, except it should be 6 months beyond the baseline data. We then asked them to repeat the procedure at month 12 and finally at month 18 post-graduation. We used one-way ANOVA to determine if there were differences in the mean DQA scores over time.

DCA scores. The DCA is an end-to-end data integrity process. Because DCA focuses on the entire surveillance network, we did not ask the graduates to cross-check all data points in the system. We asked them to do a check of the indicators and counts used in their field projects. The first end is the generation of data at the health facility level. The middle is the county record, where the health facilities report their weekly and monthly tallies to the county health department (CHD) using MoH form 753. Then those data are entered into the district health information system (DHIS) by the county health records and information officer (HRIO). The DCA process is outlined in Fig. 2.

The goal is to detect inconsistencies as data travel through the surveillance system and identify root causes for these inconsistencies and to develop solutions, at the most granular level of the surveillance system – the health facility. The DCA score indicates the depth of the inconsistencies. A low score indicates high inconsistencies, whereas a high score indicates low inconsistency levels of the data between the 3 repositories of surveillance data. We used one-way ANOVA to determine if there were differences in the mean DCA scores over time.

Timeliness of reporting. Timeliness is a key performance measure of public health surveillance systems. Timeliness can vary by disease, intended use of the data, and public health system level. The participants, as part of their field projects, had to evaluate the timeliness of reporting for the condition, disease, or health priority that was the focus of their field project. We used the results from the field project as baseline OTR measures. Then we followed up at 6 months to assess the proportion of reports submitted on time for the previous quarter. We repeated the procedure at 12 months post-graduation and then a final query at least 18 months post-graduation to examine the proportion of OTRs for the prior quarter. We used one-way ANOVA to determine if there are differences in the mean OTR scores over time.

Semi-quantitative measures. At the beginning of each training course, we asked participants to score their knowledge and skills in 8 key competencies on a Likert scale from 1 to 5, with 1 representing limited

knowledge/skills and 5 representing expertise (Table 3). At the end of the 3-month training, we asked them to gauge their knowledge skills in each of those areas now that they have sat through 30 hours of didactic training, received hands-on coaching and mentoring from FELTP faculty, and completed a 5-week field project. We use the pre-post difference as our comparison point when we followed up after 18 months and asked them to rate their knowledge and skills now in terms of practical applications to their day-to-day work. We also asked their supervisors and colleagues to score the graduates' skills and knowledge and practical application in each of those competencies. We used those scores to gauge the impact of FETP-F training on knowledge, skills, and change in work methods. We conducted a one-way ANOVA to determine if there was a difference in the scores between the 3 groups.

Table 3
Pre- and post-self-assessment score sheet, public health competencies, n = 103

Before Training					Self-assessment of Your Knowledge and Skills Related to:	After Training				
1	2	3	4	5	I can calculate basic measures of central tendency: mean, median, & mode	1	2	3	4	5
1	2	3	4	5	I can calculate basic statistical measures of dispersion: range, variance, & standard deviation	1	2	3	4	5
1	2	3	4	5	I have an understanding of descriptive epidemiology	1	2	3	4	5
1	2	3	4	5	I have an understanding of basic disease surveillance	1	2	3	4	5
1	2	3	4	5	I can use MS-Excel to enter and manipulate basic data	1	2	3	4	5
1	2	3	4	5	I know how to use formulas in MS-Excel	1	2	3	4	5
1	2	3	4	5	I know how to look at data and analyze it by person, place, and time	1	2	3	4	5
1	2	3	4	5	I know the basic components of a functional disease surveillance system	1	2	3	4	5
1	2	3	4	5	I know the difference between active and passive surveillance systems	1	2	3	4	5
1	2	3	4	5	I know the difference between qualitative and quantitative data	1	2	3	4	5
1	2	3	4	5	I can develop a case definition to use during a field investigation	1	2	3	4	5
1	2	3	4	5	I can analyze field and surveillance data and apply appropriate statistics to describe what the data show	1	2	3	4	5
1	2	3	4	5	I can audit health facility data for accuracy and completeness	1	2	3	4	5

Qualitative measures. The qualitative portion of the evaluation used grounded theory to determine the impact of the training, mentoring, and supervision on behavior, work method, and application of training to work duties [Reeves et al., 2015]. The grounded theory approach allowed us to develop our inquisitive instruments and then draw theory from them as we analyzed the interview transcripts, the interviewers' field notes from observations, and the previous responses to other evaluation tools as the graduates went through the 3-month training.

Semi-structured interviews were conducted with randomly selected graduates from groups 1–6. Because we wanted to examine the impact of the training at least 1.5 years post-graduation, so that we could, at most, look at the first 6 groups to go through the FETP-F process. Those groups enrolled between July 2014 and July 2015. All interviews were recorded with the consent of the interviewees. All interviewees had to provide written and verbal consent to the interview.

Participant observation. Participant observation allowed the field investigators to establish rapport with the person being interviewed so that the interviewee would provide more honest answers and opinions (vs answering what they think the interviewer wants to hear) [Laurier, 2016]. We used a checklist for field workers to note the presence of monitoring charts, active use of the DQA and DCA tools, and presence of log books, interactions with colleagues, presence of operational tools such as laptops, printers, case definition posters. Field investigators had to submit their field notes and completed checklists along with the consent forms and audio files to the evaluation cloud site. Audio files were sent out for transcription to parties unaffiliated with FELTP, the participants, CDC-Kenya, and other stakeholders.

Developing the interview instrument. In September 2016, FETP-F initiated development of the semi-structured interview and guide that would provide data to support the impact evaluation. The questionnaire was field tested in South Sudan for clarity and then tested with medical education partnership initiative (MEPI) graduates in Kenya. We did not develop questions that could use a Likert scale to measure these 2 attributes of the graduates. Instead, we used open-ended questions that would help us getting a fuller picture of the impact on multiple areas of their personal and work lives [Melovitz et al., 2018].

Trainings to prepare for evaluation. In September 2016, FETP-F hosted a one-week workshop on qualitative research methods in Nairobi, Kenya, facilitated by CDC/DGHT (Atlanta, Georgia, USA). The workshop participants were FELTP faculty and CDC-Kenya staff who supervised the evaluation field workers, who would conduct the interviews for the evaluation.

Field training. In December 2016, 17 MEPI graduates underwent one week of residency-based training by CDC-Kenya staff on interview and observation methods. Kenya FELTP staff also showed them how to conduct the confirmatory DQA and DCA exercises. Each MEPI graduate had to complete the online FHI-360 ethics training module before graduating from the training course [Aalborg et al., 2016].

Results

Demographics of survey respondents. For the quantitative analyses, 103 graduates were included in the analyses (Table 4). Most (55%) were male and 60% (n = 62) had < 10 years of public health work experience. Of the geographical regions, 36% were from central region (groups 2 and 5); 23% were from Nyanza region (group 4); 8% were from northern regions (group 3); 5% were from coast (group 6), and 5% were from the National Public Health Laboratory (group 1). The break down by cadre were 20% medical officers, 15% veterinary officers, 25% public health officers, 15% laboratory, 15% nursing, and 10% other.

Table 4
Participant characteristics of those who completed the impact evaluation measures
(n = 103)

Characteristic	N (%)
Years of public health service	
< 5 years	38 (37)
5–9 years	24 (23)
> 9 years	41 (40)
Gender	
M	57 (55)
F	46 (45)
Regional affiliation	
Central	40 (39)
Nyanza	24 (23)
Northern	19 (18)
Coastal	15 (15)
National	5 (5)
Program area	
NCDs	7 (7)
HIV	21 (20)
TB	17 (17)
All	46 (45)
MCH/ANC	12 (12)
Agency type	
County HD	24 (23)

Percentage reported for valid, non-missing data
Some percentages do not sum to 100% due to rounding
Examples of other program areas include family planning, surgery, immunizations
Examples of other cadres include HRIOs, EPI logistician

Characteristic	N (%)
Sub-county HD	48 (47)
Health facility	26 (25)
National	5 (5)
Cadre	
MO	21 (20)
VO	15 (15)
PHO	26 (25)
Lab	15 (15)
Nurse	16 (15)
Other	10 (10)
Percentage reported for valid, non-missing data	
Some percentages do not sum to 100% due to rounding	
Examples of other program areas include family planning, surgery, immunizations	
Examples of other cadres include HRIOs, EPI logistician	

DQA scores. Descriptive analyses of 103 DQA scores from baseline to 18-months post-graduation showed an increase in the mean DQA score from 75.64% at baseline to 84.53% at 18-months post-graduation (Fig. 3). This shows a 10.5% improvement in the mean DQA score for this sample of health facilities and programs. The subsequent ANOVA analyses on the 103 respondents showed that although the improvement was only 10.5%, that it still represented a significant improvement in DQA mean scores since baseline (Table 5).

Table 5
ANOVA results, n = 103

A. ANOVA results, DQA mean scores*		
Time interval post-graduation	DQA mean	SD
Baseline	75.64	8.05
6 months	74.88	9.00
12 months	75.08	5.21
18 months	84.53	8.82
B. ANOVA results, DCA mean scores**		
Time interval post-graduation	DCA mean	SD
Baseline	73.22	27.59
6 months	68.11	13.42
12 months	78.22	21.46
18 months	82.66	21.37
C. ANOVA results, OTR mean scores***		
Time interval post-graduation	OTR mean	SD
Baseline	29.66	15.58
6 months	70.11	23.39
12 months	70.83	180.1471
18 months	74.88	624.3399
*Between-groups: F-statistic = 70.71; f-crit = 2.61; p < 0.0001;		
** Between groups: F-statistic = 0.765; f-crit = 2.90; p = 0.52;		
*** Between-groups: F-statistic = 20.37, f-crit = 2.74, p < 0.0001.		
ANOVA = analysis of variance; DQA = data quality assessment; SD = standard deviation; DCA = data consistency assessment; OTR = on-time reporting.		

DCA scores. Descriptive analyses of DCA scores showed that there was an 11.4% improvement in DCA scores between baseline and 18 months post-graduation. However, upon further analyses using ANOVA, results showed that the increase was not significant (Table 6).

Table 6

Comparison of pre-post score differences among the graduates, their supervisors, and their colleagues.

	Self	Supervisor	Colleague
Statistics	1	2	2
Epidemiology	3	2	3
Surveillance	2	2	2
MS-Excel	1	3	2
Data analysis	2	1	2
Field investigations	2	2	2
Data audits	3	2	3
Communicating PH data	2	2	0
The scale is Likert, 1–5, with 1 = minimal knowledge/skill to 5 = exceptional knowledge/skill in the competency.			
Difference scores were calculated by subtracting the “pre” score from the “post” score. All results were positive.			
Between-groups: $F = 0.76$, $f\text{-crit} = 2.90$; $p = 0.52$.			

OTR proportions. We examined the proportion of monthly reports submitted on time from health facilities to county health departments for the preceding quarter (Table 7). The descriptive analyses show that there was a 60% increase in OTR between baseline and the 18-month assessment. The ANOVA showed this to be a significant development and improvement compared to baseline values (Fig. 4).

Table 7
Changes in knowledge and skills, FETP-F graduates, 2014–2015, n = 103

Competency	Time of measurement, mean (SD)		
	Pre-training	Post-training	Follow-up
Statistics	2.77 (0.81)	3.69 (0.61)	4.35 (0.72)
Epidemiology	2.68 (0.72)	4.11 (0.45)	3.74 (0.69)
Surveillance	2.82 (0.73)	3.84 (0.59)	3.99 (0.51)
MS-Excel	1.86 (0.75)	3.81 (0.55)	3.97 (0.62)
Data analysis	2.55 (0.96)	3.95 (0.69)	3.56 (0.49)
Field investigations	2.32 (0.89)	3.47 (0.82)	2.66 (0.74)
Data audits	2.86 (0.99)	3.89 (0.55)	3.82 (0.62)
Communicating PH data	2.73 (0.58)	3.94 (0.31)	4.02 (0.47)
The ordinal scale ranged from 1 to 5 (1 = no knowledge, 2 = little knowledge, 3 = average, 4 = good, and 5 = mastery). SD = standard deviation; pre-assessment, mean scores before the training; post-assessment, means scores immediately after completing the 3-month training process; follow-up, mean scores at least 18 months post-graduation from FETP-F. Between-groups: F = 30.02; f-crit = 3.47; $p < 0.0001$.			
FETP-F = field epidemiology training program-frontline; SD = standard deviation; PH = public health			

Semi-quantitative self-assessment of learning scores (pre-post difference) compared to assessment by supervisors and colleagues showed significant increases (Table 8). Knowledge/skill levels within the 8 competencies were relatively low before the training. After training, we noted significant increases in the mean knowledge/skill scores in each of the 8 competencies. During the site visits, field workers also interviewed supervisors of the graduates and at least one colleague regarding any notable changes (positive or negative) after the graduate resumed his/her normal work duties. We used the same assessment scale as with the graduates. Results are outlined in Fig. 5.

Table 8
Processes of analyses of qualitative data, n = 38

Level	Description
1	Process of open coding of the transcript to reduce the qualitative data to a more manageable focus.
2	We created categories from the level 1 open codes. This means that multiple level 1 codes were lumped together to create level 2 codes.
3	We then re-examined the level 2 codes to look for patterns, key words and “intent” of statements to generate themes and dimensions. This helped us tabulate the data and prepare it for subsequent analyses, such as inclusion in hierarchical regression models.
4	Reviewed the codes, themes, and dimensions to generate theory regarding the why and how (or why not) the training had an impact on its graduates and their affiliated organizations

There was not much variation in the self-assessments of the graduates when compared to the assessments of competencies provided by their supervisors and colleagues. However, the supervisors and colleagues noted a marked increase in MS-Excel skills knowledge and expertise post-graduation.

For the larger group of graduates (n = 103), we examined via online survey the mean skills and knowledge changes (pre-post) in the key competencies before training (pre-assessment), immediately after the 3-month session ended (post-assessment) and 18 months after training (follow-up) (n = 103) [Table 9].

Table 9

Results of analyses of the transcripts generated by interviews with the FETP-F graduates (n = 19), their supervisors (n = 12), and their colleagues (n = 7).

Dimension	Level 2 codes and themes	Example quotes
Personal	<ul style="list-style-type: none"> • Now seen as an organizational and health sector asset • CDC certification helped career advancement • More independent • More confident • Serves as mentor to others and teach them how to use/apply the information from FETP-F • Uses course materials to train others • Acquired short-term contracts and research opportunities • Greatly improved computer skills • Better public speaking and communications skills • Became a better leader • Now interested in writing abstracts and reports because of the opportunities that come with that skill 	<ul style="list-style-type: none"> • Which there before I used to do it guessing of what I thought was the right thing. But right now, I can comfortably analyse data and interpret it and present it to whoever needs it. • Then I also love the way we were put to task on how to present public health data • I have a lot of confidence; I like where people do presentations and I will really feel to be very keen because I can interrogate data also, I got some skills of just trying to see somebody's data and get a question out of one at least. • I think more of it is coming to the projects that we had. The projects that we had –actually from what we learnt, it was – almost after the basic training, there was an outbreak in Muranga for cholera. And because we had undergone that training, we were able to handle the case and we were able to isolate –to do the cultures and isolate and meet the identifications of the organism causing cholera. So, we were able to handle that outbreak within the county. • Yea, because previously I would not know how to do the data analysis and also how to do the –how to write a paper, but now I have the confidence even to do a paper and also present it in the forums. • Yes, like world bank we have been taking data from the...other facilities we have also been taking surveillance and given them a feedback and due to this I have also been going to other countries like Tanzania doing epidemiology and also collecting data and giving it to the world bank.

Dimension	Level 2 codes and themes	Example quotes
Organizational	<ul style="list-style-type: none"> • Data management improved • On time reporting rate improved • Improved response capacity to deal with repeated cholera outbreaks • Agency can make scientifically informed decisions about interventions 	<ul style="list-style-type: none"> • I feel she is prepared becauseas we have said, she has the knowledge now and let me site the same example, that of cholera. She was able to act very fast because the case was reported at around 4 pm and by five she was already here at the facility and she gathered everybody here and we were able to manage the same. • She is better prepared, the reasons I had said is that what happens is that she is able to pick weekly data for –like we have diseases that we focus ion so much especially diarrhea infections and we have diarrhea prone areas. So, the facility catchment areas in those areas, she is able to tell us that this facility is now reporting abnormal numbers. Then she is also able to tell us—to sound a warning if those data are not real data. Because you know she is even the surveillance coordinator. She can appoint that this data looks like it was cooked.
FETP-F processes	<ul style="list-style-type: none"> • Knowledgeable facilitators • Selection bias – favors physicians • Poor follow-up with graduates • Good adult learning techniques • Should provide refresher trainings for graduates • Training was rushed – too much material covered in short time period • Residency-based training was ideal • Difficulty of the course added to its value • Relationships with faculty improved quality of the training program 	<ul style="list-style-type: none"> • The comments are that we thought this training would pick again some people, continuity of training, could pick again others so that we bring others on board, people continue, and I didn't see that • Okay they have been very effective because considering the composition of the basic training that we undergo; it is inclusive of all cadres. Therefore, you can interact with people with different backgrounds of training and learn some new things from them • They have been very effective also because the entire faculty, one it is friendly, they are very professional and they are also facilitative because like for example if you have any project that you need to work on and present in a scientific conference, they are always willing to support you, and also facilitate you to attend the conference.

Qualitative results. Field investigators visited 19 sites and conducted 38 one-on-one private interviews (graduate, supervisor, colleague). We analyzed the transcripts of all interviews (n = 19 graduates, n = 12 supervisors, and n = 7 colleagues). After transcription, we conducted 3 levels of analysis. The coding

process was iterative and involved multiple stages that prepared and formatted the raw data so that they are available for evaluation. Each level of analysis is outlined in Table 10.

The results of 3 levels of analysis are outlined in Table 11. After conducting the level 1 analyses using key word searches and generation of word clouds, we had a list of 107 codes. During the 2nd level analyses, we reduced this to 37 codes that we later grouped into 25 themes. After the 3rd level review, we noted that the themes fell into 3 key dimensions. Graduates, their supervisors, and their colleagues' comments were associated with the "personal" (benefits to self), organizational (benefits to the agency or organization where the graduate worked or health partners in the graduate's community), and the FETP process itself (feelings and perspectives on the nominations/selection process, the execution of the course inclusive of its contents; and feedback on the quality of the faculty and facilitators).

Discussion

This evaluation results provide support for the effectiveness of a localized field epidemiology and data management training process for improving skills and capacity of frontline health workers. During the interviews, most graduates, their supervisors, and colleagues reported that the course had helped them to make scientifically based decisions and improved their overall capacity to deal with a spectrum of public health challenges, from calculating thresholds to responding to cholera cases. Additionally, they report that the course helped them to become better leaders by improving their communications skills, more evidenced-based decision making, and showing colleagues how to practically interact more critically with the data they generate at their agencies.

Some limitations of the current evaluation should be noted. First, the bulk of the data collected are self-reported, reporting DQA, DCA, and OTR scores as well as measuring respondents' perceptions of learning and impacts. It is possible that participants over- or under-rated their skills and knowledge when responding to survey items online. Second, the time gap from delivery of the course to data collection could have affected the information that graduates gave to us. Additionally, the data collection had to be rushed due to pending funding cuts. The program was not allocated ample funds for a more thorough evaluation process. This will hinder sub analyses of the formative and summative evaluation data over the life of the project. Further efforts are needed to determine if skills and/or benefits from the course change over time and whether the documented improvements in health facility data quality, consistency, and on-time reporting change over time, particularly as replications continue and the time gap since training widens and we don't have a steady flow of their colleagues who can participate in such training.

Further, many of the graduates did not respond to the measures surveys, so we don't have relevant data about them and thus are not able to conclude that respondents consistent of a representative sample of graduates.

Finally, we know that FETP-F's participants take part in a never-ending array of trainings, so we don't know how those other trainings have impacted the findings that we have documented [Beer et al., 2016]. We also don't know the spectrum of participants' involvement in support networks, how the doctors' and

nurses’ strikes affected outcomes, the role of politics in who is nominated to participate in the training, local rates of job turnover, and how that the FETP-F does not award a diploma affects uptake among some younger public health workers.

Conclusions

FETP-F is a viable and effective method for improving Kenya’s public health workforce’s skills, knowledge, and practices in key competencies. This evaluation suggests many benefits and lessons on frontline field epidemiology training including 1) the advantage of focusing on local health workers who are more familiar with contextual issues to allow tailoring of the training; 2) enhanced collaboration among multiple practice cadres to create a forum for networking and new partnership opportunities; 3) a more convenient method of training that eliminates the need to bring in external trainers or for participants to travel outside of their region; and 4) specific examples of how to improve future iterations of this kind of training. This evaluation suggests that the FETP-F model has increased the capacity of local health workers trained in field epidemiology and data analytics while maintaining fidelity with the original objectives and frameworks of the original model, the advanced-level field epidemiology training program. The FETP-F met its aims and objectives satisfactorily, and resulted in positive shifts in knowledge, attitudinal, and behavioral intentions of local health workers who graduated from the program. This suggests that this training strategy was effective and feasible in improving the capacity of local public health workers of all cadres.

Abbreviations

ANOVA	Analysis of variance
CDC	U.S. Centers for Disease Control and Prevention
CHD	County health department
DCA	Data consistency assessment
DGHT	Division of Global HIV and TB
DHIS	District health information system
DQA	Data quality assessment
FELTP	Field epidemiology and laboratory training program
FETP-F	Field epidemiology training program – frontline
HRIO	Health records and information officer
OTR	On-time reporting

Declarations

Ethics approval and consent to participate

Informed consent was obtained from all FETP-F graduates who agreed to an evaluation visit. Personal identifiers were not included in the recorded data. Permission to conduct this evaluation was sought from and granted by the Ethical Review Board of the Ministry of Health (FAN: IREC 1795). This evaluation did not involve any animal subjects. The evaluation did not collect human subject data nor any human specimen samples. All subjects provided signed and oral consent for participation.

Consent for publication

Informed consent included consent to publish findings of this evaluation research. This research did not use any images, names, or other identifying information of any of those who consented for interview and participation in the evaluation. Therefore, a consent for publication was not needed from any of the research subjects.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors report no competing interests.

Funding

Funding for FETP-F activities was provided by the World Bank, CDC-DGHP, GCE-South Sudan, and DTRA.

Authors' contributions

Z.G., J.G., E.O., W.B., and J.R. conceived and developed the evaluation tools and overall plan. J.G., E.O., E.K., and W.Q. supervised the implementation and collection of the evaluation data. J.R., Z.G., E.O., and W.B. cleaned and analyzed the data. J.R. and Z.G. contributed to interpretation of the results. J.R. took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript.

Acknowledgments

The authors acknowledge the Kenya FELTP, CDC-Kenya, and CDC-Atlanta (DGHP and DGHT), the MEPI graduates for collecting data, the interview transcribers, and AFENET.

Authors' information

¹Field Epidemiology and Laboratory Training Program, Nairobi, Kenya; ²Food and Agricultural Organization, United Nations, Nairobi, Kenya; ³Piret Partners Consulting, Washington, DC, USA.

References

1. Roka, Z.G., Githuku, J., Obonyo, M., Boru, W., Galgalo, T., Amwayi, S., Kioko, J., Njoroge, D. and Ransom, J.A., 2017. Strengthening health systems in Africa: a case study of the Kenya field epidemiology training program for local frontline health workers. *Public health reviews*, 38(1), p.23.
2. Jones, D.S., Dicker, R.C., Fontaine, R.E., Boore, A.L., Omolo, J.O., Ashgar, R.J. and Baggett, H.C., 2017. Building global epidemiology and response capacity with field epidemiology training programs. *Emerging infectious diseases*, 23(Suppl 1), p.S158.
3. Kostova, D., Husain, M.J., Sugerman, D., Hong, Y., Saraiya, M., Keltz, J. and Asma, S., 2017. Synergies between communicable and noncommunicable disease programs to enhance global health security. *Emerging infectious diseases*, 23(Suppl 1), p.S40.
4. Cheburet, S.K. and Odhiambo-Otieno, G.W., 2016. Process factors influencing data quality of routine health management information system: case of Uasin Gishu County Referral Hospital, Kenya. *Int Res J Public Environ Heal*, 3, p.6.
5. Reeves, S., Boet, S., Zierler, B. and Kitto, S., 2015. Interprofessional education and practice guide no. 3: evaluating interprofessional education. *Journal of Interprofessional Care*, 29(4), pp.305-312.
6. Laurier, E., 2016. 11 Participant and Non-participant Observation. *Key Methods in Geography*, p.169.
7. Melovitz Vasan, C.A., DeFouw, D.O., Holland, B.K. and Vasan, N.S., 2018. Analysis of testing with multiple choice versus open-ended questions: Outcome-based observations in an anatomy course. *Anatomical sciences education*, 11(3), pp.254-261.
8. Aalborg, A., Sullivan, S., Cortes, J., Basagoitia, A., Illanes, D. and Green, M., 2016. Research ethics training of trainers: developing capacity of Bolivian health science and civil society leaders. *Acta Bioethica*, 22(2).
9. Beer, M., Finnström, M. and Schrader, D., 2016. Why leadership training fails—and what to do about it. *Harvard Business Review*, 94(10), pp.50-57.

Figures

Figure 1

The current pyramid model of field epidemiology training and public health workforce development being implemented as part of KFETP.

Figure 2

The DCA process

Figure 3

Descriptive analyses of 103 DQA scores from baseline to 18-months post-graduation showed an increase in the mean DQA score from 75.64% at baseline to 84.53% at 18-months post-graduation

Figure 4

The ANOVA showed this to be a significant development and improvement compared to baseline values

Figure 5

. Comparisons of mean difference scores among FETP-F graduates (n=19), their supervisors (n=12), and their colleagues (n=7) in 8 competency areas, February – April 2017. Scale is 1-5, Likert.