Reproducibility of Results and Performance of TB Diagnostics in East Africa Public Health Laboratories Networking Project - in Kenya

*Githui W.A[1], Mwangi M[2], Sang W.K[3], Juma E[1], Aberi R.M[1], Kiptoo M[4], Wanzala P[2]

1. *Kenya Medical Research Institute-Centre for Respiratory Disease Research*
2. *Kenya Medical Research Institute-Centre for Public Health Research*
3. *Kenya Medical Research Institute-Centre for Microbiology Research*
4. *South Eastern Kenya Universit, Kenya*

***Corresponding Author**:* Githui W.A. Email: wgithui@yahoo.com

Summary

INTRODUCTION

Reproducibility of laboratory results and performance of diagnostic tools form major part of quality assurance in diagnosis. This is key to care of patients with tuberculosis (TB) in Kenya. However, to our knowledge, no study has documented comparison of reproducibility of laboratory results and performance of these TB diagnostic tools in different geographical settings.

OBJECTIVE

To determine reproducibility of laboratory results and performance of ZN and new TB diagnostic tools in different geographical settings in the East Africa Public Health laboratories Networking Project - Operational Research (EAPHLNP-OR) Tuberclosis(TB) study in Kenya. Equally it was to evaluate the impact of new tuberculosis diagnostics on patient's health outcomes in Kenya.

METHODOLOGY

A cross-sectional study was conducted between February 2013 and October 2016. People presumed to have TB aged 18 years and above were enrolled in nine selected public health facilities in Busia, Kisii-Keroka, Kitale, Lamu Machakos, Malindi, Nyahururu, Olololunga-Narok and Wajir and selected to participate in that EAPHLNP-OR TB project. Spot and morning sputum specimens were collected from participants presenting at each of the nine health facilities on two consecutive days. A total of 5715 sputum specimens were collected from all the study sites.

A proportion of the specimen at each site was aseptically processed for Ziehl Neelsen (ZN), Flourescence Microscopy (FM) and GeneXpert. The remaining specimen portion was triple packaged and shipped to the TB research laboratory at the Kenya Medical Research Institute (KEMRI), in Nairobi.

Processing using the three tests including Lowenstein Jensen (LJ) culture according to standard procedures was done at KEMRI in Nairobi. The laboratory personnel were blinded of the previous study site results. Data processing was done using *MySQL* and *IBM SPSS version 24 software*. Reproducibility of laboratory results where the specimen was the unit of analysis, was determined by Kappa values on homogenous specimens and performance by diagnostic values (sensitivity, specificity, positive/ negative predictive values).

Results from the same specimens at the study sites were compared with those from KEMRI research laboratory LJ culture was used as a gold standard to evaluate performance, where the patient was the unit of analysis.

RESULTS

Generally, excellent Kappa values for GeneXpert (0.855 (95% CI : 0.834 - 0.876), n = 660) was significantly higher than ZN microscopy (0.721 (95% CI : 0.708 - 0.734), n = 3252). FM Kappa value (0.749 (95% CI : 0.736 - 0.762), n = 2816) indicated substantial agreement. Specific results for the three diagnostic tools varied across the sites for microscopy but were not significantly different for GeneXpert.

Cumulatively, there was marginal significant incremental sensitivity of microscopy at the study sites ZN (69.9% (95% CI: 64.3 - 75.5); n = 259) and FM (76.7% (95% CI: 71.1 - 82.3) ; n = 219) than at KEMRI ZN (68.7% (95% CI: 63.1-74.4); n = 259) and FM (70.8% (95% CI:64.8 - 76.8)%; n = 219).

Apparently, the sensitivity of GeneXpert at the study sites (81.4% (95% CI:71.4 - 91.3) ; n = 59) was not significantly different from that of GeneXpert at KEMRI (81.4% (95% CI: 71.4 - 91.3); n = 59). Specificity of GeneXpert at the site was not significantly different to that of KEMRI but significantly lower than microscopy both at site and KEMRI. Microscopy results varied across specific study sites but not significantly different for GeneXpert. A similar pattern was observed for positive / negative predictive values.

CONCLUSION

Results generated by GeneXpert MTB/RIF indicated excellent reproducibility but there was no significant difference in performance, regardless of geographical setting in Kenya. This suggests that under ideal conditions GeneXpert MTB/RIF is a reliable diagnostic tool irrespective of the facility setting.

Due to higher specificity and positive/negative predictive values of microscopy to GeneXpert, microscopy could compliment GeneXpert in detection of mycobacteria, especially in settings with inadequate capacity, including infrastructure, human resource and high workload. There is need for continuous training in microscopy to enhance reproducibility of laboratory results and performance.

[***Afr. J.*** Health Sci. 2019 32(4) : 7 - 22]

Introduction

Reproducibility is the ability of an entire experiment or study to be duplicated, either by the same researcher or by someone else working independently.

A particular experimentally obtained value is said to be reproducible if there is a high degree of agreement between measurements and observations conducted on replicate specimens in different locations by different people-that is, if the experimental value is found to have a high precision. (ASTM International, ASTM E177-2014).[1]

Reproducibility of laboratory results and performance of diagnostic tools form a major part of quality assurance in laboratory diagnosis, which is key in patient management.

Performance of a diagnostic tool results in diagnostic values (Sensitivity, Specificity, PPV and NPV) generated using a gold standard. In medical diagnosis, test sensitivity is the ability of a test to correctly identify those with the disease (True Positive Rate), whereas test specificity is the ability of the test to correctly identify those without the disease True Negative Rate).

Quality assured diagnosis is key to care of patients with tuberculosis (TB). Research on new TB diagnostic tools and evaluation on their performance has taken centre stage nowdays.

In Kenya, performance of such tools including Led Emitted Diode fluorescent microscopy (LED-FM) and Xpert MTB/ RIF® (GeneXpert), have recently been documented. Reproducibility of laboratory results and performance of diagnostic tools form major part of quality assurance in diagnosis of tuberculosis (TB) in patients.

Regardless, of our knowledge, no study has documented comparison of reproducibility of laboratory results and performance of these TB diagnostic tools in different geographical settings.

Objective

To determine reproducibility of laboratory results and performance of ZN and new TB diagnostic tools in different geographical settings in the East Africa Public Health Laboratories Networking Project-Operational Research **(EAPHLNP-OR)** TB study in Kenya.

Methodology

Study Design

A cross sectional design involving determination of the reliability and diagnostic test values of ZN and new TB diagnostics, LED-FM and GeneXpert.

Study Area

Nine study sites; Busia, Kisii-Keroka, Kitale, Lamu Machakos, Malindi, Nyahururu, Olololunga-Narok and Wajir were selected to participate in the main EAPHLN-OR TB project. The main objective was to evaluate the impact of new tuberculosis diagnostics on patients' health outcomes in Kenya.

Study Population

A cross-sectional study was conducted between February 2013 and October 2016. People presumed to have TB, aged 18 years and above were enrolled at the nine selected public health facilities.

A total of 5715 sputum specimens were collected from all the study sites. Consenting age was eighteen (18) years and above to be enrolled in the public health facilities.

Spot and morning sputum specimens were collected from participants presenting at each of the nine health facilities on two consecutive days.

At each site, a proportion of each specimen was aseptically processed for ZN, FM and GeneXpert (where available). The remaining portion of the specimen was triple packaged and shipped to the TB research laboratory at the Kenya Medical Research Institute (KEMRI), Nairobi. Here, they were processed using the three tests including Lowensen Jensen (LJ) culture in accordance with standard procedures.

The laboratory personnel at KEMRI, were unaware of the study sites results. Data processing was done using MySQL and IBM SPSS version 24 software.

Reproducibility of laboratory results where the specimen was the unit of analysis was determined by Kappa values on *homogenous* specimens and performance by diagnostic values (Sensitivity, Specificity, Positive / Negative Predictive Values).

Results from the same specimens at the study sites were compared with those from KEMRI research laboratory. LJ culture was used as a gold standard to evaluate performance, where the patient was the unit of analysis.

Results

Figure 1 presents patients flow from enrolment for sputum sample analysis in the laboratory.

A Total of 3075 people presumed to have TB were enrolled between February 2013 and October 2016 at all the study sites. Of these, 2928(95.2%) respondents produced specimens that were accepted for analysis based on a set criterion.

A total of 5715 specimens were processed as below:

2782 / 2928	(95.0%)	Patients had both (spot and morning specimens.
146 / 2928	(5.0%)	Patients produced one sputum specimen (105 spot, 41 morning).

The specimens were processed at the study sites using Ziehl Neelsen (ZN), LED Fluorescent (LED-FM) microscopy and GeneXpert**,** and at KEMRI using ZN, LED-FM, GeneXpert and LJ solid culture.

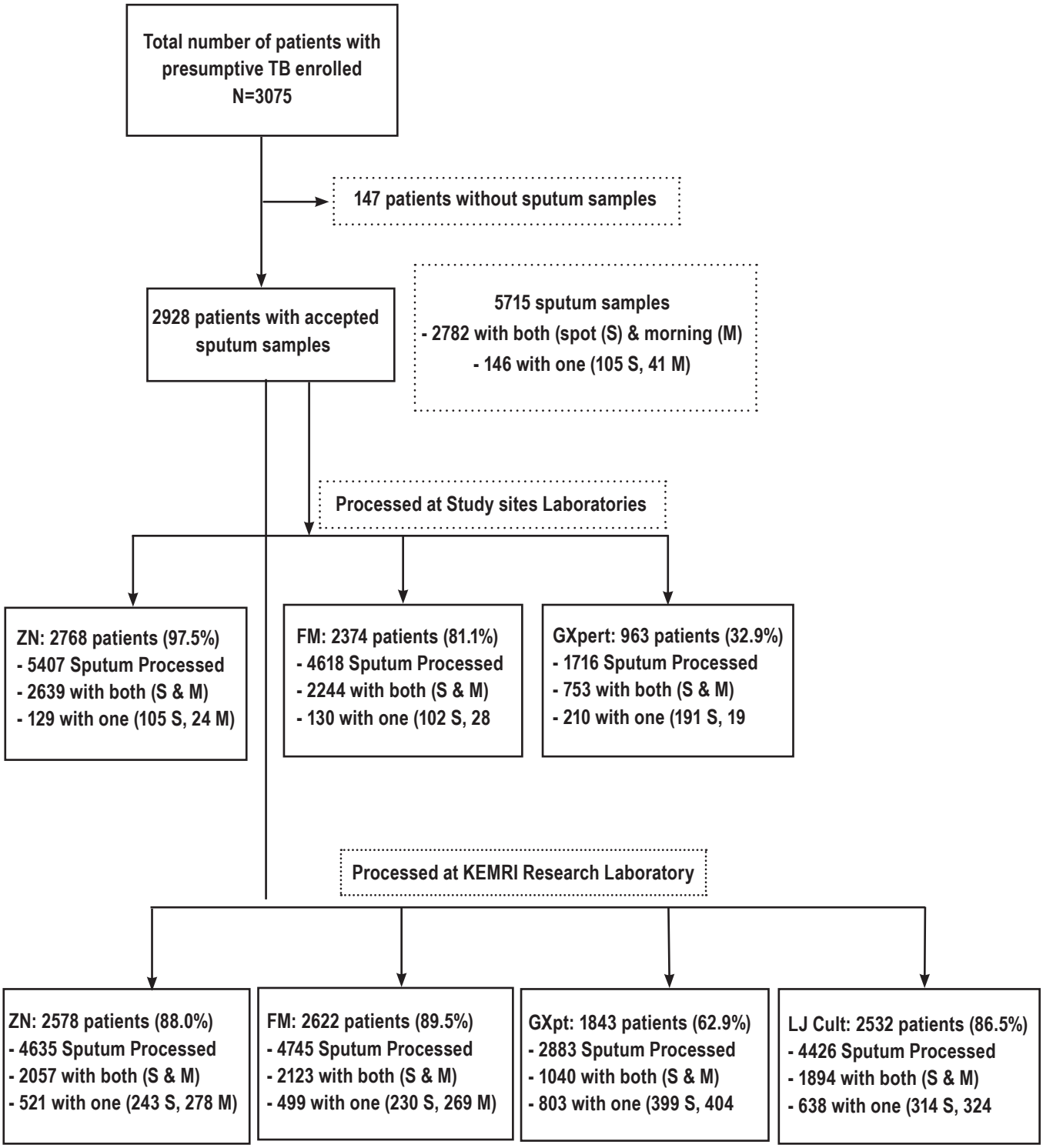

Figure 1: *Patients Flow From Enrolment To Sputum Sample Analysis In The Laboratory*

*(**Table 1**)* presents selected demographic characteristics of all the enrolled participants. A total of 3075 people presumed to have TB were enrolled into the study.

Selected demographic characteristics of all the enrolled participants are presented in ***(Table 2)*** The proportion of male (47.6%) to female (42.9%) was visible with more male than female. Mean age of the people presumed to have TB was 39 (+16.5SD) ranging between 18 and 98 years. Majority of the study participants (43.6%) were in the age group of between 20 and 39 years.

Table 1: *Enrolled People Presumed To Have TB Per Site and Their Selected Demographic Characteristics*

Characteristics	n = 3075	%
Study sites		
Busia	592	19.3
Kitale	631	20.5
Machakos	167	5.4
Malindi	1112	36.1
Wajir	227	7.4
Kisii-Keroka	104	3.4
Nyahururu	137	4.5
Narok	75	2.4
Lamu	30	1.0
Gender		
Male	1465	47.6
Female	1320	42.9
Not documented	290	9.5
Age in years		
<20 years	213	6.9
20 - 29 years	630	20.5
30 - 39 years	709	23.1
40 - 49 years	456	14.8
50 - 59 years	290	9.4
60 - 69 years	208	6.8
70 - 79 years	181	5.9
80 and above	388	12.6

Table 2 shows the prevalence of tuberculosis among the study participants (with both spot and morning samples) according to different diagnostic tools and the testing site.

At the study sites laboratories, GeneXpert detected more participants as TB positive (23.9%) followed by FM (16.8%) and ZN (15.6%). Similarly, analysis at the KEMRI research laboratory confirmed that, GeneXpert detected more study participants as TB positive (23.3%) followed by ZN (16.5%) and FM (15.8%). The prevalence of TB was comparable according to the microscopy diagnostic tools (ZN and FM) but significantly high according to GeneXpert at both study sites and KEMRI research laboratories.

The prevalence of TB by LJ culture was significantly lower than GeneXpert at both study sites and at the KEMRI laboratory.

Table 2: *Prevalence Of Tuberculosis Among The Study Participants According To Different Diagnostic Tools (With Both Spot and Morning Samples)*

Point Of Analysis	Diagnostic Tool	Prevalence (95% CI)
Study Sites Laboratories	ZN Microscopy (n=1626)	15.6(13.8-17.4)
	FM Microscopy (n=1408)	16.8(14.9-18.8)
	GeneXpert (n=330)	23.9(19.3-28.5)
Kemri Research Laboratory	ZN Microscopy (n=1626)	16.5(14.7-18.3)
	FM Microscopy (n=1408)	15.8(13.9-17.7)
	GeneXpert (n=330)	23.3(18.7-27.9)
	Culture: LJ (n=1728)	15.6(13.9-17.3)

Reproducibility of results for different TB diagnostic tools between study sites and KEMRI research laboratory was determined as presented in ***Table 3.*** Generally, excellent Kappa value for GeneXpert (0.855(95%CI: 0.834-0.876), n=660) was significantly higher than ZN microscopy (0.721(95%CI: 0.708-0.734), n=3252). FM Kappa value (0.749(95%CI: 0.736-0.762), n=2816) indicated substantial agreement. Results varied across the sites for both ZN and FM microscopy but were not significantly different for GeneXpert.

Table 3: *Reproducibility Of Results For Different TB Diagnostic Tools Between Study Sites and KEMRI Research Laboratory*

Diagnostic tool		**ZN KEMRI**			
		Positive + n	**Negative - n**	**Total**	**Kappa value (95% CI)**
ZN sites	Positive	345	125	470	0.721(0.708-0.734)
	Negative	93	2689	2782	
	Total	438	2814	3252	
		FM KEMRI			
FM sites		**Positive + n**	**Negative - n**	**Total**	**Kappa value (95% CI)**
	Positive	324	116	440	0.749(0.736-0.762)
	Negative	61	2315	2376	
	Total	385	2431	2816	
		GeneXpert KEMRI			
GeneXpert sites		**Positive + n**	**Negative - n**	**Total**	**Kappa value (95% CI)**
	Positive	130	19	149	0.855(0.834-0.876)
	Negative	14	497	511	
	Total	144	516	660	

Key: *n = Number, N = Total number*

Table 4 shows reproducibility of ZN results for different TB diagnostic tools between individual study sites and KEMRI research laboratory. Reproducibility of ZN microscopy results varied significantly across some sites, ranging from Kappa value = 0.308 (95% CI: 0.215 - 0.401) in Lamu to Kappa value = 0.830 (95% CI: 0.779 - 0.881) in Nyahururu.

Table 4: *Reproducibility Of ZN Results For Different TB Diagnostic Tools Between Study Sites and KEMRI Research Laboratory*

ZN KEMRI							95% CI	
Study site	**Diagnostic tool**		**positive + n**	**Negative - n**	**Total**	**Kappa value**	**L**	**U**
Busia	ZN site	Positive	63	20	83	0.693	0.663	0.723
		Negative	25	490	515			
		Total	**88**	**510**	**598**			
Kitale	ZN site	Positive	86	42	128	0.714	0.688	0.740
		Negative	13	593	606			
		Total	**99**	**635**	**734**			
Machakos	ZN site	Positive	34	9	43	0.756	0.714	0.798
		Negative	8	185	193			
		Total	**42**	**194**	**236**			
Malindi	ZN site	Positive	95	30	125	0.747	0.722	0.772
		Negative	25	938	963			
		Total	**120**	**968**	**1088**			
Wajir	ZN site	Positive	26	14	40	0.669	0.624	0.714
		Negative	6	176	182			
		Total	**32**	**190**	**222**			
Kisii	ZN site	Positive	14	2	16	0.773	0.710	0.836
		Negative	5	125	130			
		Total	**19**	**127**	**146**			
Nyahururu	ZN site	Positive	23	2	25	0.830	0.779	0.881
		Negative	5	88	93			
		Total	**28**	**90**	**118**			
Narok *	ZN site	Positive	3	5	8	0.347	0.288	0.406
		Negative	4	80	84			
		Total	**7**	**85**	**92***			
Lamu *	ZN site	Positive	1	1	2	0.308	0.215	0.401
		Negative	2	14	16			
		Total	3	15	18*			

Key*: *Sample size inadequate,* ***L*** *= Lower limit,* ***U*** *= Upper limit,* ***n*** *=Number,* ***N*** *= Total number*

Table 5 shows reproducibility of FM results for different TB diagnostic tools between study sites and KEMRI research laboratory. Reproducibility of FM results was significantly different across some study sites, ranging from Kappa value = 0.667 (95% CI: 0.602 - 0.732) in Kisii to Kappa value = 0.836 (95% CI: 0.789 - 0.883) in Nyahururu. The kappa value on results for samples from Wajir was negative (-0.200 (95% CI: - 0.229 to - 0.171).

Table 5: *Reproducibility Of FM Results Between Study Sites and KEMRI Research Laboratory*

FM KEMRI							95% CI	
STUDY SITE	**Diagnostic tool**		**Positve + n**	**Negative - n**	**Total**	**Kappa value**	**L**	**U**
Busia	FM site	Positive	63	27	90	0.720	0.690	0.750
		Negative	13	493	506			
		Total	**76**	**520**	**596**			
Kitale	FM site	Positive	77	37	114	0.727	0.699	0.755
		Negative	8	448	456			
		Total	**85**	**485**	**570**			
Machakos	FM site	Positive	41	8	49	0.793	0.754	0.832
		Negative	8	177	185			
		Total	**49**	**185**	**234**			
Malindi	FM site	Positive	103	34	137	0.764	0.740	0.788
		Negative	20	933	953			
		Total	**123**	**967**	**1090**			
Wajir *	FM site	Positive	0	2	2	-0.200	-0.229	-0.171
		Negative	1	5	6			
		Total	**1**	**7**	**8***			
Kisii	FM site	Positive	12	6	18	0.667	0.602	0.732
		Negative	4	122	126			
		Total	**16**	**128**	**144**			
Nyahururu	FM site	Positive	26	2	28	0.836	0.789	0.883
		Negative	6	124	130			
		Total	**32**	**126**	**158**			
Lamu*	FM site	Positive	2	0	2	0.765	0.596	0.934
		Negative	1	13	14			
		Total	**3**	**13**	**16***			

Key*: *Sample size inadequate,* **L** *= Lower limit,* **U** *= Upper limit,* **n** *=Number,* **N** *= Total number*

Table 6: shows reproducibility of GeneXpert results between study sites and KEMRI research laboratory. Reproducibility of GeneXpert results on samples from Wajir (Kappa value = 0.614 (95% CI: 0.506-0.722)) was significantly lower than other study sites, ranging from Kappa value = 0.871 (95% CI: 0.829 - 0.912) in Kisii to Kappa value = 0.895 (95% CI: 0.831 - 0.959) in Machakos.

Table 6: *Reproducibility Of GeneXpert Results Between Study Sites and KEMRI Research Laboratory.*

GeneXpert KEMRI							95% CI	
Study site	**Diagnostic Tool**		**Positive**	**Negative**	**Total**	**Kappa value**	**L**	**U**
Busia	GeneXpert site	Positive	31	4	35	0.871	0.829	0.913
		Negative	3	124	127			
		Total	**34**	**128**	**162**			
Kitale	GeneXpert site	Positive	79	12	91	0.859	0.832	0.886
		Negative	7	270	277			
		Total	**86**	**282**	**368**			
Machakos*	GeneXpert site	Positive	12	0	12	0.895	0.831	0.959
		Negative	2	34	36			
		Total	**14**	**34**	**48***			
Malindi*	GeneXpert site	Positive	4	1	5	0.874	0.766	0.982
		Negative	0	33	33			
		Total	**4**	**34**	**38***			
Wajir*	GeneXpert site	Positive	4	2	6	0.614	0.506	0.722
		Negative	2	36	38			
		Total	**6**	**38**	**44***			

Key: **Sample size inadequate,* ***L*** *= Lower limit,* ***U*** *= Upper limit,* ***n*** *=Number,* ***N*** *= Total number*

Table 7 shows the performance of different TB diagnostic tools at study sites and KEMRI research laboratory.

Cumulatively, there was marginal significant incremental sensitivity of microscopy at the study sites ZN (69.9% (95% CI : 64.3 - 75.5) % ; n = 259) and FM (76.7% (CI: 71.1 - 82.3) %; n = 219) than at KEMRI ZN (69.9% (CI: 64.3-75.5) %; n = 259) and FM (70.8% (95% CI: 64.8 - 76.8) % ; n = 219).

However, the sensitivity of GeneXpert at the study sites (81.4% (95% CI : 71. 4 - 91.3) %; n = 59) was not statistically significantly different from that of GeneXpert at KEMRI (81.4% (95% CI : 71.4 - 91.3)% ; n = 59).

Specificity of GeneXpert at the site was not statistically significantly different from that at KEMRI but lower than that of Microscopy both at site and KEMRI.

Microscopy results varied across specific study sites but not statistically significantly different for GeneXpert. A similar pattern was observed for positive/ negative predictive values.

Table 7: *Performance Of Different TB Diagnostic Tools At Study Sites and KEMRI Research Laboratory*

Diagnostic Tool	n	PPV Sen.	(95% CI)		n	NPV Spec	(95% CI)		n	PPV Spec	(95% CI)		n	NPV Spec	(95% CI)		Total N
ZN sites	259	69.9	64.3	75.5	1367	94.7	93.5	95.9	253	71.5	66.0	77.1	1373	94.3	93.1	95.5	1626
ZN KEMRI	259	68.7	63.1	74.4	1367	93.3	92.0	94.7	269	66.2	60.5	71.8	1357	94.0	92.8	95.3	1626
FM sites	219	76.7	71.1	82.3	1189	94.2	92.9	95.5	237	70.9	65.1	76.7	1171	95.6	94.5	96.8	1408
FM KEMRI	219	70.8	64.8	76.8	1189	94.4	93.1	95.7	222	69.8	63.8	75.9	1186	94.6	93.3	95.9	1408
GeneXpert sites	59	81.4	71.4	91.3	271	88.6	84.8	92.4	79	60.8	50.0	71.5	251	95.6	93.1	98.2	330
GeneXpert KEMRI	59	81.4	71.4	91.3	271	89.3	85.6	93.0	77	62.3	51.5	73.2	253	95.7	93.1	98.2	330

Key*: 95%* ***CI:*** *Confidence Interval,* ***PPV:*** *Positive Predictive Value,* ***Sen:*** *Sensitivity,* ***Spec:*** *Specificity,* ***N:*** *Total number* ***NPV:*** *Negative Predictive Value,* ***n:*** *Number,*

Table 8 shows performance of ZN at study sites and KEMRI laboratory. Sensitivity of ZN microscopy at the satellite sites ranged from 50.0% (95% CI: 26.9 - 73.1) % in Wajir to 78.8% (95% CI: 71.4 - 86.2) % in Kitale.

Specificity was generally not statistically significantly different within the study sites. The lowest observed in Wajir (84.8% (95% CI: 79.9 - 89.7) %).

Sensitivity of ZN microscopy at non-satellite site ranged from;-

Narok: 40.0% (95% CI: 9.6 - 70.4) %
to
Nyahururu: 90.9% (95% CI: 78.9 - 100.0) %

Specificity was generally not statistically significantly different within the study sites. The lowest was observed in Kisii - Keroka (94.4% (95% CI: 90.3 - 98.4) %).

However, due to the small number of enrolled patients, especially in Lamu, sensitivity and specificity of ZN microscopy was estimated with wider confidence intervals.

Therefore, with reduced precision apart from Kitale, where the sensitivity of ZN is significantly higher than at KEMRI, the reproducibility of ZN was comparable with the study sites and KEMRI.

Study site	Diagnostic Tool	N	PPV Sen.	(95% CI)		n	NPV Spec	(95% CI)		n	PPV Spec	(95% CI)		n	NPV Spec	(95% CI)		N
Busia	ZN: site	63	69.8	58.5	81.2	535	92.7	90.5	94.9	83	53.0	42.3	63.7	515	96.3	94.7	97.9	598
Busia	ZN: KEMRI	63	68.3	56.8	79.7	535	91.6	89.2	93.9	88	48.9	38.4	59.3	510	96.1	94.4	97.8	598
Kitale	ZN: site	118	78.8	71.4	86.2	616	94.3	92.5	96.1	128	72.7	64.9	80.4	606	95.9	94.3	97.5	734
Kitale	ZN: KEMRI	118	61.0	52.2	69.8	616	95.6	94.0	97.2	99	72.7	64.0	81.5	635	92.8	90.7	94.8	734
Machakos	ZN: site	39	64.1	49.0	79.2	197	90.9	86.8	94.9	43	58.1	43.4	72.9	193	92.7	89.1	96.4	236
Machakos	ZN: KEMRI	39	66.7	51.9	81.5	197	91.9	88.1	95.7	42	61.9	47.2	76.6	194	93.3	89.8	96.8	236
Malindi	ZN: site	112	72.3	64.0	80.6	976	95.5	94.2	96.8	125	64.8	56.4	73.2	963	96.8	95.7	97.9	1088
Malindi	ZN: KEMRI	112	68.8	60.2	77.3	976	95.6	94.3	96.9	120	64.2	55.6	72.7	968	96.4	95.2	97.6	1088
Wajir	ZN: site	18	50.0	26.9	73.1	204	84.8	79.9	89.7	40	22.5	9.6	35.4	182	95.1	91.9	98.2	222
Wajir	ZN: KEMRI	18	55.6	32.6	78.5	204	89.2	85.0	93.5	32	31.3	15.2	47.3	190	95.8	92.9	98.6	222
Kisii	ZN: site	22	45.5	24.6	66.3	124	95.2	91.4	98.9	16	62.5	38.8	86.2	130	90.8	85.8	95.7	146
Kisii	ZN: KEMRI	22	54.5	33.7	75.4	124	94.4	90.3	98.4	19	63.2	41.5	84.8	127	92.1	87.4	96.8	146
Nyahururu	ZN: site	22	90.9	78.9	102.9	96	94.8	90.3	99.2	25	80.0	64.3	95.7	93	97.8	94.9	100.8	118
Nyahururu	ZN: KEMRI	22	86.4	72.0	100.7	96	90.6	84.8	96.5	28	67.9	50.6	85.2	90	96.7	93.0	100.4	118
Narok	ZN: site	10	40.0	9.6	70.4	82	95.1	90.5	99.8	8	50.0	15.4	84.6	84	92.9	87.3	98.4	92
Narok	ZN: KEMRI	10	40.0	9.6	70.4	82	96.3	92.3	100.4	7	57.1	20.5	93.8	85	92.9	87.5	98.4	92
Lamu	ZN: site	2	50.0	19.3	119.3	16	93.8	81.9	105.6	2	50.0	19.3	119.3	16	93.8	81.9	105.6	18
Lamu	ZN: KEMRI	2	0.0	0.0	0.0	16	81.3	62.1	100.4	3	0.0	0.0	0.0	15	86.7	69.5	103.9	18

Key: **Sample size inadequate, **95%CI:** Confidence Interval, **Sen:** Sensitivity, **Spec:** Specificity, **PPV**: Positive Predictive Value, **NPV:** Negative Predictive Value, **n**: Number, **N:** Total number*

Table 8: *Performance Of ZN Microscopy At Study Sites and Kemri Research Laboratory.*

Table 9 shows performance of FM at study sites and KEMRI laboratory. Sensitivity of FM Microscopy at study sites ranged from 68.2% (95% CI: 58.5 - 77.9) % in Kitale to 87.5% (95% CI: 80.6 - 94.4) % in Kitale.

Specificity was generally not statistically significantly different within the study sites. The lowest observed in Machakos (89.2%(95% CI:84.9-93.6) %). Sensitivity of FM Microscopy at the study sites within non-satellite sites ranged from 45.5% (95% CI:24.6-66.3) % in Kisii to 91.7% (95% CI: 80.6 - 100) % in Nyahururu.

Specificity was generally not statistically significant different within the study sites with the lowest observed in Kisii-Keroka (92.5% (95% CI: 88.1 - 97.0) %).

Due to the small number of enrolled patients, especially in Wajir and Lamu, sensitivity and specificity of FM microscopy was estimated with wider confidence intervals. Therefore, with reduced precision.

Table 9: *Performance Of FM Microscopy At Study Sites and Kemri Research Laboratory.*

Study site	Diagnostic tool	n	PPV Sen.	(95% CI)		n	NPV Spec	(95% CI)		n	PPV Spec	(95% CI)		n	NPV Spec	(95% CI)		Total N
Busia	*FM: site*	61	82.0	72.3	91.6	535	92.5	90.3	94.8	90	55.6	45.3	65.8	506	97.8	96.6	99.1	596
Busia	*FM: KEMRI*	61	68.9	57.2	80.5	535	93.6	91.6	95.7	76	55.3	44.1	66.4	520	96.3	94.7	98.0	596
Kitale	*FM: site*	88	87.5	80.6	94.4	482	92.3	89.9	94.7	114	67.5	58.9	76.1	456	97.6	96.2	99.0	570
Kitale	*FM: KEMRI*	88	68.2	58.5	77.9	482	94.8	92.8	96.8	85	70.6	60.9	80.3	485	94.2	92.2	96.3	570
Machakos	*FM: site*	39	76.9	63.7	90.1	195	90.3	86.1	94.4	49	61.2	47.6	74.9	185	95.1	92.0	98.2	234
Machakos	*FM: KEMRI*	39	71.8	57.7	85.9	195	89.2	84.9	93.6	49	57.1	43.3	71.0	185	94.1	90.6	97.5	234
Malindi	*FM: site*	110	78.2	70.5	85.9	980	94.8	93.4	96.2	137	62.8	54.7	70.9	953	97.5	96.5	98.5	1090
Malindi	*FM: KEMRI*	110	72.7	64.4	81.1	980	95.6	94.3	96.9	123	65.0	56.6	73.5	967	96.9	95.8	98.0	1090
Wajir*	*FM: site*	0	n/a	n/a	n/a	8	75.0	45.0	100	2	0.0	0.0	0.0	6	100	100	100	8*
Wajir*	*FM: KEMRI*	0	n/a	n/a	n/a	8	87.5	64.6	100	1	0.0	0.0	0.0	7	100	100	100	8*
Kisii	*FM: site*	22	45.5	24.6	66.3	122	93.4	89.1	97.8	18	55.6	32.6	78.5	126	90.5	85.4	95.6	144
Kisii	*FM: KEMRI*	22	50.0	29.1	70.9	122	95.9	92.4	99.4	16	68.8	46.0	91.5	128	91.4	86.6	96.3	144
Nyahururu	*FM: site*	24	87.5	74.3	100	134	94.8	91.0	98.5	28	75.0	59.0	91.0	130	97.7	95.1	100	158
Nyahururu	*FM: KEMRI*	24	91.7	80.6	100	134	92.5	88.1	97.0	32	68.8	52.7	84.8	126	98.4	96.2	100	158
Lamu*	*FM: site*	2	50.0	0.0	100	14	92.9	79.4	100	2	50.0	19.3	100	14	92.9	79.4	100	16*
Lamu*	*FM: KEMRI*	2	50.0	0.0	100	14	85.7	67.4	100	3	33.3	20.0	86.7	13	92.3	77.8	100	16*

Key: **Sample size inadequate, 95% **CI:** Confidence Interval, **Sen:** Sensitivity, **Spec:** Specificity, **PPV**: Positive Predictive Value, **NPV:** Negative Predictive Value, **n**: Number, **N:** Total number*

Table 10 shows performance of GeneXpert at study sites and KEMRI laboratory. The Sensitivity of GeneXpert at study sites ranged from 66.7% (95% CI: 13.3 - 100.0) % in Malindi to 94.9% (95% CI: 70.7 - 100) % in Kitale. Specificity was generally comparable with the lowest observed in Machakos (75.6% (95% CI: 63.0 - 88.1) %).

However, due to the small number of specimens analyzed using GeneXpert at both study sites and KEMRI, especially in Wajir, Malindi and Machakos, the sensitivity and specificity of GeneXpert was estimated with wider confidence intervals, in this case, with reduced precision.

Table 10: *Performance Of GeneXpert at Study Sites and KEMRI Research Laboratory.*

Study site	Diagnostic tool	n	PPV Sen	(95% CI)		n	NPV Spec	(95% CI)		n	PPV	(95% CI)		n	NPV	(95% CI)		N
Busia	GeneXpert: site	21	85.7	70.7	100	141	87.9	82.6	93.3	35	51.4	34.9	68.0	127	97.6	95.0	100	162
Busia	GeneXpert: KEMRI	21	85.7	70.7	100	141	88.7	83.4	93.9	34	52.9	36.2	69.7	128	97.7	95.0	100	162
Kitale	GeneXpert: site	59	94.9	89.3	100	309	88.7	85.1	92.2	91	61.5	51.5	71.5	277	98.9	97.7	100	368
Kitale	GeneXpert: KEMRI	59	93.2	86.8	100	309	90.0	86.6	93.3	86	64.0	53.8	74.1	282	98.6	97.2	100	368
Machakos *	GeneXpert: site	3	100	100	100	45	80.0	68.3	91.7	12	25.0	0.5	49.5	36	100	100	100	48*
Machakos *	GeneXpert: KEMRI	3	100	100	100	45	75.6	63.0	88.1	14	21.4	0.0	42.9	34	100	100	100	48*
Malindi *	GeneXpert: site	3	66.7	13.3	100	35	91.4	82.2	100	5	40.0	0.0	82.9	33	97.0	91.1	100	38*
Malindi *	GeneXpert: KEMRI	3	66.7	13.3	100	35	94.3	86.6	100	4	50.0	1.0	99.0	34	97.1	91.4	100	38*
Wajir *	GeneXpert: site	4	0.0	0.0	0.0	40	85.0	73.9	96.1	6	0.0	0.0	0.0	38	89.5	79.7	99.2	44*
Wajir *	GeneXpert: KEMRI	4	0.0	0.0	0.0	40	85.0	73.9	96.1	6	0.0	0.0	0.0	38	89.5	79.7	99.2	44*

Key: ***95% CI:*** *Confidence Interval,* ***Sen:*** *Sensitivity,* ***Spec:*** *Specificity,* ***N:*** *Total number*
PPV: *Positive Predictive Value,* ***NPV:*** *Negative Predictive Value,* ***n:*** *Number,*

Discussion

This study the first to document comparison on reproducibility of laboratory results and performance of the TB diagnostic tools in different geographical settings. Suggesting that, there was no statistically significant difference between reproducibility and performance of GeneXpert in all study sites including KEMRI.

This suggests that, under ideal conditions GeneXpert MTB/RIF is a reliable diagnostic tool irrespective of the facility setting.

Furthermore, the observations support the sentiments expressed in the WHO document on "Frequently Asked Questions on GeneXpert MTB/RIF assay" which indicates that the "new technology is a relatively low throughput technology and easy to use, it may be better suited at lower levels of the health care system (e.g. at County or Sub - County level).

Access to appropriate treatment for a case diagnosed with the new technology is also essential, and implementation of the new technology needs to be linked to treatment services" (www.who.int/tb/labosratory/xrepert_faqs.pdf).

However, there are several pertinent factors that require redress prior to implementation of GeneXpert. These include adequate infrastructure without interrupted power supply, guranteed supply of cartridges, other potential consumables, frequent scheduled machine services such as maintenance and calibrations.

These factors if not adequately addressed and monitored could cause unnecessary delays in the diagnosis and subsequent mismanagement of TB patients.

Furthermore, the significantly lower specificity of GeneXpert than microscopy both at site and KEMRI, indicate that, microscopy could compliment GeneXpert. This is especially in settings with inadequate capacity including Infrastructure, Human Resource and High workload.

Generally, reproducibility of results for the ZN and FM microscopy varied across the sites but were not significantly different for GeneXpert.

The performance of ZN and FM in this study indicated that the results are consistent with those from a previous study conducted using the same population and it was shown that although sensitivity of FM was higher than that of ZN, the difference was not statistically significant [2]

Similar observations were made in a systematic review by Steingart KR, et al, 2006. The wider confidence intervals observed in sensitivity and specificity of GeneXpert observed in some study sites was mainly due to low numbers of samples involved in the analysis resulting in reduced precision in estimates.

There is need for continuous training in microscopy to enhance reproducibility of laboratory results and performance.

Acknowledgement

We thank the Ministry of Health officials especially:

(i). The Medical Superintendents from Busia, Kitale, Machakos, Malindi, Wajir, Kisii, Nyahururu, Narok and Lamu referral hospitals.
(ii) Their Laboratory staff for the technical support

in the preparation of research materials for the East African Public Health Laboratories Networking Project (EAPHLNP),

The East, Central and Southern Africa Health Community (ECSA-HC) - for coordination of the project,

The Director KEMRI and staff for facilitating the smooth running of the project.

Finally The World Bank through the Ministry of Health- Kenya, for the financial support

References

1. Sub-committee E11.20 on Test Method Evaluation and Quality Control. Standard Practice for Use of the Terms Precision and Bias in ASTM Test Methods, (2014), *ASTM International, ASTM E177.* From Wikipedia, the free encyclopedia.

2 **Githui W.A1, Mwangi M., Orina F., Kiptoo M., Ogaro T , Wanzala P., Sang W.K., Omar S., Kariuki J.N**. Performance of Ziehl-Neelsen Microscopy, Light Emitting Diodes – FM and Xpert MTB/RIF in the Diagnosis of Tuberculosis in People with Presumptive TB from EAPHLNP study sites in Kenya. *AJHS* October - December 2014 ISSN 1022-9272 Volume 24, Number 4 (Supplement):433.

3. Frequently Asked Questions on Xpert MTB/RIF assay ... GeneXpert, the test device ... sensitivity superior to that of conventional microscopy or culture on solid media. www.who.int/tb/laboratory/xpert_faqs.pdf.

4 **Keeler E, Perkins MD, Small P, Hanson C, Reed S, Cunningham J,** *et al.* Reducing the global burden of tuberculosis: the contribution of improved diagnostics. Nature 2006;444:49–57,

5. **Habte D, Melese M, Hiruy N, Gashu Z, Jerene D, Moges F, Yifru S, Girma B, Kassie Y, Kebede HY, Suarez PG, Tessema B.** The additional yield of GeneXpert MTB/RIF test in the diagnosis of pulmonary tuberculosis among household contacts of smear positive TB cases. *International Journal of Infectious Diseases* 49 (2016) 179-184.

6. **Steingart KR, Ramsay A, Pai M.** Optimizing sputum smear microscopy for the diagnosis of pulmonary tuberculosis. *Expert review of anti-infective therapy* 2007;5(3):327–31.

7. **Steingart KR1, Henry M, Ng V, Hopewell PC, Ramsay A, Cunningham J, Urbanczik R, Perkins M, Aziz MA, Pai M**. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. *Lancet Infect* Dis. 2006 Sep;6(9):570-81.