Strengthening quality management systems of clinical laboratories in Cambodia and the role of quality mentors: a quantitative evaluation of program outputs and outcomes

Grant Donovan

A thesis

Submitted in partial fulfillment of the

requirements for the degree of

Master of Public Health

University of Washington

2019

Committee:

Lucy A. Perrone

Annette L. Fitzpatrick

Michael L. Astion

Patricia Sadate-Ngatchou

Program Authorized to Offer Degree:

Department of Global Health

University of Washington

Abstract

Strengthening quality management systems of clinical laboratories in Cambodia and the role of quality mentors: a quantitative evaluation of program outputs and outcomes

Grant Donovan

Chair of the Supervisory Committee:

Lucy Perrone, MSPH, PhD

Department of Global Health

Programs to improve laboratory quality management systems (LQMS) in resource-limited countries have been shown to be an effective means to meet international accreditation standards for improved health outcomes. In the context of Cambodia, where most clinical laboratories have not yet met these standards, much can be learned from the evaluation of current LQMS improvement programs and their effectiveness at achieving measurable change. Toward this end, this study describes and evaluates the methods, outputs, and outcomes of the International Training and Education Center for Health's Laboratory Health Systems Strengthening Program in 12 clinical laboratories between January 2018 to April 2019. The program goal is to improve laboratory operations in Cambodia for enhanced disease detection, surveillance, and biosecurity through improved quality assurance and management practices. In participating laboratories, the program used a combination of methods, including formal LQMS training, on-site technical

consults, and remote mentoring by trained quality mentors through video conference technologies. For evaluation of this program within the logical framework of program activities in participating laboratories, activity outputs were quantified from the activity reports of program trainers and mentors; direct program outcomes were measured by pre- and post-implementation audits of LQMS conformity to international standards; and changes in laboratory quality as indirect outcomes were assessed through a set of common indicators of laboratory quality.

Program activities achieved an average output per laboratory of 23±2 trainings of target personnel, 11±2 site visits, and 6027±2454 minutes of video-conference activity participation with quality mentors. Participating laboratories significantly improved performance in pre- and post-implementation audits as a program outcome over a 15-month evaluation period by a mean percent difference of 22±12% ($p=0.002$). A separate cross-sectional comparison found that 2019 overall scores of intervention laboratories (median=56.5) were significantly higher ($p=0.0001$) than those of a sample of laboratories with no intervention (median=22%). Significant differences between comparison groups were found in all individual audit sections addressed directly through program activities. In subsequent analysis of physician satisfaction ratings, corrected report frequencies, test turnaround times, and external quality assessment performance, changes in these measures of laboratory quality have not yet demonstrated the desired long-term outcome. This study demonstrates, however, the potential of the Cambodia laboratory information system and national external quality assessment program as long-term monitors of quality and provides recommendations for improved data collection and process improvement.

Regarding the role of quality mentors, a test for correlation between pre- and post-audit score differences and laboratory personnel participation time in remote mentoring activities identified a strong monotonic relationship ($r_s=0.66$, $p=0.02$). This relationship highlights the role

of remote mentoring through video communication technologies as a potentially valuable approach for future program development in similar, resource-limited settings. The study as a whole shows that LQMS performance and conformity in participating laboratories has improved significantly in Cambodia after a mixture of remote and on-site mentoring along with a formal training curriculum, suggesting that replication of these approaches may be beneficial if expanded to non-intervention laboratories, which currently demonstrate significantly lower performance in LQMS audits.

Table of Contents

Acknowledgements

I would like to give my sincerest thanks to my advisors, Drs. Lucy Perrone, Michael Astion, Annette Fitzpatrick, and Patricia Sadate-Ngatchou for their invaluable guidance throughout this project. Thank you as well to the I-TECH Cambodia team, Dr. Siew Kim Ong, Sophanna Song, Chhay Heng Leang, and Malin Chao for all of their contributions to this project and for answering my many questions. Finally, to Mr. Vanra Ieng of the World Health Office in Cambodia for his time and effort to provide data for this evaluation, and to the Cambodia Ministry of Health Bureau of Medical Laboratory Services and National Institute of Public Health for their cooperation and shared information, thank you.

This work was supported by the US Department of Defense Threat Reduction Agency through contract number HDTRA1-08-D-0007, awarded to Principal Investigator, Lucy Perrone at the International Training and Education Center for Health. The contents of this study are the responsibility of the author's and are not the official views or statements of the Defense Threat Reduction Agency or the US Government.

Introduction

Laboratory quality management systems in Cambodia

Developing strong laboratory quality management systems (LQMS) is a key component of strengthening health systems for improved health outcomes and disease surveillance in resource-limited countries, requiring standardization and strategic planning[1-3]. ISO 15189 accreditation, which is the international standard for medical laboratory quality management systems recommended by the World Health Organization (WHO), provides standardization of LQMS requirements with a strong technical foundation for health, safety, and conformity[4]. These standards are stringent, however, and have required a variety of approaches for laboratories at different levels of development to achieve them[4,5].

In Cambodia, a national effort to meet International Health Regulations and improve health outcomes has culminated in the endorsement of national laboratory quality standards in 2011 and a national laboratory strategic plan in 2010 with an objective to improve quality management systems of national and regional medical laboratories[6,7]. Subsequent progress has included the integration of external quality assessments (EQA) through private and public partnerships, and the development of a National Laboratory Information System (CamLIS) in collaboration with WHO[8]. In 2013, however, a comprehensive assessment of laboratory quality management capacity in Cambodia using the WHO Laboratory Assessment Tool (LAT) revealed general indicator scores of laboratory capacity to remain deficient[9]. The assessment measured 11 indicators of laboratory capacity, identifying a low average score of 36% in 22 laboratories, with indicators of LQMS averaging only 47% due to a lack of quality management systems, trained quality assurance managers, or continuous improvement practices[9,10].

There are several examples of successful LQMS development programs demonstrated in Cambodia[8,10,11]. In 2011 the US Center for Disease Control (CDC) partnered with the Cambodia Ministry of Health (MOH) to implement the Strengthening Laboratory Management Toward Accreditation (SLMTA) program to meet the need for LQMS development[11]. The partnership implemented baseline audits of 12 laboratories in 2012, subsequently providing a series of trainings and learning projects with mentored support for LQMS development[11]. The Stepwise Laboratory Improvement Process Towards Accreditation (SLIPTA) audit tool used to monitor the SLMTA program has since been adapted by the Cambodia MOH to develop the Cambodia Laboratory Quality Management System Checklist for Accreditation (CamLQMS), which is now used by the MOH as a primary method of LQMS assessment through the Bureau of Medical Laboratory Services, otherwise referred to as BMLS[12] (BMLS Cambodia, unpublished presentation, 2018). Successes of the SLMTA program in Cambodia and globally have shown the impact that structured and mentored LQMS programs can have in resource-limited healthcare systems, suggesting promise for other such programs in Cambodia in the future[12].

The International Training and Education Center for Health (I-TECH), a non-profit organization affiliated with the University of Washington, partnered with the MOH in Cambodia in 2014 to improve and expand LQMS implementation nationally as part of the organization's Laboratory Health System Strengthening (LHSS) Program. This program began with a mentored and stepwise approach to accreditation, complementary to SLMTA's approach, and later transitioned to a combination of training, mentorship, and advocacy that supported the current strategy to LQMS development[7,10]. The current mentored LQMS approach of the LHSS Program in Cambodia from July 2017 to April 2019 is the primary focus of this study, which aims to

describe and evaluate the process and outcomes of program implementation in participating laboratories.

Program background

During the initial project of the LHSS program in Cambodia, participating laboratories utilized the Laboratory Quality Stepwise Implementation (LQSI) toolkit with the support of a local I-TECH Cambodia mentor, trained in quality management system implementation[10]. The LQSI was developed by the Royal Tropical Institute for WHO as a public resource for medical laboratories in resource limited settings to implement LQMS and achieve ISO 15189 accreditation, and the tool was adapted by I-TECH for use by Cambodian laboratories[10,13]. This program has since been implemented in 12 laboratories and has demonstrated that consistent on-site mentoring in the local language with an action plan and stepwise checklist toward accreditation standards enhances LQMS development without interrupting regular laboratory services[10].

One finding from a follow-up assessment of the 2013 independent LAT assessment in 2015 determined that laboratories participating in I-TECH LHSS Program's mentored LQSI project improved mean general indicator scores from a baseline of 52% in 2013 to 69% in 2015, while those of non-mentored laboratories improved from a baseline of 49% to 50% during the same period[9,10] (WHO Cambodia, unpublished data, 2016). While this study and further program evaluation have demonstrated that I-TECH's past program approaches are valuable to the field of laboratory quality improvement, recent external evaluation recommendations and changes to the national strategy to LQMS development have prompted a change in approach to I-TECH's LHSS Program in Cambodia[7,8]. The primary goal of the program is

> *"to improve laboratory operations in Cambodia for enhanced disease detection, surveillance, and biosecurity through improved laboratory quality assurance and management practices in tertiary level and national referral hospital laboratories, in line with national and international quality standards and regulations" (L.A. Perrone, PhD, unpublished presentation, April 2019).*

The general program objectives were carried out through a mixture of advocacy for legal and regulatory framework, curriculum development, continuing mentorship and training of participating laboratory managers and Quality Assurance Officers (QAO), and the support of additional quality assurance oversight and networking.

One notable shift in implementation of the program includes a complete switch to the nationally driven CamLQMS checklist as the primary assessment tool for stepwise LQMS development to meet accreditation standards. In January of 2018, a baseline CamLQMS assessment in the twelve participating mentored laboratories was completed in partnership between I-TECH Cambodia and the Cambodia BMLS. This audit was followed by a dissemination meeting and training workshop where QAOs and laboratory managers were trained in action planning to eliminate gaps or deficiencies among the 12 sections of quality assessed within the audits. These 12 sections are derived from the 12 Quality System Essentials (QSE) established by the Clinical and Laboratory Standards Institute as follows[14]:

a. Documents and records
b. Management reviews
c. Organization and personnel
d. Client management and customer service
e. Equipment
f. Evaluation and audits
g. Purchasing and inventory
h. Process control and internal and external quality assessment
i. Information management
j. Corrective action
k. Occurrence management and process improvement
l. Facilities and safety

I-TECH Cambodia then implemented a series of eleven trainings that further targeted gaps in these quality management essentials within participating laboratories and continued to support laboratories through on-site reviews along with frequent video conference training and mentoring, occurring at least weekly. These activities were followed by a second CamLQMS audit, giving post-intervention LQMS capacity data for each laboratory. This collection of data, along with an assessment of outputs and additional quality indicator data during the same period, is the foundation of the program's evaluation.

Methods for evaluation

This study is a quantitative evaluation, describing the I-TECH Cambodia LHSS Program activities, outputs, and outcomes in participating laboratories during the evaluation period of January 2018 to April 2019. The evaluation used an uncontrolled longitudinal study to assess changes in LQMS compliance to international standards and changes in quality between baseline and endpoint measurements. A cross-sectional analysis was then used to compare post-implementation LQMS performance and conformity of intervention laboratories to a select group of non-intervention laboratories. The logic model, which provides context for the evaluation, is demonstrated in Figure 1.

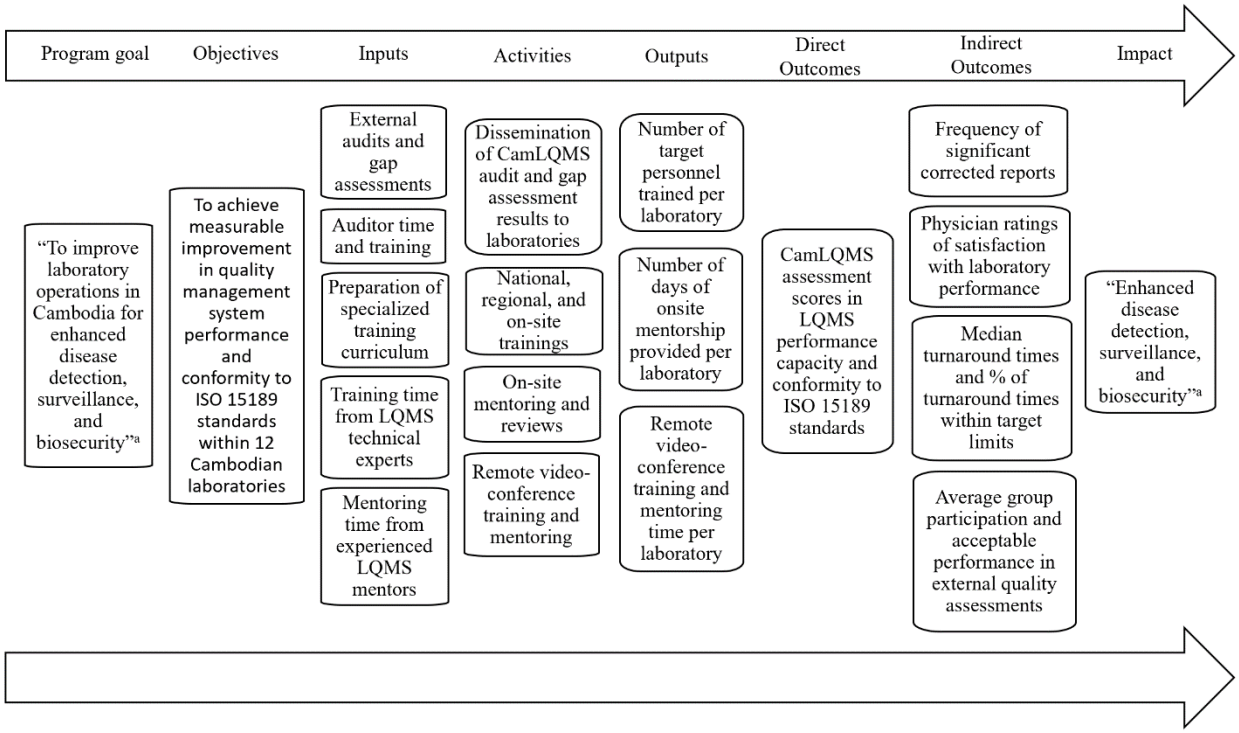

Figure 1. Program evaluation logic model
[a](L.A. Perrone, PhD, unpublished presentation, April 2019)

The logic model demonstrates the hypothesized relationship between activity inputs, outputs, direct outcomes, indirect outcomes, and program impact. The methods presented here quantify indicators of select components of the implementation process by first assessing the program outputs, then improved LQMS capacity as a direct outcome, and then changes to quality indicators as indirect outcomes of program activities. Indirect outcomes are hypothesized to be longer term changes in quality as a result of direct changes in improved LQMS conformity to quality standards and readiness for accreditation.

This study utilized retrospective data to evaluate the impact of the program intervention on laboratory quality management systems, with no elements of the study aimed at generating conclusions about human subjects or their behavior. As such, the University of Washington

Human Subjects Division has determined that it does not include human subjects research and is exempt from further University of Washington IRB review.

Data management and basic descriptive statistics for all evaluation methods were performed using Microsoft Excel for Office 365. All complex calculations of statistics and hypothesis testing were performed using STATA 14 statistical software.

Methods for the evaluation of program activity outputs

For the description and enumeration of activity outputs, the study utilized monitoring and evaluation records collected from I-TECH Cambodia. Outputs of interest, as listed in the logic model in Table 1, include (1) the number of trainings attended by personnel of the intended audience per laboratory, (2) the number of days of onsite mentorship provided per laboratory, and (3) the amount of video conference training and mentoring time recorded during the evaluation period. Materials for output calculations were provided in the form of (1) attendance records for the 12 completed trainings, (2) the I-TECH Cambodia project calendar and mentor site-visit reports, and (3) Zoom usage reports, extracted by I-TECH Cambodia from the team's Zoom accounts for evaluation and reviewed to match laboratory and position details to meeting participants. Supplementary records of remote mentoring were provided to corroborate usage reports and to identify missing information. Datasets from all three material sources were organized into separate spreadsheets for review and quantitative analysis.

The training curriculum implemented by I-TECH Cambodia in partnership with BMLS and the University of British Columbia involved eleven intensive multi-day sessions of quality management and quality assurance training. Attendance records for all 11 training events were organized by meeting date and analyzed for the count of personnel trained who matched the criteria for the intended audience of each training event. The intended audience generally

included laboratory managers and QAOs but at times included directors or administrators, biosafety officers, equipment officers, or stock officers. Attendance entries that met inclusion criteria for the intended audience were included as a count of one intended audience training. Counts were calculated by laboratory and event, and totals and averages were calculated for the group.

On-site visits to laboratories by I-TECH quality mentors were aimed to meet the specific technical needs of laboratories to address the larger gaps in quality management. Records of on-site mentoring were similarly organized into a separate spreadsheet, organized by laboratory visited. On site mentoring reports contained sufficient detail for a summation of the number of unique days on-site for a minimum of one half-day of mentoring with laboratory quality assurance personnel. Totals and averages were similarly calculated for the overall group.

Because program mentors used Zoom Video Conference Calling for most remote mentoring and training, meeting and participant data were automatically recorded through the report feature of the conference tool and available for extraction. These reports were compiled into a dataset including a join time, a leave time, and a duration of participation for each user ID during a meeting as a representative sample of remote mentoring activities. Within this dataset, formal training attendance logs were used to match user IDs with the participants' laboratories and positions where a name was indicated, using mentor reports as supplement records to match and attribute 98% of participation time to participating laboratory personnel, to I-TECH Cambodia staff or mentors, or to other participating stakeholders.

Due to the use of multiple devices by some participants during meetings, a dynamic Gantt Chart was used to visually and methodically identify duplicate, overlapping usernames, the duplicates of which were then recategorized as "device only" regarding position and laboratory

to exclude them from analysis. All user logins that indicated multiple participants associated with a user ID were duplicated to reflect attendance of those participants. Minutes of participation time were grouped by laboratory and summed to calculate total participation time of unique attendees from each laboratory within the sample over the evaluation period. Records of additional remote training or mentoring events held outside of tracked video conferences were reviewed to determine how representative the sample is out of the total estimate of events.

Total video conference participation times per laboratory were plotted in a scatter diagram against the direct outcome of percent differences in pre- and post-intervention audit scores, described under the methods for CamLQMS outcome evaluation, to be visually inspected for a linear or monotonic relationship between the two variables and then tested for the strength of that relationship by using Spearman's rank order correlation coefficient. Spearman's rank was chosen as a non-parametric test due to the small sample size of the intervention group (n=12), which was expected to increase the test sensitivity to moderate outliers in a Pearson's Test for Correlation. Because formal trainings and site visits were restricted from random variation, this study was unable to provide similar correlation assessments between these activity outputs and direct program outcomes.

Methods for the evaluation of CamLQMS audit score changes as a direct program outcome

For the assessment of improved LQMS capacity in participating laboratories as a direct outcome of program activities, the primary data collection tool used was the nationally accepted CamLQMS checklist for accreditation, which is an electronic adaptation of the WHO AFRO SLIPTA audit tool[11]. The CamLQMS checklist is divided into 12 sections of laboratory quality and 117 questions, each assigned a numerical value that contributes to the audit score within

each section and in the whole, as shown and described in the example scorecard in Appendix A (BMLS, unpublished document, 2017). Auditors were thoroughly trained by I-TECH and BMLS to identify areas of conformity and nonconformity to CamLQMS standards.

Mentored LQMS laboratories completed baseline CamLQMS audits in January of 2018 and outcome assessments in April of 2019 for a case-series analysis of the improvement of LQMS compliance scores over time. Additionally, a comparison group of representative non-mentored, non-LQMS laboratories was selected by BMLS at the request of I-TECH for an additional cross-sectional study, with the requested criteria that all comparison laboratories should be selected from similar hospitals in terms of the complementary package of activities or services (CPA) offered. The complete list of laboratories included in this study is shown in Table 1, including the demographics of each laboratory's workforce, inpatient capacity, workload, and CPA category of services provided by the laboratory hospital.

Table 1 *. List of laboratories included in 2019 CamLQMS audit assessments, including demographics of laboratory workload and capacity provided by each laboratory*

Primary partnership support	Laboratory name	# of lab staff	Inpatient occupancy capacity (# of beds)	Average inpatient occupancy (% of beds used)	Average number of daily OPD[a] visits	Average number of lab tests per day	CPA[b] capacity type
I-TECH/UW	Prey Veng Provincial Referral Hospital Laboratory	9	200	80	35	200	CPA-3
	Svay Rieng Provincial Hospital Laboratory	13	168	133	84	173	CPA-3
	Preah Sihanouk Provincial Referral Hospital Laboratory	10	160	100	83	90	CPA-3
	Kampot Provincial Referral Hospital Laboratory	12	155	112	119	429	CPA-3
	Preah Kossamak Hospital Laboratory	36	400	85	11	914	CPA-3 National
	Takeo Provincial Referral Hospital Laboratory	15	250	110	84	311	CPA-3
	Chey Chumneas Referral Hospital Laboratory (Kandal)	12	190	85	35	250	CPA-3
	Kratie Provincial Referral Hospital Laboratory	10	150	86	70	115	CPA-3
	Kampong Cham Provincial Referral Hospital Laboratory	19	260	110	35	500	CPA-3
	Khmer Soviet Friendship Hospital Laboratory (KSFH)	33	600	140	1052	1614	CPA-3 National
	Preah Ang Duong Hospital Laboratory	14	260	100	700	1000	CPA-3
	National Pediatric Hospital Laboratory	28	300	120	210	600	CPA-3 National
Asian Development Bank	Pea Reang Referral Hospital Laboratory	4	130	100	40	80	CPA-2
	Neak Loeung Referral Hospital Laboratory	4	120	50	120	110	CPA-2
	Kampong Trabek Referral Hospital Laboratory	4	120	30	45	60	CPA-2
	Romeas Heak Referral Hospital Laboratory	4	70	137	30	40	CPA-2
	Kirivong Referral Hospital Laboratory	8	120	140	120	250	CPA-2
	Koh Thom Referral Hospital Laboratory	6	74	65	10	30	CPA-2
	Tbong Khmum Referral Hospital Laboratory	5	90	110	No OPD[a]	40	CPA-2
	Ponhea Krek Referral Hospital Laboratory	8	80	96	10	200	CPA-2
	Kampong Trach Referral Hospital Laboratory	8	70	50	50	15	CPA-2
	Angkor Chey Referral Hospital Laboratory	7	60	57	30	20	CPA-2

[a]*OPD visits is short for outpatient department visits*
[b]*CPA is short for the complementary package of activities or services provided by the hospital*

Within the two groups studied, CPA-2 hospitals generally have 60-100 beds, basic obstetric care, emergency care, major surgery, blood transfusion, and other specialized services, while CPA-3 hospitals generally have a capacity of 100-250 beds and additional specialized services to those of CPA-2 hospitals[15]. Because most CPA-3 hospital laboratories in Cambodia are now implementing LQMS improvement programs as part of the established I-TECH or SLMTA initiatives, all comparison laboratories in this list are of the CPA-2 category.

Both sets of 2018 and 2019 audit results have been provided for analysis in this study, from which audit scores were calculated as a percentage of the total added value of checklist

items for each section and overall for each laboratory. A Wilcoxon Sign-rank non-parametric comparison for paired samples was performed to determine the strength of the difference between 2018 and 2019 audit scores of mentored LQMS laboratories for each section and total overall. Mean audit scores and standard deviations were calculated by section for a visual presentation of comparisons between laboratory groups, and all sections with statistically significant differences between years were documented with the level of significance.

Simple calculations of the percent difference in overall audit scores between years were measured in each section to present the magnitude of change visually. These percent differences were used as the primary outcome variables for a Spearman's rank correlation assessment of the relationship strength between audit score improvement and laboratory personnel participation time in video conference activities.

An assessment of the statistical difference between audit scores of mentored LQMS laboratories and non-mentored, non-LQMS comparison laboratories was performed using the Wilcoxon Rank-sum test for two independent samples. Comparisons were made for overall audit scores and scores for individual audit sections, and all sections with statistically significant differences between groups were again documented with the level of significance.

The decision to use non-parametric tests over parametric alternatives was based on tests for basic assumptions for parametric testing and ultimately was determined by tests for normality within the distributions of the variables. A Shapiro-Wilks test for normality was performed for both pre- to post-audit comparisons and intervention to non-intervention comparisons of overall scores, then for comparisons of scores within individual audit sections. Although normality was found for overall audit score distributions in both sets of comparisons, the same could not be

found for comparisons of individual sections between intervention and non-intervention laboratories. For consistency, nonparametric tests were chosen for both comparisons.

Methods for the evaluation of quality indicators as indirect program outcomes

To assess whether quality has changed during the evaluation period as an indirect outcome of program activities, a selection of common quality indicators was chosen to comprehensively detect changes in different aspects of quality. The use of quality indicators is a requirement of ISO 15189 accreditation standards, and has long been used as a tool both for monitoring quality for corrective action and for monitoring and driving improvement[16,17]. Indicators were selected based on their ability to detect changes in different areas of laboratory quality, on the quality of data from laboratories, and on the availability of retrospective data within a relevant timeframe to the evaluation period. The following indicator measures are presented with their respective analysis methods as follows:

Physician satisfaction surveys

The use of physician satisfaction surveys as a quality indicator of laboratories is one of the standards and requirements of ISO 15189 accreditation and measures timeliness and effectiveness of laboratory services[16,18]. Physician satisfaction ratings on a Likert scale of 1 to 5 were measured using a brief physician satisfaction survey that was distributed to health providers being served by the 12 participating hospital laboratories. In contrast to other quality indicators used in this study, a baseline dataset was not available for a proximal timepoint to the evaluation period. Baseline data were collected instead from a former survey that was successfully collected by I-TECH in 2015 and a smaller version of this same survey was distributed again in February 2019 in the national language, Khmer, with the same methodology. An English translation of the condensed 2019 survey can be seen in Appendix B. Anonymous surveys for both years were

collected and assigned a unique identifier by I-TECH mentors. These surveys were then assigned to a translator for all free form entry to be translated into English, then all data were entered into a spreadsheet for analysis. In 2019, response data from 11 of the 12 laboratories were collected successfully and combined with 2015 data to be compared by year.

An exclusion criterion was required in the protocol to exclude ambiguous or incomplete data. Respondents who failed to list their positions or who listed a position in their hospital as anything other than an ordering healthcare provider were excluded from the study. Questions were analyzed independently, excluding missing or ambiguous responses from the analysis of each variable. Responses that were left blank, incomplete, or answered ambiguously were removed from data analysis and response totals. The shortened survey included two categories of Likert scale questions: those related directly to satisfaction and those related to physician utilization of services. As a measure of physician satisfaction, however, only those questions related to satisfaction (Appendix B, questions 4-5) were included in the final analysis for the physician satisfaction indicator.

Means and standard deviations (SD) were calculated for all included variables in baseline and endpoint surveys, and median ratings were calculated for Likert scale variables for each category of physician satisfaction. An eighth category was generated as a composite measure of overall satisfaction from an average of satisfaction ratings provided by each respondent in each category. Counts and response frequencies of each of the five Likert scale ratings were then calculated and plotted in a stacked bar graph to present the percent frequency of each rating out of the total number of responses for each category. These frequencies were analyzed for obvious differences between years. Because of the ordinal nature of the data, the non-parametric, two-sample Wilcoxon Rank-sum (Mann-Whitney) test was used to determine whether physician

satisfaction ratings had changed significantly in each category between surveys. Additionally, a non-parametric Spearman's rank order correlation test was performed to determine whether a significant relationship exists between the variations in the composite measure of overall physician satisfaction and overall 2019 CamLQMS audit scores in the 11 laboratories for which complete both sets of data are available.

Corrected report frequencies

Corrected report frequency is another common indicator of quality in medical laboratories in regard to patient safety[18]. This study aimed to assess corrected report frequencies over time as an indicator of analytical error in participating laboratories, identifying errors that were made but not identified until after release of the initial reported result, and determining whether corrected report frequencies decreased or improved over a segment of the mentored LQMS implementation period. Past studies have shown that errors identified from corrected reports are associated with adverse outcomes for patients, and laboratories are encouraged to investigate corrected reports, identifying root causes of analytical error in the interest of quality improvement[19]

Data used for corrected report frequency analysis were extracted from the Cambodia Laboratory Information System (CamLIS), which is a web based platform that allows laboratories to record patient results in an electronic database synchronized to the MOH's national Health Management Information System[7]. CamLIS was developed and is continually being improved and maintained by the WHO office in Cambodia, who have extracted and provided data for the benefit of this study with permission from the Cambodia MOH. CamLIS is functional in 11 of the 12 participating laboratories and has been capable of collecting and storing corrected report data since mid-July of 2018.

Six complete months of corrected report data were provided with the laboratory name, the test type, the original and corrected results, the normal reference range, and the dates of the original and corrected report entries. Corrected report frequencies were analyzed from a subset of 31 common biochemistry and hematology tests, selected to calculate a representative mean frequency of corrected report occurrences per month. A separate dataset containing the number of patients tested per month for each respective test type was used as the denominator for a frequency of significant corrected reports made per 1000 patients per month.

The dataset was filtered to include only those corrected reports which had changed interpretation between normal and abnormal results upon correction. Remaining reports, which had changed in interpretation, were defined as significant corrections for the purpose of this study. This definition is lenient by more ideal standards that would compare values to laboratory critical ranges and whether the subsequent interpretations by clinicians were impacted by the initial error[19]; however, the criteria are limited here to the interpretation of data relative to the reference range. Data for more stringent criteria was unavailable for this study.

From corrected reports that met the defined criteria, significant corrected error frequencies were calculated as the mean number of corrected reports per 1000 patients tested per month. These frequencies were then plotted over time and assessed for monotonic trends using a locally weighted scatterplot smoothing (LOWESS) curve. A Spearman's rank order test for correlation was performed to determine the strength of any monotonic trends in mean corrected report frequencies over time for individual laboratories and for the overall group.

Median test turnaround times and the percentage of turnaround times within standard limits

Changes in median turnaround times (TAT) and the percentage of TAT within target limits were analyzed as an indicator of timeliness in participating laboratory testing. Data for analysis were

provided again by the WHO office in Cambodia, extracted from CamLIS and provided with the receipt and completion times for all creatinine and CBC orders during the evaluation period, which were used to calculate a TAT for each test completed. Creatinine and CBC were chosen as surrogate markers of TAT because they are two of the most commonly ordered tests for core biochemistry and hematology[20].

Prior to analysis of TATs for participating laboratories, an inquiry was made as to the record keeping practices of the laboratories to determine the quality of the data. I-TECH Cambodia mentors reported that many of the laboratories did not yet meet accurate standard practices for recording receipt times of samples immediately upon arrival. Rather, technicians have been known to record the receipt times immediately before reporting results. Of the 11 participating laboratories with available sample receipt and reporting time data, three were recommended as having sufficiently accurate recording practices for inclusion in the study. Laboratory TATs were further screened for data quality by plotting frequency distributions of TATs in histograms to confirm that distributions of included laboratories demonstrated a non-gausian, unimodal, positively skewed curve, typically expected in TAT analysis[21]. In addition, because testing for creatinine requires an initial processing time of five to ten minutes according to required centrifugation times and minimum time on the instruments, a high frequency of turnaround times within this processing period was considered indicative of falsified recording practices, excluding the laboratory from the analysis.

Median monthly TATs and the monthly percentage of TATs within routine and rapid turnaround time targets were calculated and plotted in line graphs over time for the evaluation period between January 2018 and February 2019. Plots were evaluated for monotonicity of each

category of data over time with the intent to test for monotonic trends using a Spearman's rank correlation coefficient if a monotonic trend was suspected.

Proficiency testing participation and compliance over the evaluation period

Proficiency testing (PT) or external quality assessment (EQA) is a common indicator of patient safety and accurate testing and is another requirement of ISO 15189 accreditation standards[14,18]. Past studies have demonstrated the relationship between proficiency test performance and testing accuracy, though it is widely considered limited due to a tendency to measure only the best possible outcomes rather than a random sample of outcomes[22].

Cambodia has a national EQA program, largely supported by external funding and implementing organizations, consisting of assessments in microbiology, hematology, biochemistry, serology, and blood banking two to three times per year. Results for all participating laboratories between 2012-2018 were provided for this evaluation by the Cambodia National Institute of Public Health (NIPH), who serve as the focal point for distribution of EQA materials and for the collection of assessment responses from laboratories. This analysis is used conservatively to present whether intervention laboratories are improving in compliance with EQA standards over time by assessing the percentage of laboratories that are participating in EQA and that have achieved 80% or better in EQA assessments, and by assessing the mean EQA scores of I-TECH mentored LQMS laboratories over time compared to the mean scores of all other Cambodia laboratories.

In order to remain in compliance with national standards for a given period, laboratories are expected to pass EQA assessments with a score of 80% or greater in any given round of assessments. As an internal indicator of improved quality, individual laboratory and overall interlaboratory mean performance scores in Hematology, Biochemistry, Serology, and

Microbiology testing were organized by EQA implementation period and assessed for a passing score of 80% or greater in participating laboratories. A percentage of the 12 laboratories participating and passing EQA per assessment period was calculated for each of the six EQA programs for which data was made available. These percentages for each EQA category were plotted over the chronological sequence of EQA rounds and a linear regression was fitted to the data for all categories to present changes in group EQA compliance over time.

To assess whether mean EQA scores of I-TECH mentored laboratories have improved over time compared to all other Cambodia laboratories assessed through the national EQA program, scores for all I-TECH mentored LQMS laboratories were averaged per round of implementation for each category and overall. The same calculations were made separately for all other participating Cambodia laboratories to create an independent comparison group and the average scores for each group were plotted in sequence for trend analysis. All scores of 0% due to a laboratory's failure to participate during an assessment period were excluded from this analysis in order to reflect the laboratories' actual performance. EQA scores for blood banking were excluded due to the extremely small sample size of one blood bank capable laboratory within the I-TECH LHSS program.

Results

Results for program output and outcome evaluation

Table 2 shows the numbers of outputs for all three measured program activities and the corresponding increase in CamLQMS audit score as the direct program outcome. An output of 274 (mean=23±2) target personnel trainings, 72,321 (mean=6027±2454) minutes of video conference training, and 130 (mean=11±2) visits to laboratories resulted in an average positive percent difference of 23±12% between overall scores of the 2018 and 2019 audits. Video

conference participation time was calculated from a sample of 153 Zoom meetings with trackable usage reports out of a total of 261 meetings identified from supplemental mentor reports and program calendar entries. In terms of staff inputs, formal training and video conference activities included two primary mentors, two mentor trainees, the country project coordinator, and three laboratory systems technical and senior technical specialists. Additionally, several government officials from BMLS and the National Institute of Public Health in Cambodia collaborated in the implementation of all formal trainings and in numerous video conference activities.

Table 2. *Individual laboratory and group mean calculations of the three primary activity outputs and the overall outcome achieved within the evaluation period*

	Number of completed trainings of intended participants (total participants)	**Mentor days on site per laboratory**	**Video conference participation time (minutes)**	**Audit score percent difference**
Prey Veng	19 (25)	9	3766	4%
Svay Rieng	21 (26)	10	5855	9%
Preah Sihanouk	24 (25)	10	2742	14%
Kampot	23 (25)	10	6320	31%
Preah Kossamak	25 (29)	13	9302	43%
Takeo	24 (37)	13	9664	37%
Kandal	23 (24)	9	6800	25%
Kratie	21 (27)	13	5290	26%
Kampong Cham	28 (36)	13	7210	26%
KSFH	22 (24)	8	4434	29%
Preah Ang Duong	22 (26)	10	8675	15%
National Pediatric	22 (26)	12	2263	6%
Group mean ± SD	23± 2 (28±4)	11±2	6027±2454	22%±12%

Table 3 lists the overall and individual section scores of intervention laboratories for 2018 and 2019 CamLQMS audits with a color scale to visually highlight and scale audit scores from low (red), to medium (yellow), to high (green). The table reveals a large percent difference between assessments and a high amount of variation between laboratories. A Wilcoxon Signed-Ranks

Test indicated that overall audit scores for LQMS program laboratories in 2019 (median score=56.5%) were significantly higher than overall audit scores for the same laboratories in 2018 (median score=40.0%), $z=3.06$, $p=0.002$. In a comparison of scores for individual audit sections between years, the Wilcoxon Signed-Ranks Test indicated that mean 2019 scores for 11 out of 12 audit sections have improved significantly ($p<0.01$), with *information management* being the exception, which had been maintained but not improved from an already high performance at baseline. Table 3 indicates sections for which differences are statistically significant. A bar graph for a visual comparison of overall audit scores between 2018 and 2019 for individual laboratories is presented in Figure 2, including overall scores of non-mentored, non-LQMS laboratories for additional comparison. A bar graph of calculated mean scores for each section in all comparison groups is presented in Figure 3.

Table 3. *Summary of 2018 and 2019 CamLQMS audit scores overall and by individual section*

		Documents and records[a]	Management reviews[a]	Organization and personnel[a]	Client management and customer service[a]	Equipment[a]	Evaluation and audits[a]	Purchasing and inventory[a]	Process control and internal and external quality assessment[a]	Information management	Corrective action[a]	Occurrence management and process improvement[a]	Facilities and safety[a]	Overall score[a]
2019 mentored LQMS laboratories	Prey Veng	46%	14%	41%	40%	33%	13%	46%	50%	87%	5%	17%	44%	38%
	Svay Rieng	64%	14%	45%	60%	55%	13%	58%	57%	87%	16%	42%	70%	52%
	Preah Sihanouk	54%	43%	50%	70%	33%	40%	50%	47%	73%	16%	25%	60%	47%
	Kampot	82%	43%	68%	70%	79%	47%	54%	73%	60%	26%	50%	74%	65%
	Preah Kossamak	75%	71%	100%	90%	82%	100%	96%	94%	93%	79%	100%	84%	85%
	Takeo	61%	93%	100%	90%	82%	100%	83%	97%	100%	79%	100%	88%	88%
	Kandal	57%	21%	45%	50%	64%	7%	50%	57%	100%	11%	17%	53%	48%
	Kratie	68%	36%	45%	70%	58%	20%	67%	60%	42%	37%	58%	72%	56%
	Kampong Cham	71%	100%	100%	90%	82%	100%	96%	91%	100%	79%	100%	98%	91%
	KSFH	82%	57%	82%	60%	76%	73%	88%	57%	47%	21%	75%	70%	67%
	Preah Ang Duong	46%	14%	55%	70%	61%	20%	67%	66%	73%	47%	67%	72%	57%
	National Pediatric	61%	14%	45%	70%	61%	47%	75%	66%	73%	16%	33%	60%	55%
2018 mentored LQMS laboratories	Prey Veng	43%	7%	36%	30%	31%	13%	46%	50%	42%	0%	0%	51%	29%
	Svay Rieng	54%	7%	41%	50%	37%	7%	50%	44%	90%	0%	8%	65%	38%
	Preah Sihanouk	50%	7%	36%	20%	23%	13%	38%	38%	90%	0%	0%	37%	29%
	Kampot	43%	7%	36%	30%	26%	7%	33%	33%	95%	0%	0%	47%	30%
	Preah Kossamak	54%	7%	45%	60%	29%	7%	75%	47%	90%	5%	8%	47%	39%
	Takeo	50%	7%	45%	60%	54%	13%	71%	56%	90%	21%	8%	67%	45%
	Kandal	32%	7%	32%	0%	14%	0%	42%	30%	47%	0%	0%	26%	19%
	Kratie	43%	7%	27%	20%	20%	13%	33%	38%	53%	0%	0%	49%	25%
	Kampong Cham	50%	7%	45%	90%	71%	47%	88%	72%	90%	58%	17%	86%	60%
	KSFH	39%	7%	36%	30%	37%	7%	54%	53%	67%	5%	0%	51%	32%
	Preah Ang Duong	36%	7%	41%	50%	49%	13%	42%	50%	90%	0%	8%	60%	37%
	National Pediatric	46%	7%	55%	80%	43%	20%	75%	50%	81%	5%	17%	63%	45%

A color scale is used to visually highlight and scale audit scores from low (red), to medium (yellow), to high (green) variations

[a]Audit score difference between comparison groups for indicated categories is highly significant at the $p<0.01$ significance level

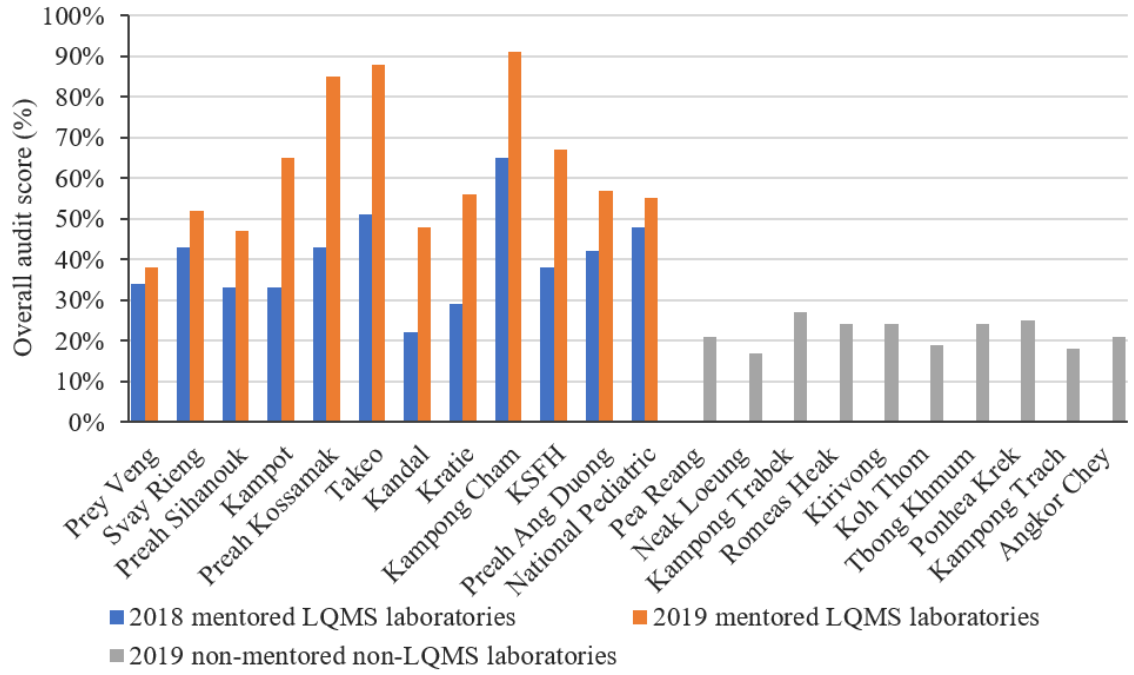

Figure 2. Overall 2018-2019 CamLQMS audit scores for intervention and comparison laboratories. No 2018 non-mentored, non-LQMS baseline audits were available for analysis in this study.

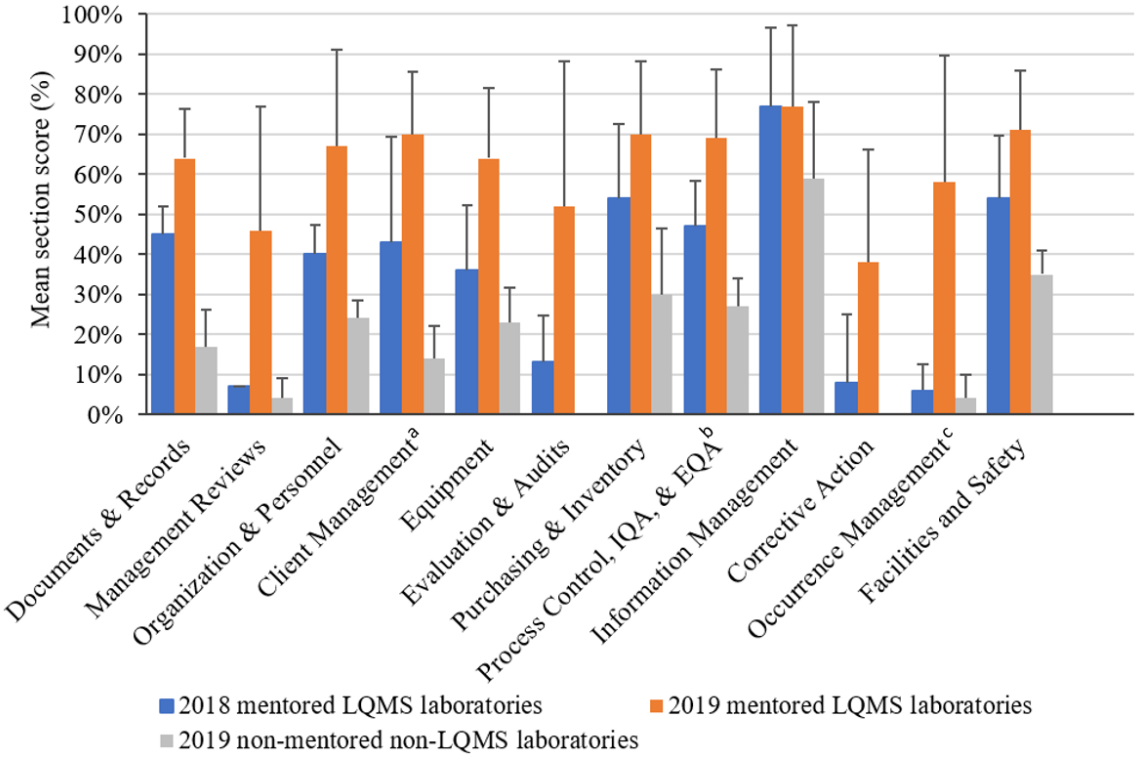

Figure 3.Mean 2019 CamLQMS audit scores for participating mentored LQMS program laboratories compared to 2018 baseline and comparison laboratories. Error bars represent the absolute standard deviations from the mean score of each section.

[a] *Client Management & customer service*
[b] *Process Control & Internal and external quality assessment*
[c] *Occurrence management and process improvement*

In a cross-sectional comparison of the 2019 audit performance of mentored LQMS laboratories to the sample of non-mentored, non-LQMS comparison laboratories, Table 4 again provides a summary of the overall and individual section scores of laboratories in both groups, color scaled for visual comparison. The table again shows a large contrast in scores between the two groups with a high amount of variation between laboratories. A two-sample Wilcoxon rank-sum (Mann-Whitney) test indicated that overall audit scores for mentored LQMS laboratories in 2019 (median=57%) were significantly higher than overall audit scores for non-mentored, non-LQMS laboratories (median=23%)in the same year ($z=3.96$, $p=0.0001$). Bar graphs of overall audit scores for non-intervention laboratories are visualized in Figure 2 with those of the intervention group. Mann-Whitney tests for individual audit section comparisons similarly found significant differences in 11 of the 12 sections ($p<0.001$) between intervention and non-intervention laboratories, *information management* again being the exception, which was significantly different at the $p<0.05$ level. Figure 3 presents a bar graph visualization of individual section scores for comparison between intervention and non-intervention groups. In terms of percent difference in mean section scores between groups, client management and customer service as well as occurrence management and process improvement demonstrated the largest differences of 56% and 54% between groups. Section 9, information management, again showed the smallest percent difference between groups of 18%.

Table 4. *Summary of CamLQMS audit scores overall and by individual section for both 2019 mentored LQMS and non-mentored, non-LQMS laboratories*

		Documents and records[b]	Management reviews[b]	Organization and personnel[b]	Client management and customer service[b]	Equipment[b]	Evaluation and audits[b]	Purchasing and inventory[b]	Process control and internal and external quality assessment[b]	Information management[a]	Corrective action[b]	Occurrence management and process improvement[b]	Facilities and safety[b]	Overall score[b]
2019 mentored LQMS laboratories	Prey Veng	46%	14%	41%	40%	33%	13%	46%	50%	87%	5%	17%	44%	38%
	Svay Rieng	64%	14%	45%	60%	55%	13%	58%	57%	87%	16%	42%	70%	52%
	Preah Sihanouk	54%	43%	50%	70%	33%	40%	50%	47%	73%	16%	25%	60%	47%
	Kampot	82%	43%	68%	70%	79%	47%	54%	73%	60%	26%	50%	74%	65%
	Preah Kossamak	75%	71%	100%	90%	82%	100%	96%	94%	93%	79%	100%	84%	85%
	Takeo	61%	93%	100%	90%	82%	100%	83%	97%	100%	79%	100%	88%	88%
	Kandal	57%	21%	45%	50%	64%	7%	50%	57%	100%	11%	17%	53%	48%
	Kratie	68%	36%	45%	70%	58%	20%	67%	60%	42%	37%	58%	72%	56%
	Kampong Cham	71%	100%	100%	90%	82%	100%	96%	91%	100%	79%	100%	98%	91%
	Khmer Soviet Friendshi	82%	57%	82%	60%	76%	73%	88%	57%	47%	21%	75%	70%	67%
	Preah Ang Duong	46%	14%	55%	70%	61%	20%	67%	66%	73%	47%	67%	72%	57%
	National Pediatric	61%	14%	45%	70%	61%	47%	75%	66%	73%	16%	33%	60%	55%
2019 non-LQMS comparison laboratories	Pea Reang	11%	7%	18%	0%	24%	0%	25%	33%	64%	0%	0%	37%	21%
	Neak Loeung	21%	7%	27%	10%	24%	0%	17%	17%	46%	0%	0%	19%	17%
	Kampong Trabek	29%	0%	27%	10%	36%	0%	71%	18%	64%	0%	0%	35%	27%
	Romeas Heak	11%	0%	23%	10%	30%	0%	42%	25%	91%	0%	0%	40%	24%
	Kirivong	25%	0%	27%	30%	24%	0%	17%	28%	67%	0%	17%	37%	24%
	Koh Thom	18%	0%	18%	10%	15%	0%	42%	17%	55%	0%	0%	30%	19%
	Tbong Khmum	29%	7%	27%	10%	33%	0%	25%	33%	40%	0%	8%	33%	24%
	Ponhea Krek	25%	0%	27%	20%	30%	0%	21%	30%	85%	0%	0%	35%	25%
	Kampong Trach	0%	7%	23%	20%	15%	0%	33%	18%	37%	0%	0%	35%	18%
	Angkor Chey	18%	14%	32%	10%	9%	0%	38%	28%	37%	0%	8%	33%	21%

A color scale is used to visually highlight and scale audit scores from low (red), to medium (yellow), to high (green) variations

[a]Audit score difference between comparison groups for indicated categories is significant at the $p<0.05$ significance level

[b]Audit score difference between comparison groups for indicated categories is very highly significant at the $p<0.001$ significance level

In an assessment of the relationship between mean audit score differences in 2018-2019 assessments and the amount of participation time by individual laboratories in zoom video conference training a Spearman's rank correlation showed a strong monotonic relationship between the two variables ($r_s=0.66$, $p=0.02$) with significant certainty. A locally weighted

scatterplot smoothing (bandwidth=0.8) presents the monotonic but non-linear relationship between the variables in Figure 4.

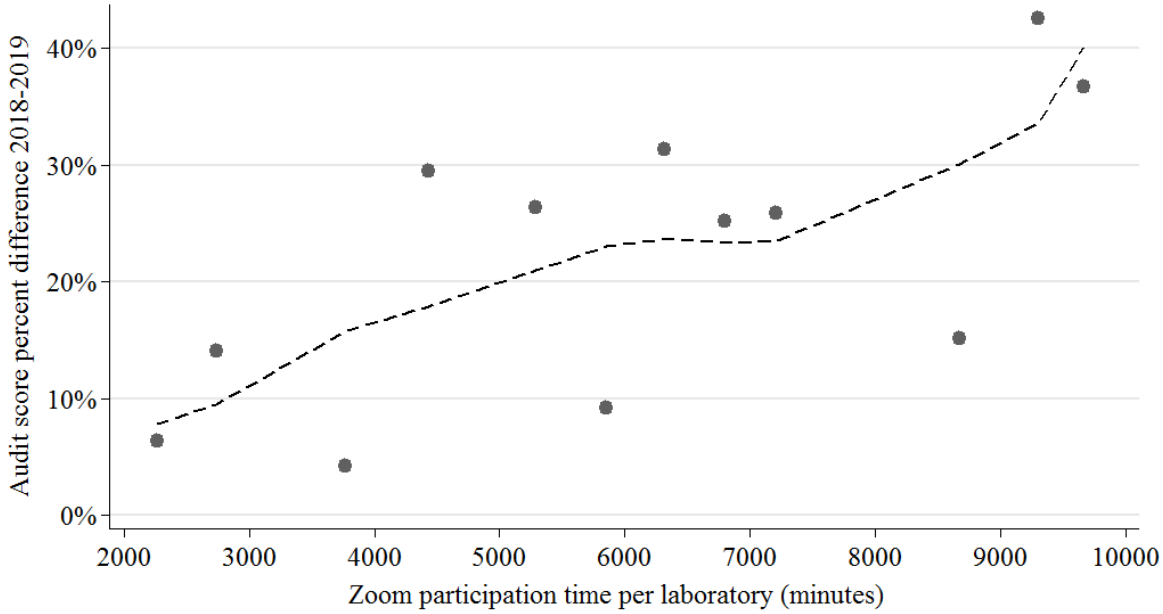

Figure 4. *Scatter gram of the percent difference in mentored LQMS audit scores and video conference participation time by lab. Spearman's rank coefficient r_s=0.66, p=0.02. The line represents a locally weighted scatterplot smoothing curve to represent the monotonic but non-linear relationship of the variables (bandwidth=0.8).*

Results for quality indicator changes as indirect outcome measurements

Physician satisfaction survey analysis

Respondents did not vary noticeably between years regarding the proportions of physicians indicating positions of leadership or not. Average physician estimates of the percentage of patients who require laboratory testing, interestingly, remained almost unchanged from 62±34% in 2015 to 63±33% in 2019, whereas physician estimates of the number of patients seen per day increased from 13±9 in 2015 to 17±15.

Response frequencies for physician satisfaction survey categories are organized in Table 5, including the total counts of each rating per category and the total frequency of ratings out of the total number of responses provided. The numbers of responses, the mean ratings, and the median

ratings are included for reference. Figure 5 presents Likert scale response frequencies in a stacked bar graph, with each rating representing a proportional frequency out of 100% of the total responses.

A two-sample Wilcoxon Rank-sum (Mann-Whitney) test indicated that median Likert-scale ratings of satisfaction with laboratory services did not change significantly from baseline 2015 ratings in the service categories of result quality ($z=0.80$, $p=0.4$) and service ($z=1.4$, $p=0.2$). Median ratings for all other service categories of ease of test ordering, good communication, turnaround time, timely calling of critical results, result corroboration with clinical findings, and the combined measure of overall satisfaction, however, have been found to have significantly decreased significantly since 2015 ($p<0.05$). In terms of changes in frequencies of ratings, survey categories of turnaround time and timely calling of critical results demonstrated the largest percent difference of good ratings (fours and fives) between surveys, decreasing by a percent difference of -23% and -26%. These two categories similarly presented the highest increase in poor rating frequencies (ones and twos), with percent differences of 17% and 8%.

Table 5*. Physician satisfaction response frequencies for all service categories in both years. Categories are paired by year for direct comparison.*

Survey Response Category	No. of ratings (N)	Mean rating	Median rating	Rating 5 No. (%)	Rating 4 No. (%)	Rating 3 No. (%)	Rating 2 No. (%)	Rating 1 No. (%)
2015 Ease of test ordering	157	3.9	4	50 (32%)	50 (32%)	45 (29%)	8 (5%)	4 (3%)
2019 Ease of test ordering	329	3.6	4	30 (9%)	165 (50%)	119 (36%)	11 (3%)	4 (1%)
2015 Quality results	155	3.7	4	31 (20%)	62 (40%)	51 (33%)	10 (6%)	1 (1%)
2019 Quality results	327	3.7	4	25 (8%)	183 (56%)	101 (31%)	17 (5%)	1 (0%)
2015 Good turnaround time	155	3.7	4	36 (23%)	49 (32%)	57 (37%)	8 (5%)	5 (3%)
2019 Good turnaround time	326	3.1	3	20 (6%)	83 (25%)	140 (43%)	62 (19%)	21 (6%)
2015 Good communication	156	3.8	4	45 (29%)	53 (34%)	47 (30%)	7 (4%)	4 (3%)
2019 Good communication	327	3.6	4	43 (13%)	156 (48%)	100 (31%)	23 (7%)	5 (2%)
2015 Good Service	156	3.7	4	39 (25%)	47 (30%)	56 (36%)	11 (7%)	3 (2%)
2019 Good Service	328	3.6	4	38 (12%)	144 (44%)	116 (35%)	26 (8%)	4 (1%)
2015 Timely calling of criticals	162	3.9	4	54 (33%)	59 (36%)	36 (22%)	9 (6%)	4 (2%)
2019 Timely calling of criticals	321	3.4	3	36 (11%)	106 (33%)	129 (40%)	40 (12%)	10 (3%)
2015 Labs match clinical findings	163	3.8	4	43 (26%)	58 (36%)	49 (30%)	10 (6%)	3 (2%)
2019 Labs match clinical findings	325	3.5	4	29 (9%)	141 (43%)	130 (40%)	22 (7%)	3 (1%)

Ratings for timely calling of critical results and whether labs match clinical findings are based on the question of how well physicians agree with the statement in the question, where 1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5=strongly agree. All other questions ask physicians to rate laboratory service categories on a scale of 1=very poor, 2=poor, 3=average, 4=good, and 5=very good.

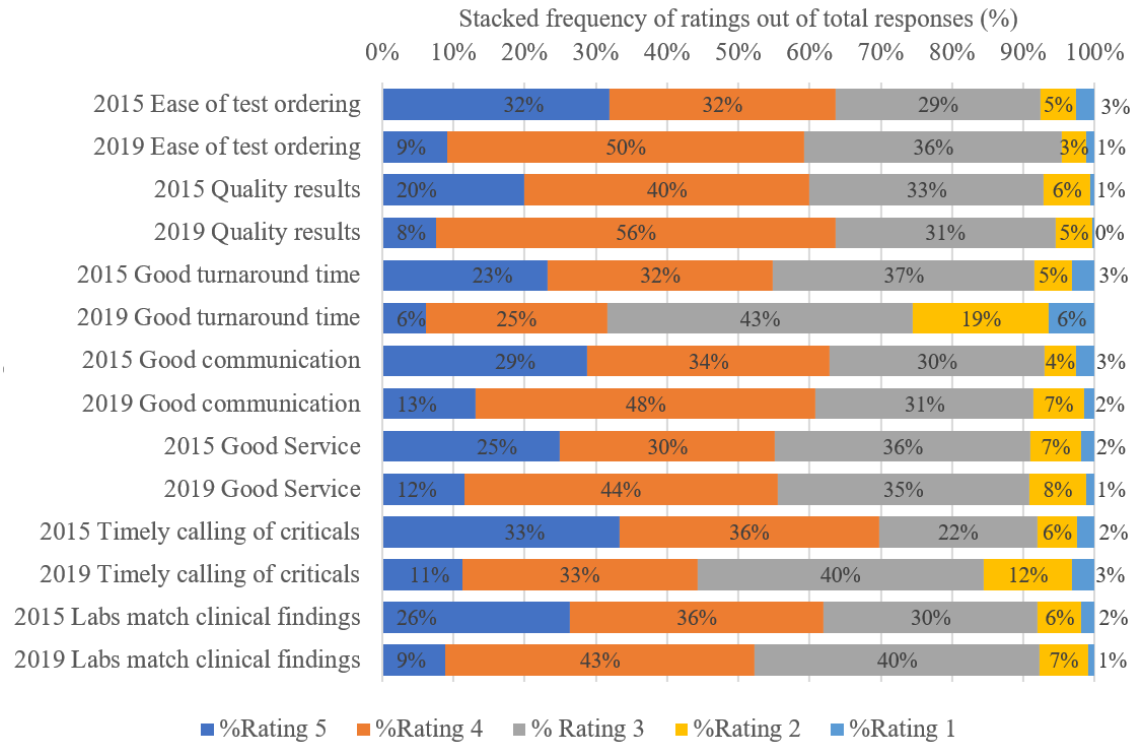

Figure 5. Stacked bag graph of individual rating frequencies per physician satisfaction service category

To answer the question of whether variations in current physician satisfaction ratings for individual laboratories are related to current laboratory performance in CamLQMS audits, a spearman rank test for correlation found no significant relationship (r_s=-0.22, p=0.52) when using the composite mean indicator for overall physician satisfaction as the dependent variable.

Corrected report frequency analysis

A summary of the mean corrected report frequencies for individual laboratories and the overall group is presented in Table 6. Of the individual laboratories analyzed for monotonic trends over time, only two laboratories, Prey Veng (r_s=-1, p=<0.001) and Kratie (r_s=-0.83, p=0.04)

provinces, showed clear monotonic reduction in corrected report frequencies at significant levels over the six-month period. The overall mean corrected report frequency, however, along with all other individual laboratory frequencies, did not show monotonic trends with any statistical significance. Graphical representations of the mean corrected report frequencies over time are presented in Figures 6 and 7 for Kratie and the overall mean as examples.

Table 6. Mean corrected report frequencies for intervention laboratories between August 2018-January 2019. Spearman statistics are listed to present whether any observed trend can be considered significant.

	Month						Spearman r_s	
Laboratory	**Aug-2018**	**Sep-2018**	**Oct-2018**	**Nov-2018**	**Dec-2018**	**Jan-2019**	**r_s**	**p**
Prey Veng	4.3	0.37	0.058	0.30	0.21	0.034	-0.83	0.04
Svay Rieng	13	55	51	92	26	130	0.6	0.2
Preah Sihanouk	2.0	0.79	2.3	1.7	0.66	0.21	-0.71	0.1
Kampot	1.8	1.1	1.1	1.4	2.0	1.0	-0.26	0.6
Preah Kossamak	320	480	370	330	330	320	-0.086	0.9
Takeo	16	9.7	7.6	7.7	1.9	37	-0.086	0.9
Kandal	0.84	1.4	4.8	0.93	0.46	0.82	-0.54	0.3
Kratie	3.4	0.83	0.52	0.25	0.18	0.051	-1.00	<0.001
Kampong Cham	35	24	33	19	37	25	-0.03	1.0
Preah Andoung	290	270	260	230	340	330	0.37	0.5
National Pediatric	7.3	6.2	9.7	3.6	2.8	6.1	-0.66	0.2
Mean±SD	63±120	77±155	67±125	62±112	67±133	78±130	0.43	0.4

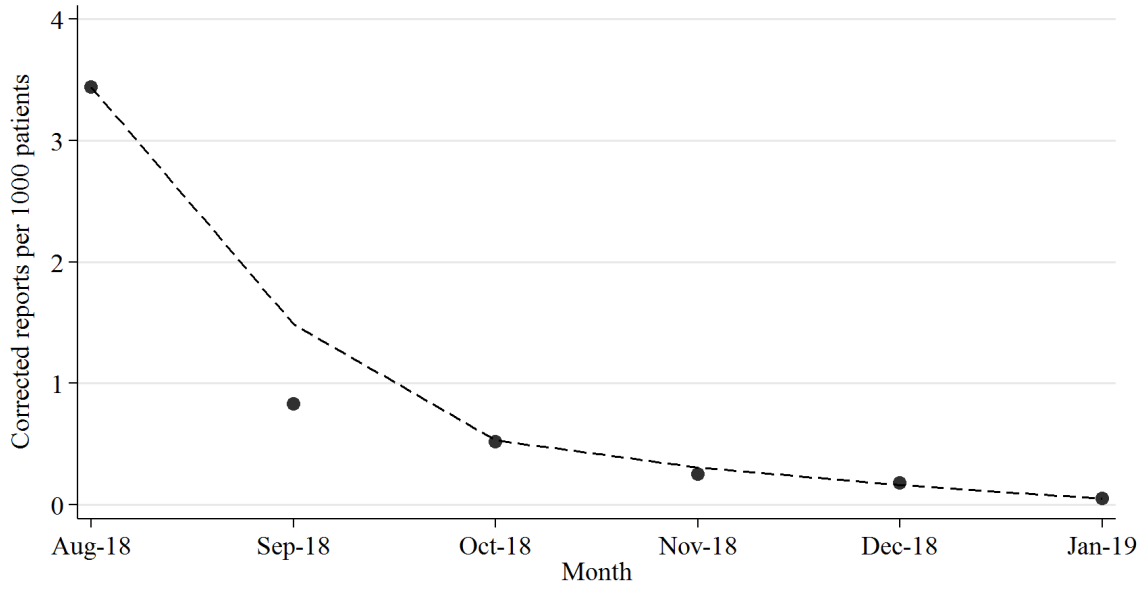

Figure 6. *Kratie Provincial Regional Hospital mean corrected error report frequencies per 1000 patients over a six-month period. Spearman's rank coefficient r_s=1, p=<0.001. The line*

represents a locally weighted scatterplot smoothing curve to represent the monotonic trend in the data over time (bandwidth=0.8).

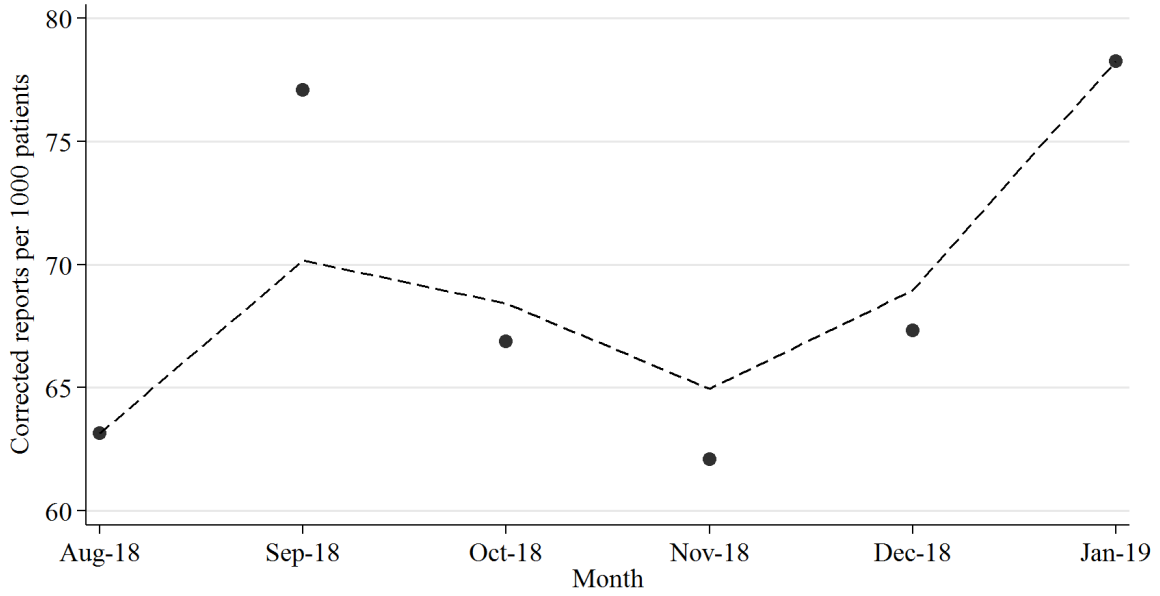

Figure 7. Mean corrected report frequencies per 1000 patients over a six-month period for the overall group of intervention laboratories. Spearman's rank coefficient r_s=-0.43, p=<0.40. The line represents a locally weighted scatterplot smoothing curve, demonstrating that no monotonic trend exists for this data (bandwidth= 0.8).

Turnaround time analysis

Of the three laboratories recommended for analysis of TATs based on acceptable recording practices, only Preah Kossamak demonstrated the monomodal, non-gausian, positively skewed distribution expected for typical turnaround time distributions. All other laboratories demonstrated a distribution with a sharp peak below the expected minimum processing time for testing. These were excluded from analysis based on the low quality of data.

For Preah Kossamak laboratory, median TATs and percentage of testing completed in the laboratory's routine and rapid ranges of under four and two hours are plotted and presented in Figure 8 and Figure 9. A visual assessment of both plots suggested the trend over time to be neither linear nor monotonic, having multiple peaks and valleys over the total time period. A Spearman's rank test for a monotonic correlation of median TATs and percentages of testing

within four- and two-hour limits over the implementation period showed no significant positive or negative trends over time for either Creatinine or CBC testing ($p>0.1$).

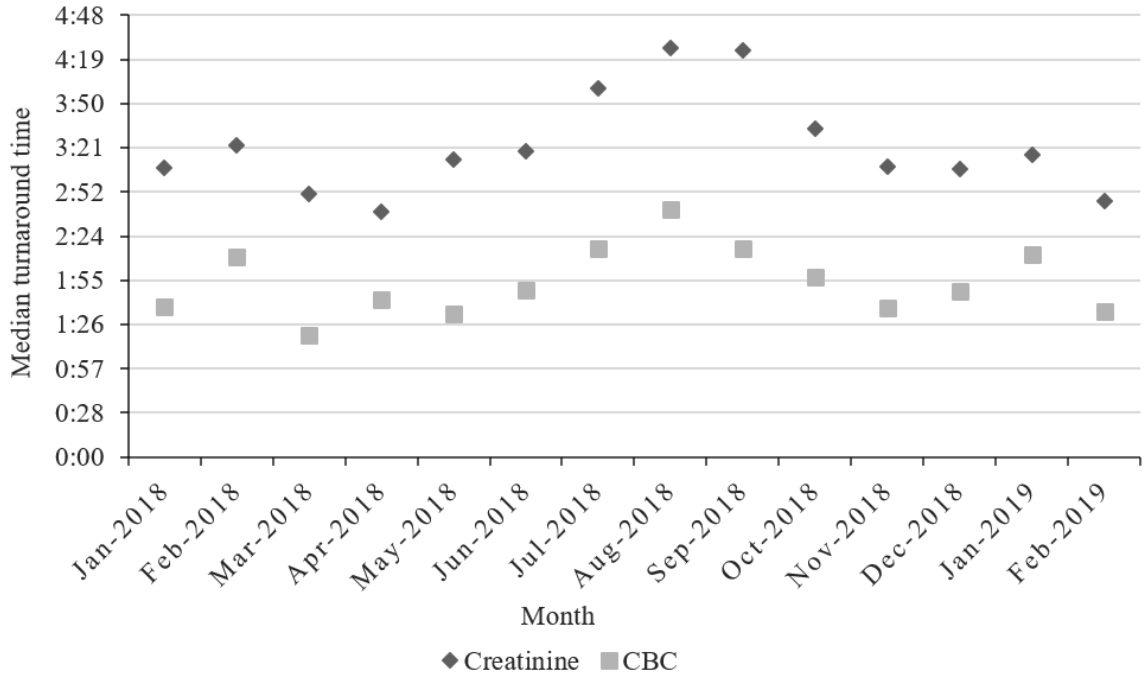

***Figure* 8.** *Median turnaround times for creatinine and complete blood count testing for Preah Kossamak Regional Hospital Laboratory between January 2018 to February 2019.*

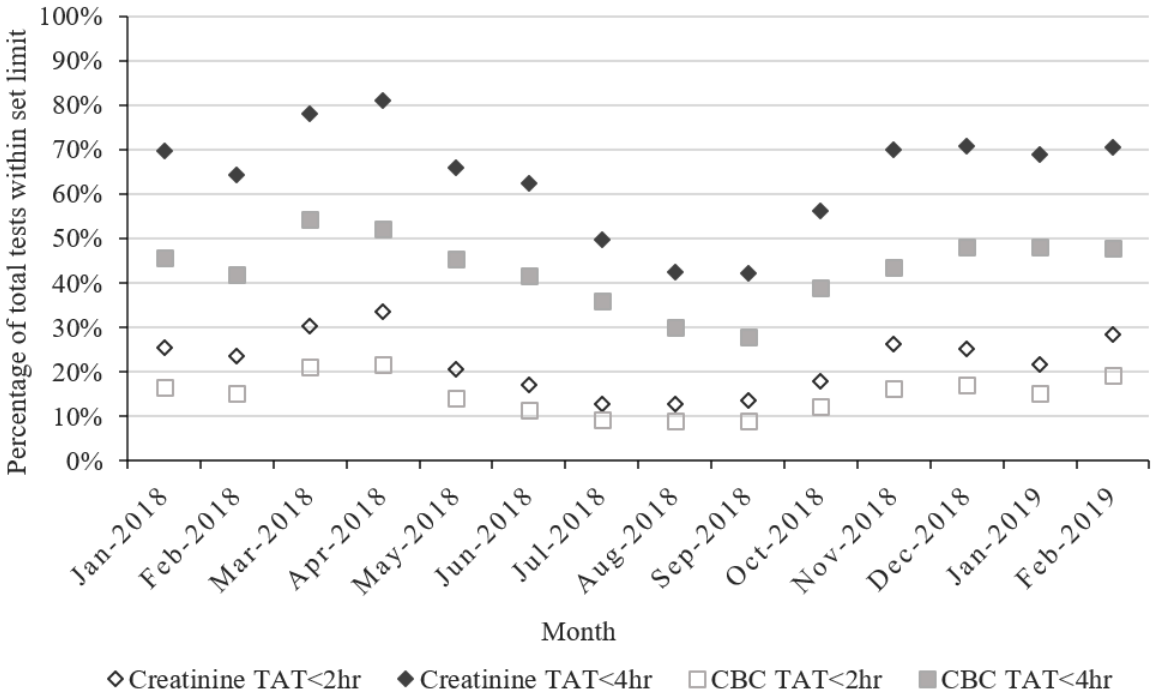

Figure 9. *Percentage of tests completed within target turnaround time ranges of two and four hours for creatinine and complete blood count tests in Preah Kossamak Referral Hospital between January 2018 to February 2019.*

Proficiency testing analysis

Calculated percentages of the 12 intervention laboratories participating and passing external quality assessments in each category of testing are plotted chronologically by implementation period in Figure 10. The various program schemes for EQA according to the specific laboratory sections were implemented by three different EQA providers, including the Pacific Paramedical Training Center, who have provided assessments for microbiology and serology as well as past hematology and biochemistry assessments until these categories were taken over by One World Accuracy in 2016 and Sysmex in 2017. A simple linear regression was fitted to all points in the data set in order to describe the trend in group compliance over time, finding the equation y=0.0269x+0.3127. Group compliance overall was found to increase at a rate of approximately 2.7% per round of implementation (8.1% per year). Some rounds of EQA were not completed or provided between 2016-2018, adding an additional limitation to the analysis.

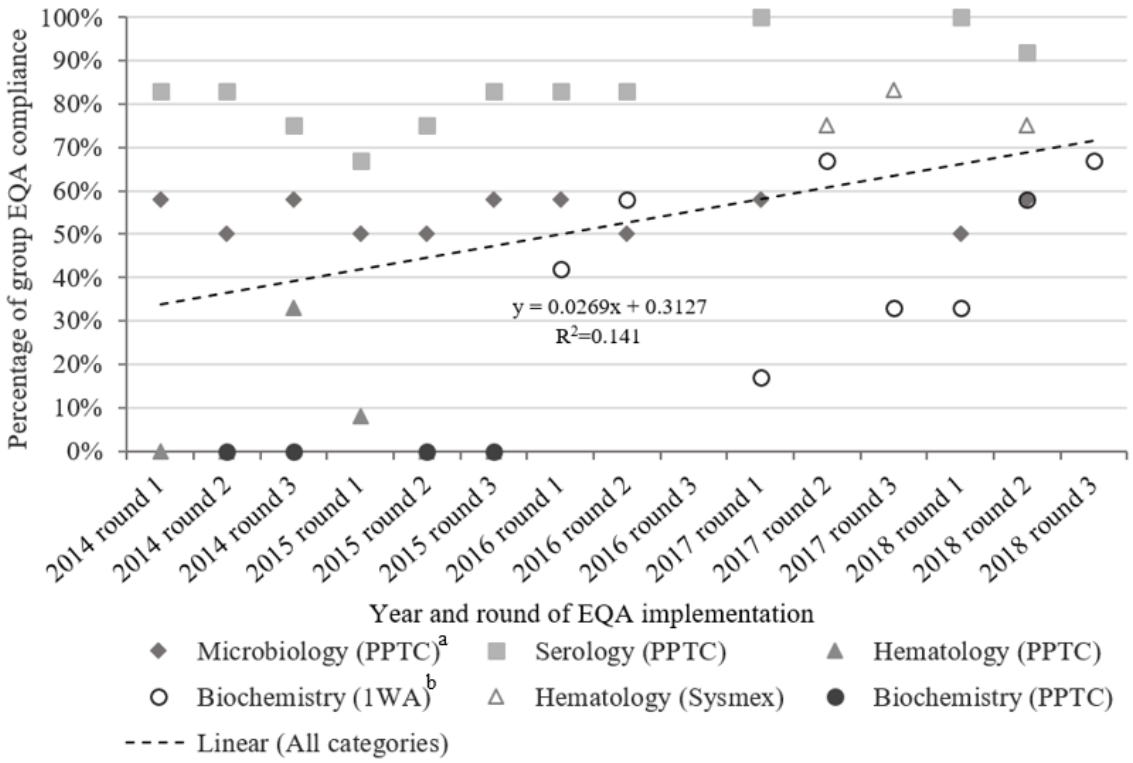

Figure 10. *Percentage of intervention laboratories participating and meeting passing criteria for EQA compliance per round of assessment for each EQA category and provider.*
[a]PPTC is an abbreviation for the organization known as the Pacific Paramedical Training Center.
[b]1WA is an abbreviation for the organization known as One World Accuracy.

A plot of mean EQA performance scores in all categories over sequential rounds of implementation is presented in Figure 11 for I-TECH mentored LQMS laboratories compared to an independent group of all other Cambodia hospital laboratories. The plot shows no apparent divergence of the two groups in mean EQA performance over time. A comparison of individual EQA categories between groups had similar results (data not shown).

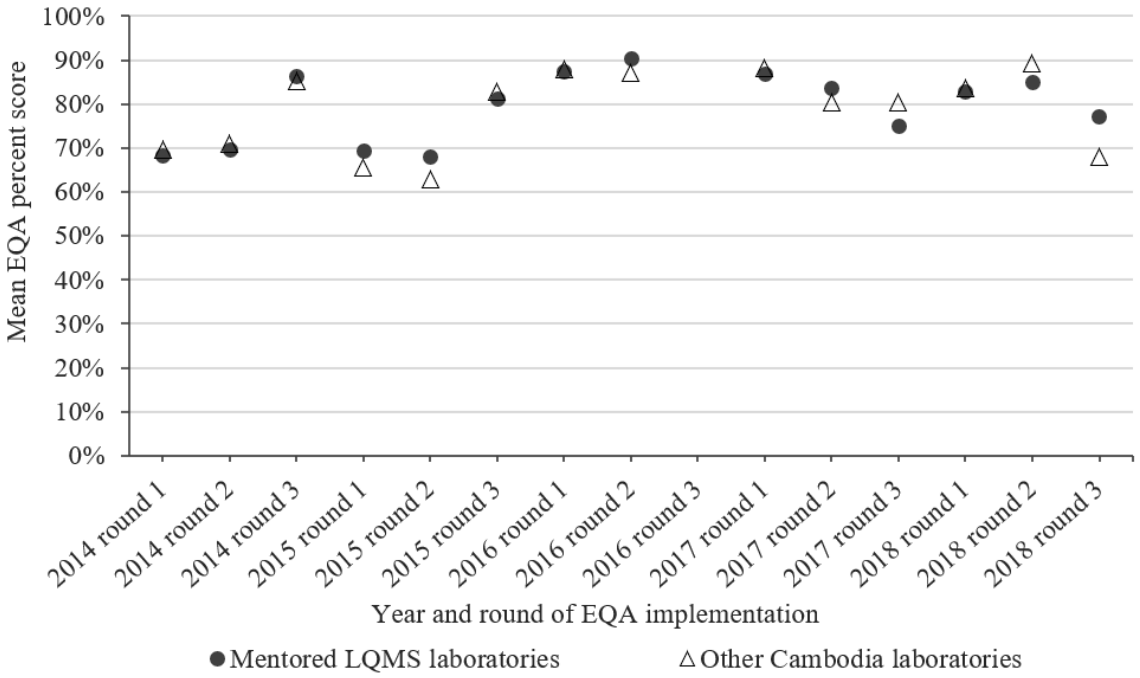

Figure 11. Mean external proficiency scores of I-TECH mentored LQMS laboratories compared to all other Cambodia MOH hospital laboratories sequential rounds of assessment implementation, averaged from all assessments in all categories.

Discussion

Key findings

Key findings of this evaluation indicate that quality management systems of laboratories participating in the I-TECH LHSS Program in Cambodia have improved significantly between baseline and post implementation CamLQMS assessments, demonstrating that laboratories are conforming to a greater proportion of LQMS standards for competent quality management systems since the start of the current implementation period. Participating laboratory CamLQMS performance is also found to be significantly higher in all categories of laboratory quality than the performance in a sample of Cambodia medical laboratories with no history of any LQMS improvement program being implemented. Although the two groups of laboratories function at different tiers of services, the evaluation suggests that Cambodia's system of provincial hospital

laboratories may benefit from a national expansion of LQMS training and mentorship programs of a similar design, utilizing the curriculum and training methodologies. Without 2018 baseline data for non-intervention laboratories, these results serve as a cross-sectional comparison of CamLQMS performance in the two groups. A controlled before-and-after study with the current 2019 data serving as baseline would provide additional information regarding the differences in improvement rates between intervention and control laboratories.

When compared to the number of outputs generated through program activities, the positive correlation between improved CamLQMS assessment scores and trainer contact time with laboratories via video conference activities suggests that remote mentoring of laboratories with trained local LQMS technical experts can be a strong driver of laboratory strengthening. Often referred to as telementoring, studies have shown the use of video communication technologies to be an efficient and effective tool to provide seasoned or specialized surgical expertise to learning surgeons of remote or resource-limited medical facilities. As a tool for laboratory mentorship, a recent medical laboratory strengthening program in Southeast Europe utilized monthly mentorship through telecommunication to improve laboratory quality in five different countries and demonstrated measurable progress within the six laboratories supported[24]. Qualitative evaluation of the program indicated improved accountability, collaborative problem solving, and increased awareness of the importance of laboratory quality to be important drivers of the program's success[24].

Notably, although attendance to formal training and the number of on-site visits did not vary enough for correlation studies, they too were expected to have had an impact on LQMS improvement. In particular, the program's content in the formal training curriculum reflects improvements in individual audit sections well. The program generally excluded the topic of

information management due to topical overlap with CamLIS implementation by WHO Cambodia. *Facilities and safety* was further excluded as a training topic due to overlap with another national program. These two areas corresponded closely to the smallest mean percent differences of 0% and 17% between audit years. *Purchasing and inventory* likewise received little attention in the curriculum content, resulting in a minor 16% improvement between years. In contrast, topics such as *documents and records*, *management review*, and *occurrence management and process control* received a heavy amount of focus in formal trainings. Although *documents and records* showed only a marginally higher improvement, *occurrence management and process control* improved by a mean percent difference of 52%, and *management review* improved by 39%.

During site visits, mentors worked closely with laboratories on specific technical needs of laboratories such as improved use of quality indicators (*occurrence management*), quality control testing (*process control*), *equipment* verification, and *corrective action*. In later stages of the implementation period, mentors coached lab personnel on internal auditing in preparation for the second round of CamLQMS audits. Mentors also worked with laboratories to develop a quality manual, which satisfied one of the key requirements for *documents and records* conformance.

Video conference activities through Zoom were informal and often more specific to the questions and answers of laboratory participants. Primarily providing mentorship and specific training to QAOs and laboratory managers, these video conferences occurred weekly (at least once minimum per week), consisting of technical expertise from quality mentors, and enhanced by the networking and shared experience of participating laboratory quality assurance personnel. Many smaller meetings with one or two laboratories at a time were scheduled in between,

providing one on one consultation from mentors that spanned multiple topics of quality management.

Although LQMS performance and conformity to standards have improved significantly, physician satisfaction surveys demonstrated decreased ratings in several categories of satisfaction with laboratory services between 2015 and 2019, particularly in the categories of test TATs and timely calling of critical results. These findings can seem contradictory, giving the impression that the quality of laboratory services is decreasing; however, these surveys are best interpreted as objective perceptions of quality rather than precise measures of quality itself. Previous studies have found categories of communication and reporting such as these to be uniformly low rated areas of satisfaction among physicians, despite process improvement efforts[21,25]. Physician satisfaction ratings are particularly difficult to associate with monitoring practices of common quality indicators for process improvement such as TAT monitoring, critical result notifications, and corrected reports,[26]. Satisfaction ratings have, however, shown a negative association with the size of the hospital, with larger hospitals receiving lower overall satisfaction scores, although the mechanism of this association is not defined[26]. It may be that increased workload or even physician expectations of laboratory services as hospitals expand and improve have outpaced improvements in quality between surveys. These findings emphasize the need for improved occurrence management and process improvement systems to accurately measure and improve quality in these laboratories to adapt with changing health systems.

To that aim, the study identifies an important deficiency in record keeping practices by laboratories, leading to an inability to record accurate turnaround times and monitor improvement. TAT analysis is a valuable tool for monitoring the timeliness of laboratory services and guiding corrective action for laboratories to improve quality[21]. It is furthermore

shown to be steadily important for physician satisfaction with laboratory services with the standard expectation for 90% of results or more to be completed within 60 minutes in many institutions[25,26]. The laboratory evaluated for TAT changes in this study found that less than 60% of CBC results were completed within four hours in any month measured, with no trend of improvement over the course of the evaluation period. To resolve this problem, the electronic CamLIS database provides a convenient and effective method for collection and storage of a ready source of data, and it should be utilized as a monitoring tool by laboratories to improve quality. Improving documentation of receipt times for laboratory samples is a practical step toward accurate TAT analysis and improved timeliness of results in Cambodian laboratories.

Regarding corrected report frequencies study also served to identify extreme error correction frequencies in Preah Ang Duong and Preah Kossamak Hospitals, measuring frequencies of over 45% errors in some sample periods. Further investigation into the root causes of these errors and delays in their identification may help to improve patient safety and laboratory quality, as well as to identify systemic drivers of high corrected report frequencies.

Finally, analysis of proficiency testing results suggests that the proportion of mentored LQMS laboratories participating and meeting compliance standards in Cambodia's national EQA program is increasing over time. Both participation and mean EQA scores in participating laboratories show a positive trend over time; however, because it could not be demonstrated that I-TECH mentored laboratories are improving at a greater rate than other laboratories, these improvements could not be attributed directly to the improved quality management systems of these laboratories. Further studies may identify common factors between intervention and non-intervention laboratories that may be impacting EQA performance.

Study limitations

Overall audit scores, although they serve a descriptive purpose to approximate the percentage of a list of standards conformed to by clinical laboratories, should be interpreted with caution as a standalone measure. Some individual sections are given a greater value than others and are not represented equally in the overall score. A laboratory may receive a high overall score but still perform poorly in a lower valued section such as *client management and customer service*. Some questions are also excluded from analysis if it is not applicable to the laboratory's normal functions, lowering the total score possible for that individual section and overall. This may slightly increase the scores of smaller laboratories that may have not met the excluded standard otherwise. These results are most accurate as a gap assessment for individual laboratories to drive improvement in each specific section of LQMS. Nonetheless, comparisons of overall scores serve as a useful estimate of the outcome of the evaluated program activities.

Regarding the correlation of Zoom activity participation time with percent differences in audit scores, it is possible that the internal motivation of laboratories to improve may be a confounding factor in this study, potentially driving a high rate of improvement and participation time in teleconference training independently. Analysis of video conference usage reports are further subject to misclassification bias due to an incompleteness of information regarding the association of users with individual laboratories. Approximately 2% of participation time in video conference activities could not be associated with or disassociated from individual laboratories and could potentially be biased toward one or a few laboratories prone to using unidentified devices. Further bias lies in the potential for multiple laboratory personnel to participate through a single device, thus being recorded as a single user associated with that corresponding laboratory. Many of such cases such as these were identified and accounted for

through supplemental mentor reports; however, these supplemental reports were not available for all teleconference events and were unlikely to account for all such cases.

This evaluation is further limited to a small sample size of laboratories, along with a small sample of time points for time trend analysis, especially in the case of the six-month sample of corrected report frequencies, which placed a severe limitation on the study in its ability to detect change. In the case of corrected report frequencies, the trend appears to have decreased monotonically over time in two of the laboratories studied; however, further observations over time may add additional insight into whether these changes can be sustained by the laboratories or whether the trends are seasonal, coming and going with laboratory work load trends. For all four of the measured indicators of quality included in this study, continued improvement of LQMS performance and conformity to standards may lead to demonstrated improvements given additional time, particularly in the area of *occurrence management and process improvement*.

Conclusion

The I-TECH Laboratory Health System Strengthening Program has used a combination of training, mentoring, and advocacy to achieve rapid and significant outcomes in quality management system development. Participating laboratories performed significantly better in assessments of LQMS performance and conformity than non-intervention laboratories, suggesting that an expansion of the methods in Cambodia may benefit currently non-mentored, non-LQMS laboratories significantly. The study uses a correlation of remote mentoring through video conference calling with improved external audit scores to highlight remote mentoring as one potentially effective tool for laboratory LQMS development. This finding, if confirmed or supported further, has implications for LQMS strengthening in remote, hard-to-reach hospitals, and presents a cost-effective alternative to frequent travel for on-site mentorship.

The study successfully utilized data from the Cambodian electronic laboratory information system, among other means, to measure changes in laboratory performance indicators and identify quality gaps that may guide future strengthening of laboratory systems. While significant progress has been demonstrated, accreditation has still not yet been achieved, and laboratories in Cambodia should continue to implement stepwise improvement programs toward accreditation with increased emphasis on improving the quality of performance indicator data for effective quality improvement.

Reference Page

1. Olmsted SS, Moore M, Meili RC, et al. Strengthening laboratory systems in resource-limited settings. *Am J Clin Pathol.* 2010;134(3):374-380.
2. Carter JY, Lema OE, Wangai MW, Munafu CG, Rees PH, Nyamongo JA. Laboratory testing improves diagnosis and treatment outcomes in primary health care facilities. *African journal of laboratory medicine.* 2012;1(1):8-8.
3. Najjar-Pellet J, Machuron J-L, Bougoudogo F, et al. Clinical laboratory networks contribute to strengthening disease surveillance: The RESAOLAB Project in West Africa. *Emerging health threats journal.* 2013;6:10.3402
4. World Health Organization Regional Office for South-East Asia. Laboratory Quality Standards and their Implementation. New Delhi, India. http://www.who.int/iris/handle/10665/205405. Published 2011. Accessed April 21, 2019.
5. Gachuki T, Sewe R, Mwangi J, et al. Attaining ISO 15189 accreditation through SLMTA: A journey by Kenya's National HIV Reference Laboratory. *Afr J Lab Med.* 2014;3(2):216.
6. Kingdom of Cambodia in Collaboration with the World Health Organization. National Strategic Plan for Medical Laboratory Services 2010-2015. Phnom Penh, Cambodia; https://www.globe-network.org/sites/default/files/sites/default/files/documents/national-strategic-plan-for-medical-laboratory-services-2010-2015.pdf. Published 2010. Accessed April 21, 2019.
7. Kingdom of Cambodia. National Laboratory Strategy 2015-2020. Phnom Penh, Cambodia: Department of Health Services, Cambodia Ministry of Health; 2015:7-34.
8. World Health Organization. Joint external evaluation of IHR core capacities of the Kingdom of Cambodia: mission report, 26 August - 2 September 2016. Geneva, Switzerland. http://www.who.int/iris/handle/10665/254705. Published 2017. Accessed April 21, 2019.
9. Pierson A, Orelle A. Assessment of the National Laboratory System in Cambodia: Summary Report. Integrated Quality Laboratory Services; 2013:1-38.
10. Perrone LA, Voeurng V, Sek S, et al. Implementation research: a mentoring programme to improve laboratory quality in Cambodia. *Bull World Health Organ.* 2016;94(10):743-751.
11. Nguyen TT, McKinney B, Pierson A, et al. SLIPTA e-Tool improves laboratory audit process in Vietnam and Cambodia. *Afr J Lab Med.* 2014;3(2):219.
12. Yao K, Maruta T, Luman ET, Nkengasong JN. The SLMTA programme: Transforming the laboratory landscape in developing countries. *Afr J Lab Med.* 2014;3(3).
13. World Health Organization Royal Tropical Institute. Laboratory quality stepwise implementation tool. Geneva, Switzerland. https://www.who.int/ihr/lyon/hls_lqsi/en. Available 2014. Accessed April 21, 2019.
14. Clinical Laboratory Standards Institute. Application of a quality management system model for laboratory services, approved guideline GP26-A3. 3rd ed. Wayne, Pensylvania: NCCLS; 2004.
15. World Health Organization Regional Office for the Western Pacific. The Kingdom of Cambodia Health System Review. *Health Systems in Transition*; 5(2). https://iris.wpro.who.int/handle/10665.1/11356. Published 2015. Accessed April 21, 2019.

16. ISO 15189. Medical laboratories - Requirements for quality and competence. Geneva, Switzerland: International Organization for Standardization; 2012.1-36.
17. Barth JH. Clinical quality indicators in laboratory medicine. *Annals of Clinical Biochemistry.* 2011;49(1):9-16.
18. Shahangian S, Snyder SR. Laboratory medicine quality indicators: a review of the literature. *Am J Clin Pathol.* 2009;131(3):418-431.
19. Yuan S, Astion ML, Schapiro J, Limaye AP. Clinical impact associated with corrected results in clinical microbiology testing. *Journal of clinical microbiology.* 2005;43(5):2188-2193.
20. Fanshawe TR, Ordóñez-Mena JM, Turner PJ, Bruel AVd, Shine B, Hayward GN. Frequencies and patterns of laboratory test requests from general practice: a service evaluation to inform point-of-care testing. *Journal of Clinical Pathology.* 2018;71(12):1065-1071.
21. Hawkins RC. Laboratory turnaround time. *The Clinical biochemist Reviews.* 2007;28(4):179-194.
22. Jenny RW, Jackson KY. Proficiency test performance as a predictor of accuracy of routine patient testing for theophylline. *Clinical Chemistry.* 1993;39(1):76.
23. El-Sabawi B, Magee W, 3rd. The evolution of surgical telementoring: current applications and future directions. *Annals of translational medicine.* 2016;4(20):391-391.
24. Polansky L, Chester S, Warren M, et al. Can mentorship improve laboratory quality? A case study from influenza diagnostic laboratories in Southeast Europe. *BMC Health Services Research.* 2019;19(1):49.
25. Zarbo RJ, Nakhleh RE, Walsh M. Customer Satisfaction in Anatomic Pathology. *Archives of Pathology & Laboratory Medicine.* 2003;127(1):23-29.
26. McCall SJ, Souers RJ, Blond B, Massie L. Physician Satisfaction With Clinical Laboratory Services: A College of American Pathologists Q-Probes Study of 81 Institutions. *Archives of Pathology & Laboratory Medicine.* 2016;140(10):1098-1103.

Appendix A: CamLQMS Audit Scoring Sheet

AUDIT SCORING

Cambodia Laboratory Quality Management System (CamLQMS) Checklist contains 12 main sections (a total of 117 questions for a total of 275 points. Each item has been awarded a point value of 2, 3, or 5 points—based upon relative importance and/or complexity. Responses to all questions must be, "yes", "partial", or "no".

- Items marked "*yes*" receive the corresponding point value (2, 3, or 5 points). **All elements of a question must be present in order to indicate "yes" for a given item and thus award the corresponding points.**

  NOTE: items that include "tick lists" must receive all "yes" and/or "n/a" responses to be marked "yes" for the overarching item.
- Items marked "*partial*" receive 1 point.
- Items marked "*no*" receive 0 points.

When marking "partial" or "no", notes should be written in the comments field to explain why the laboratory did not fulfil this item to assist the laboratory with addressing these areas of identified need following the audit.

Where the checklist question does not apply, indicate as NA. Subtract the sum of the scores of all questions marked NA and subtract that sum of NAs from the total of 275. Since denominator has changed, the level status is then determined using % score.

Audit Score Sheet

Section	*Total Points*
Section 1: Documents & Records	28
Section 2: Management Reviews	14
Section 3: Organization & Personnel	22
Section 4: Client Management & Customer Service	10
Section 5: Equipment	35
Section 6: Evaluation and Audits	15
Section 7: Purchasing & Inventory	24
Section 8: Process Control	32
Section 9: Information Management	21
Section 10: Identification of Non Conformities, Corrective and Preventive Actions	19
Section 11: Occurrence/Incident Management & Process Improvement	12
Section 12: Facilities and Biosafety	43
TOTAL SCORE	275

Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
(0 – 150 pts)	(151 – 177 pts)	(178 – 205 pts)	(206 – 232 pts)	(233 – 260 pts)	(261 – 275 pts)
< 55%	*55 – 64%*	*65 – 74%*	*75 – 84%*	*85 – 94%*	*≥ 95%*

Appendix B: 2019 Physician Satisfaction Survey (English Translation)

Physician Satisfaction with Laboratory Services Survey

The purpose of this survey is to collect information about the perceived reliability of laboratory services and how confident providers are in the quality of laboratory results.

A. General information

1. Job Position: _____
2. On average, how many patients do you see each day? _____
3. Approximately which % of your patients require laboratory testing? _____

B. Satisfaction with laboratory services.

4. On a scale of 1 through 5 (**1=very poor, 2=poor, 3=average, 4=good, and 5=very good**) please rate how you would rank your hospital laboratory in terms of the following (please circle one option for each line):

   a. Ease of test ordering ……...…….…….. 1 2 3 4 5

   b. Quality results ……... 1 2 3 4 5

   c. Good turnaround time ……..... 1 2 3 4 5

   d. Good communication ……1 2 3 4 5

   e. Good service ……...1 2 3 4 5

5. On a scale of 1 through 5 (**1=strongly disagree, 2=disagree, 3=neither agree nor disagree, 4=agree, 5=strongly agree**) please rate how strongly you agree with the following statements about services from your hospital laboratory (please circle one option for each line):

   a. My laboratory informs me of critical values in a timely manner

   1 2 3 4 5

   b. Laboratory test results are compatible with patient clinical observations

   1 2 3 4 5

6. On a scale of 1 through 5 **(1=Not at all important, 2=somewhat unimportant, 3=neither important nor unimportant, 4=somewhat important, 5=Very important)** please rate how important the following factors are in preventing you from ordering a laboratory test from your hospital laboratory (please circle one option for each line):

a. Test is not offered by my laboratory……..1 2 3 4 5
b. Diagnosis is based on symptoms only ……..1 2 3 4 5
c. Laboratory results are too slow ……1 2 3 4 5
d. Laboratory results are not reliable……1 2 3 4 5
e. I am not confident on test to order ……..……..1 2 3 4 5
f. Patient can't pay for the test ……..1 2 3 4 5
g. There are no national guidelines ……...1 2 3 4 5
h. I am not confident on sample collection……1 2 3 4 5
i. I am not confident in interpreting the results……1 2 3 4 5

7. Comments _____

_____