Inter / Intra Examiner Calibration of Clinical Signs and Symptoms in Comparison with Laboratory Results in the Diagnosis of Pulmonary Tuberculosis in Patients from EAPHLN Project Study Sites in Kenya

*Wanzala P.[2], Githui W.A[1], Mwangi M.[2], Kiptoo M.[3], Sang W.K.[4], Orina F[1].,Kariuki J.N.[2], Ogaro T[5].,

1. *Center for Respiratory Disease Research (CRDR), Kenya Medical Research Institute*
2. *Center for Public Health Research (CPHR), Kenya Medical Research Institute*
3. *KEMRI Graduate School of Health Sciences (KGSHS), Kenya Medical Research Institute*
4. *Center for Microbiology Research (CMR), Kenya Medical Research Institute*
5. *Division of TB, Leprosy & Lung Disease, Ministry of Health Kenya*

Corresponding Author Email : wanzap2003@yahoo.com

The East Africa Public Health Laboratory Networking Project (EAPHLNP) is a regional project involving five East African countries, Namely: Burundi, Kenya, Rwanda, Uganda and Tanzania and it is supported by the World Bank.

Summary

INTRODUCTION

The clinical diagnosis of pulmonary tuberculosis (PTB) is based on occurrence of four cardinal signs and symptoms which include current cough, night sweats, weight loss, or low grade fever. However few studies have determined the validity and reliability of these diagnoses by intra and inter-examiner calibration of clinicians for appropriateness of detection of tuberculosis (TB) in resource constrained settings.

OBJECTIVE

The study's aim was to determine the sensitivity of concordance and reliability (Kappa values) of inter-examiner and intra-examiner findings of clinicians in the EAPHLN project.

METHODOLOGY

The study was a cross-sectional study in nine sites. It included 155 patients for intra-examiner and 57 patients for inter-examiner calibrations selected from eligible people with symptoms or signs suggestive of TB during the implementation of the East African Public Health Laboratory Network Project (EAPHLNP) in Kenya. TB clinical symptoms and signs were recorded in a structured medical form included the following: productive cough, weight loss, night sweats, low grade fever (classical cardinal signs and symptoms). Using quality assurance sampling for a total population of ten thousand people with symptoms or signs suggestive of TB from the sites with a minimum defective sample acceptable of 0 and a probability of defect accepted of 1% and an alpha of 5%, the sample size of repeatable samples is 262 for total patients for the sites per year. Intra-examiner calibration involved examination of the same patient independently by the same clinician within one day interval. Inter-examiner calibration involved examination of the same patient by two clinicians independently the same day. Calibration of the clinical tools used during examination of patients was done. TB laboratory diagnosis was first done by sputum

smear microscopy Ziehl–Neelsen stain.(ZN), secondly by optimized sputum smear microscopy with a Light Emitting Diode microscope (LED) or fluorescent microscopy(FM), and thirdly by Gene-expert technique (Gene Xpert or Gx). The results from the clinicians and reference laboratory findings for these patients were entered in a computer, verified and analyzed in SPSS for reliability statistics. These unweighted Cohen Kappa scores were interpreted as follows: poor 0.01–0.20, moderate 0.21–0.40, fair 0.41–0.60, good 0.61–0.80, or excellent 0.81–1.0 based on the agreement between the intraexaminer and inter-examiner findings

RESULTS

A significant difference was found between concordant diagnosis of a least 4 signs and symptoms of TB compared to fewer by the same examiner on the same patient in all TB test/HIV status categories except the ZN positive /HIV positives and GeneXpert negative /HIV positives and HIV negative categories. The highest sensitivity rate was 81.8 %(95%CI=52.3-94.9) in the Gx+ve/HIV+ve category. The significance difference in senstivity results of TB/HIV test *vs* at least presence of the 4 signs and symptoms. However, This did not occur in ZN+ve/HIV+ve, FM+ve/HIV+ve , Gx+ve/HIV+ve Gx+ve/HIV-ve categories. Kappa values for cough and fever were consistently significantly higher than zero kappa.

CONCLUSION

Excellent kappa can be achieved in low resource settings by clinician using all four cardinal signs and symptoms of TB with laboratory results. There is possibility of using the clinical diagnosis using the four signs and symptoms where laboratory diagnosis is not present but specificity is low. Good clinical practice would improve the specificity.

[Afr J Health Sci. 2014; 27(4) Supplement : 450-4468] [***Afr. J.*** Health Sci. 2019 32(6) : 18 - 31]

Introduction

The World Health Organization(WHO) previously recommended a cough duration of two (2) or more weeks as one of the cardinal symptoms to screen for pulmonary tuberculosis (TB). TB signs and symptoms can range from none(in 10 to 20%) to the maximum.

However, this has been indicated as inadequate for HIV-associated TB, with the sensitivity as low as 50% in this patient group [1, 2, 3]. Screening tools that combine multiple symptoms have much higher sensitivity, but accompanied by low specificity.

A meta-analysis of 10,000 HIV - infected patients actively screened for TB was conducted to identify an optimum symptom screening *algorithm*. The optimum *algorithm* identified patients with at least 1 of 4 cardinal signs and symptoms (current cough, night sweats, weight loss, or low grade fever (<38.5 0)) with a sensitivity of 79% and a specificity of 50%. [4]. TB diagnosis is initiated by patients presenting with signs and symptoms at health facilities. The classification of TB disease include:

Class: 0– No infection, not infected

Class: 1– TB exposure, not infected

Class: 2– Latent TB infection, no disease

Class: 3–TB : clinically active

Class: 4–TB ; not clinically active

Class: 5–Person with symptoms or signs suggestive of TB diagnosis pending[5].

Although primarily, disease essentially exists sub-clinically, some self-limiting findings might be noticed in an assessment. Signs and symptoms of TB include insidious onset and chronic course, chest symptoms (cough, usually productive, hemoptysis, chest pain (usually pleuritic), non-specific constitutional symptoms, and extra-pulmonary symptoms.

Immuno-suppression affects presentation of these signs and symptoms with differential presentations in HIV positives with the pathophysiology resulting in lower immuno-reactivity so there is less cavitation, inflammation and endobronchial irritation. Variations in signs and symptoms may differ by the progress of disease from indolent and fulminant, local to systemic, and any organ or tissue (25% of the disease being extra-pulmonary).

However, over 90% of patients with sputum smear positive pulmonary TB develop cough at disease onset [6]. When a patient progresses to active TB, early signs and symptoms are often nonspecific [10]. Manifestations often include progressive fatigue, malaise, weight loss, and a low-grade fever accompanied by chills and night sweats. Wasting, a classic feature of TB is due to the lack of appetite and the altered metabolism associated with the inflammatory and immune responses. Wasting involves the loss of both fat and lean tissue; the decreased muscle mass contributing to the fatigue. Associated paratracheal *lymphadenopathy* may occur because the of *bacilli* spread from the lungs through the *lymphatic* system.

If the primary lesion enlarges, pleural effusion is a distinguishing finding. This effusion develops because the *bacilli* infiltrate the pleural space from an adjacent area. The effusion may remain small and resolve spontaneously, or it may become large enough to induce symptoms such as fever, pleuritic chest pain, and *dyspnea. Dyspnea* is due to poor gas exchange in the areas of affected lung tissue. Dullness to percussion and a lack of breath sounds are physical findings indicative of a pleural effusion because excess fluid has entered the pleural space. However physical sign have a broad differential diagnostic ability since they occur in other chest conditions [6, 7, 8, 9, 19].

Tuberculosis(TB), including latent TB infection and active TB infection affect the following sites:

- Respiratory (lung, bronchus, pleura, thoracic
- lymph nodes)
- meninges
- pericardium
- bone and joints
- peripheral lymph nodes
- genitourinary system or is disseminated (including military TB).

Comorbidities such as diabetes, hepatic disease, renal disease, or mental illness can occur in TB patients [7]. Immune Reconstitution Inflammatory Syndrome (IRIS) also occurs in HIV infected patients with TB [11, 12, 13, 14, 17].

Sputum smear microscopy was commonly used for diagnosis of pulmonary TB in Kenya. However, patients could also be managed on clinical diagnosis [19]

The aim of the study was to assess the correlation on cardinal TB signs and symptoms between two same clinician repeat intraexaminer calibrations and inter-examiner calibration by two clinicians at each study site on the same patient[15,16] Clinical findings were compared with laboratory test results.

Methodology

In this cross-sectional study 155 patients for intra-examiner and 57 patients for inter-examiner calibrations were selected from eligible people with symptoms or signs suggestive of TB during the implementation of the East African Public Health Laboratory Network Project (EAPHLNP) in nine sites in Kenya. TB clinical symptoms and signs were recorded in a structured medical form included the following: productive cough, weight loss, night sweats, low grade fever (classical cardinal signs and symptoms), general patient condition, body temperature, body mass index (BMI), blood pressure, pulse, respiratory rate, heart sounds (vital signs and symptoms), conjunctival presentation (aneamia), neck swellings (lymphadenopathy), difficulty in breathing, ease of air entry (respiratory signs and symptoms) and percussion and auscultation findings. These were obtained by the history of morbidities and co-morbidities and physical examination of the enrolled patients.

Using quality assurance sampling for a total population of ten thousand people with symptoms or signs suggestive of TB from the satellite sites with a minimum defective sample acceptable of 0 and a probability of defect accepted of 1% and an alpha of 5%, the sample size of repeatable samples was 262 for total patients for the satellite sites per year. This gave a sampling interval of approximately 50 patients per site per year. This translated to 1 randomly selected patient per site per week. Matching was done by the site clinical supervisor to ensure that every examiner was paired

with another examiner so that the recommended number of matched results was obtained for each examiner. The results were in custody of the site clinical supervisor and released quarterly to the study team. The serial number of which patient to be selected was computer generated. These randomly selected recruited eligible patient suspects in the study then underwent intra-examiner (same clinician making two independent examinations on same patient at recruitment and at delivery of morning sample) and between clinicians or inter-examiner (two clinicians making two independent examinations on the same patient) the same day. Calibration of the weight, height and temperature tools for consistency were also done at the different sites. TB diagnosis was;

1. First done by sputum smear microscopy Ziehl–Neelsen stain (ZN),
2. Secondly by optimized sputum smear microscopy with a Light Emitting Diode microscope (LED) or fluorescent microscopy (FM),
3. Thirdly by GeneXpert technique (GeneXpert or Gx).

The results from the clinicians and reference laboratory findings for these patients were entered in a computer, verified and analyzed in SPSS for reliability statistics of the screening performance of all combinations and separate (signs and symptoms vs TB laboratory diagnosis) variables of interest. These unweighted Cohen Kappa scores were interpreted as follows: poor 0.01–0.20, moderate 0.21–0.40, fair 0.41–0.60, good 0.61–0.80, or excellent 0.81–1.0 based on the agreement between the intra-examiner and inter-examiner findings in different TB laboratory diagnosis/ HIV categories [15, 16, 18].

Results

Concordance of the intra-examiner and interexaminer on the examination of at all least four, any combination at least any of 3 , at least any two and at least one cardinal TB signs and symptoms was determined and the sensitivity of the concordance performed by clinicians as shown in *Tables 1a,1b and Table 3.*

Table 1a: *Sensitivity of The Concordance of Different Combinations of Signs and Symptoms for Intra-Examiner Calibration*

TB test	HIV status	n;-% (95% CI)					Total N (100%)
		None	AT least 1 TB sign and symptom	AT least 2 TB signs and symptoms	At least 3 TB signs and symptoms	At least 4 TB signs and symptoms	
ZN+ve	+ve	0	0	0	5:-50(23.7- 76.3)	5:-50(23.7-76.3)	10
ZN+ve	-ve	0	1:-4.5(0.8-21.8)	2:-9.1(2.5-27.8)	3:-13.6(4.5-33.3)	16:72.7(51.986.9) a 22	
ZN-ve	+ve	0	0	2:-8(2.2-25)	3:-12(4.2-30)	20:80(60.9-91.1) a	25
ZN-ve	-ve	1:1.3(0.27.1)	1:-1.3(0.2-7.1)	5:-6.6(2.8-14.5)	12:-15.8(9.3-25.6) aa	57:75(64.2 83.4)a	76
FM+ve	+ve	0	0	0	4:44.4(18.973.3)	5:56.6(26.781.1)	9
FM+ve	-ve	0	1:-5.3(0.9-24.6)	1:-5.3(0.9-24.6)	3:-15.8(5.5-37.6)	14:73.7(51.288.2) a	19

Table 1b: *Sensitivity of The Concordance of Different Combinations of Signs and Symptoms for Intra- Examiner Calibration*

TB test	HIV status	n;-% (95% CI)					Total N (100%)
		None	AT least 1 TB sign and symptom	AT least 2 TB signs and symptoms	At least 3 TB signs and symptoms	At least 4 TB signs and symptoms	
FM-ve	+ve	0	0	2:-8(2.2-25)	3:-12(4.2-30)	20:80(60.9-91.1) [a]	25
FM-ve	-ve	1:1.5(0.38.5)	1:-1.5(0.3-8.5)	5:-7.4(3.4-17.3)	12:17.6(11.3-30.4) [aa]	44:72.1(59.679.8) [a]	63
Gx+ve	+ve	0	0	1:-9.1(1.6-37.7)	1:-9.1(1.6-37.7)	9:81.8(52.394.9) [a]	11
Gx+ve	-ve	1:4.5(0.8-21.8)	0	4:18.2(7.3-38.5)	3:-13.6(4.8-33.3)	14:63.6(43-80.3) [a]	22
Gx-ve	+ve	0	0	0	3:-50-(18.8- 81.2)	3:-50(18.8- 81.2)	6
Gx-ve	-ve	0	0	1:-11.1(2-43.5)	4:44.4(18.973.3)	4:44.4(18.973.3)	9

Key: a means p-value <0.05 considered statistically significant difference between concordant diagnosis of
[a]: least 4 signs and symptoms of TB compared to fewer by the same examiner on the same patient.
[aa]: means p-value <0.05 considered statistically significant difference between concordant diagnosis of
[a]: least 3 signs and symptoms of TB compared to fewer by the same examiner on the same patient.

ZN+ve/--ve = ***Ziehl***–Neelsen stain positive and negative.
FM+ve/--ve = Fluorescent microscopy positive and negative.
Gx+ve/--ve = Gene-Xpert positive and negative.
HIV+ve/-ve = Human Immuno-deficiency Virus positive and negative.

Significant differences were found between concordant diagnosis of a least 4 signs and symptoms of TB compared to fewer than 4 by the same examiner on the same patient in all TB test/HIV status categories except the ZN +ve / HIV +ve and GeneXpert negative / HIV positive or HIV negative categories. Statistically significant difference between concordant diagnosis of a least 3 signs and symptoms of TB compared to fewer by the same examiner on the same patient was found in the ZN negative/HIV negative category and FM negative/HIV negative category.

The highest concordance occurred in GeneXpert positive / HIV positive category. One patient had none of the four cardinal signs and symptoms of TB present

(Tables 1a, 1b).

The crude odds ratio (OR) results of odds ratios of TB test vs at least presence of the 4 signs and symptoms indicate that the crude OR for ZN is significantly different from the HIV stratum specific and adjusted ORs for this test. HIV is a confounding factor in ZN testing for TB. No significance was found between the crude stratum specific and adjusted odds ratios when the other TB tests were stratified for HIV. The values were not significantly different from those of HIV stratum specific and adjusted odds ratios for ZN *(Table 2).*

Table 2: *OR of Having TB Vs. Combination of 4 Signs and Symptoms During Intra-Examiner Calibration*

		Odds ratio (95% CI)	Odds ratio (95% CI)	p-value [a]
TB test	**HIV test**			
ZN				
	+ve	.250(.051-1.214)	4.32(1.75,-11)[aaa]	.189 [a]
	-ve	889(.304-2.598)	.601 (.254-1.421) aa	
FM				
FM	**+ve**	1.209 (.381-3.835)	0.79 (0.29,-2.28) [aaa]	.181 [a]
	-ve	0.79(0.29-2.28)	.790(.316-1.973) [aa]	
Zx				
	+ve	4.500((.491-41.248)	2.63(0.62- 11.08) [aaa]	.602 [a]
	-ve	2.188(.452-10.576)	2.777(.774-9.956) [aa]	

Key: [a]: Breslow-Day test of Homogeneity of the Odds Ratios, [aa]: Adjusted odds ratio, [aaa]: Crude odds ratio

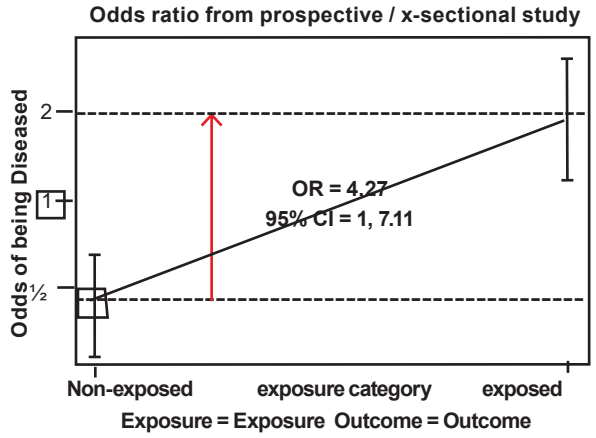

Figure1. *ZN odds in exposed and none-exposed*

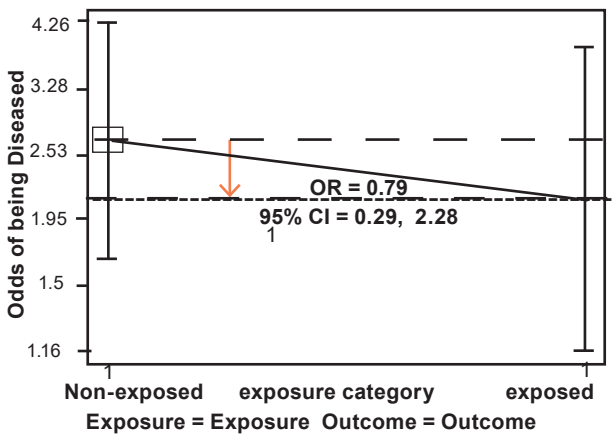

Figure 2 *FM odds in exposed and none-exposed*

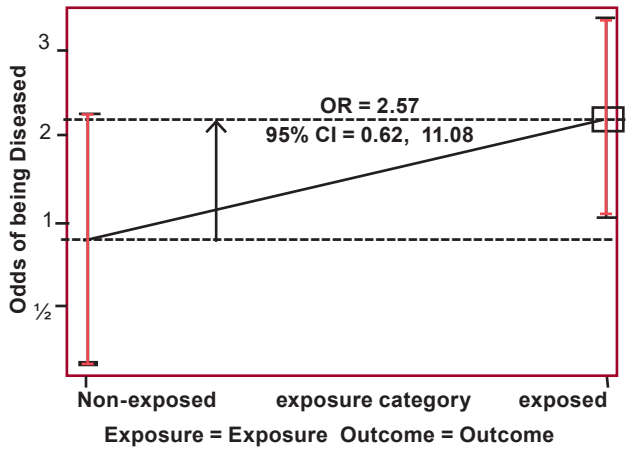

Figure 3. *GX Odds in Exposed and None-Exposed*

Key: Exposure-Positive TB test, Non-exposure-Negative TB test, Diseased- Having at least 4 of the cardinal clinical TB symptoms and signs

Figures 1, 2 and ***3*** are diagrams of the nonexposed and exposed odds including the odds ratios. The directions of the odds in nonexposed and exposed in the FM test differ in the opposite direction (increasing or decreasing) from the odds in the ZN and Gx tests. This indicates that although the sensitivity in table 1 for ZN+ve/HIV+ve and FM+ve/HIV+ve are not significant for the two tests (at least all four signs and symptoms), there is still variations in the direction of the change in the odds in non-exposure vs exposure.

Table 3a: *Sensitivity of The Concordance of Different Combinations of Signs and Symptoms for Inter-Examiner Calibration*

TB test	HIV status	n:-% (95% CI)					Total N-(100%)
		0	At least 1 TB sign and symptom	At least 2 TB signs and symptoms	At least 3 TB signs and symptoms	At least 4 TB signs and symptoms	
ZN+ve	+ve	0	0	0	0	2:-100	2
ZN+ve	-ve	0	0	1:-16.7(3-56)	4:-66.7(30-90)	1:-16.7(30-56)	6
ZN-ve	+ve	0	0	0	5:-41.7(21.3-72)	7:-58.3(35.4-84.8)	11
ZN-ve	-ve	0	0	1:-7.7(1.4-33.3)	5:-38.5(17.7-64.5)	7:-53.9(29.1-76.8)	13
FM+ve	+ve	0	0	0	0	2:-100	2
FM+ve	-ve	0	0	1:-16.7(3-56.3)	3:-50(18.8-81.2)	2:-33.3(9.7-70)	6
FM-ve	+ve	0	0	0	6:-41.7(23.2-70.9)	7:-58.3(29.1-76.8)	13
FM-ve	-ve	0	0	1:-7.7(1.4-33.3)	5:-18.5(17.7-64.5)	7:-77.8(29.1-76.8)	13 459

Table 3b: Sensitivity of The Concordance of Different Combinations of Signs and Symptoms for Inter-Examiner Calibration

TB test	HIV status	n:-% (95% CI)					Total N(100%)
Gx+ve	+ve	0	0	0	0	4:-100	4
Gx+ve	-ve	0	0	0	2:-66.7(20.8-93.1)	1:-33.3(6.1-79.2)	3
Gx-ve	+ve	0	0	0	1:-33.3(6.1-79.2)	2:-66.7(20.8-93.9)	3
Gx-ve	-ve	0	0	1:-16.7(3-56.3)	1:-16.7(3-56.3)	4:-66.7(30-90.3)	6

Key: ***ZN*+ve / --ve *Ziehl*** – Neelsen stain positive and negative,
FM+ve/--ve Fluorescent Microscopy positive and negative.

Gx+ve/--ve GeneXpert positive and negative, HIV+ve/ve Human Immunodeficiency Virus positive and negative

No statistically significant difference was found FM positive /HIV positive and Gx positive /HIV between the different combinations of signs and positive categories albeit the low numbers symptoms. The highest concordance occurred among results of those with at least 4 signs and in the ZN+ve/HIV+ve, symptoms of TB (***Tables 3a,3b)***.

Table 4: OR of having TB Vs. Combination of 4 Signs and Symptoms During Inter Examiner Calibration

TEST of TB test vs the 4 signs and symptoms	HIV test	Odds ratio (95% CI)	Odds ratio (95% CI)	p-value
TB test				
ZN				
	+VE	1.905(0.074-6.92)	2.33(0.35- 18.15) a	.392 aaa
	-VE	.171(.015-1.905)	.346((.072-1.677) aa	
FM				
	+ve	.538(.326-.891)	0.64(0.09-4.48) a	0.136 aaa
	-ve	.429(0.057-3.22)	.971(.191-4.93) aa	
Gx				
	+ve	0.667(0.3-1.484)	1.25(0.1-20.54) a	0.117 aaa
	-ve	0.25(0.03-4.729)	0.893(0.105-7.564) aa	

Key: [a]**:** Crude odds ratio estimates
[aa]**:** Mantel-Haenszel Common Odds Ratio Estimate
[aaa]**:** Breslow-Day test of Homogeneity of the Odds Ratio

Although the crude ORs results of TB test vs specific and adjusted odds ratio for this test. No the 4 signs and symptoms indicate that the significant differences were found for the HIV crude odds ratio for ZN is still higher but not stratum specific odds ratios for the HIV tests significantly different from the HIV stratum ***(Table 4).***

Table 5a: *Intra Examiner Performance Of Specific Signs and Symptoms*

TB test	HIV status	**Kappa**, p-value-(n)			
		Cough	**Weight Loss**	**Night Sweats**	**Low Grade Fever**
ZN+ve	+ve	0.615, 0.035-(10)*	N/A	N/A	0.615, 0.035-(10)*
ZN+ve	-ve	0.79, 0.00 -(113)*	0.241, 0.221-(22)	-0.048, 0,823,-(22)	0.645, 0.001(22)*
ZN-ve	+ve	N/A	1.00, 0.00-(25)*	0.359, 0.019-(25)*	0.648, 0.001-(25)*
ZN-ve	-ve	0.774, 0.00-(76)*	0.541, 0.000-(76)*	0.537, 0.000-(76)*	0.573, 0.000(76)*
FM+ve	+ve	0.609, 0.047-(9)*	0.496, 0.012-(19)*	N/A	N/A
FM+ve	-ve	N/A	N/A	N/A	0.642, 0.003-(19)*
FM-ve	+ve	0.778, 0.000-(25)*	1, 0.00 -(25)*	0.359, 0.019-(25)*	0.648, 0.001-(25)*
FM-ve	-ve	0.751, 0.000-(68)*	0.501, 0.00-(68)*	0.532, 0.000-(68)*	0.570, 0.000(68)

*

Table 5b: *Intra Examiner Performance Of Specific Signs and Symptoms*

TB test	HIV status	**Kappa,** p-value-(n)			
		Cough	**Weight Loss**	**Night Sweats**	**Low Grade Fever**
Gx+ve	+ve	1.00 , 0.014-(6)*	N/A	-0.200,0.624-(6)	N/A
Gx+ve	-ve	N/A	0.400, 0.134-(9)	N/A	1 ,0.003-(9)*
Gx-ve	+ve	0.621, 0.025-(11)*	-0.138,0.484-(4)	N/A	N/A
Gx-ve	-ve	0.748, 0.000-(22)*	O.501,0.00-(68)*	N/A	0.463,0.010,-(22)*

Key: ***ZN*** - Ziehl–Neelsen , ***FM*** - Fluorescent microscopy ***Gx*** - Gene Xpert
* means ***p-value <0.05*** is considered statistically significant. ***N/A*** – Not available

Table 6: Inter-Examiner Performance of Specific Signs and Symptoms

TB test	HIV status	Kappa, p-value-(n)			
		cough	weight loss	night sweats	low grade fever
ZN+ve	+ve	**1**, 0.157-(2)	N/A	N/A	N/A
ZN+ve	-ve	**0.22**, 0.350-(7)	N/A	**0.58,** 0.088-(7)	N/A
ZN-ve	+ve	**0.82,** 0.004-(12)*	**0.429,** 0.070-(12)	N/A	**-0.091**, 0.753-(12)
ZN-ve	-ve	**0.86** , 0.00-(33)*	**0.818,** 0.000-(33)*	**1,** 0.00-(33)*	**0.522**, 0.002-(33)*
FM+ve	+ve	**1,** 0.157-(2)	N/A	N/A	N/A
FM+ve	-ve	**0.36,** 0.21-(7)	N/A	**0.58**, 0.088-(7)	N/A
FM-ve	+ve	**0.82**, 0.004-(12)*	**0.429,** 0.070-(12)	N/A	**-0.091**, 0.753-(12)
FM-ve	-ve	**0.86**, 0.00(28)*	**0.819,** 0.00-(28)*	**1,** 0 .00-(28)*	0.512, 0.006-(28)*
Gx+ve	+ve	**1**, 0.157,-(2)	N/A	N/A	N/A
Gx+ve	-ve	N/A	N/A	N/A	N/A
Gx-ve	+ve	N/A	N/A	N/A	N/A
Gx-ve	-ve	**0.69**, 0.053-(7)	**0.588**, 0.088-(7)	**1,** 0.00-(7)*	**0.588,** 0.088-(7),

Key: ***ZN*** - Ziehl–Neelsen , ***FM -*** Fluorescent microscopy, ***Gx*** - Gene Xpert :
***** means ***p-value <0.05*** is considered statistically significant.
N/A –Not available

The recorded duration history of cough by the first examiner-15 patients (9.6%) and second examiner-16 (10.2%) included those patients who had coughed less than 2 weeks. This was not an inclusion criteria. However significantly good kappa values were calculated for all categories of TB and HIV test results where results were available. Gx+ve /HIV +ve provided an excellent kappa scores. Patients with history of weight loss had significant fair kappa scores with excellent scores occurring with ZN-ve /HIV +ve and FM-ve / HIV+ve groups. Moderate kappa scores occurred with ZN+ve/ HIV-ve group. Gx-ve/ HIV-ve kappa results were below 0. This was also observed for the ZN+ve/ HIV-ve and Gx+/HIV+ve results from the patients with night sweats. Moderate and fair results significantly different from zero were found for ZN+ve/HIV+ve, ZN+ve/HIV-ve , FM+ve/HIV+ve and FM+ve/HIV-ve categories kappa scores in patients with night sweats . Low grade fever scores were fair for ZNve/HIV-ve and FM-ve/ HIV-ve, Gx+ve/HIV-ve but significantly greater than the zero score. All other available results were good and significantly greater than zero score. The highest kappa occurred in the cough category (*Tables 5a, 5b).* For patients with a cough symptom category, kappa

values were significantly excellent in the ZN-ve and FM-ve irrespective of the HIV status. The data was sparse for individuals with ZN +ve and Gx+ve. Patients with weight loss had significantly excellent kappa values for those with ZN-ve and FM-ve among the HIV -ves. Kappa was fair in the ZN+ve and FM+ve who were HIV +ve as well in the few Gx results that had findings. Significant excellent results were observed in the calculated kappa results for patients who had night sweats and were ZN-ve, FM-ve, Gx-ve and HIV-ve among this group. ZN-ve and FM-ve kappa were fair but still significantly different from the zero score. HIV +ve patients with ZN-ve, and FM-ve had kappa values below zero *(Table 6).*

Table 7: *Summary Diagnostic Test Values of Having Positive TB Test Vs.all 4 Signs and Symptoms During Intra-Examiner Calibration*

	ZN	95% CI	FM	95% CI	Gx	95% CI
Sensitivity:	0.6563	0.4831 to 0.7959	0.6786	0.4934 to 0.8207	0.6970	0.5266 to 0.8262
Specificity:	0.2376	0.1652 to 0.3293	0.2727	0.1907 to 0.3738	0.5333	0.3012 to 0.7519
Positive likelihood ratio:	0.861	0.655 to 1.131	0.933	0.701 to 1.241	1.494	0.831 to 2.683
Negative likelihood ratio:	1.447	0.8 to 2.617	1.179	0.623 to 2.229	0.568	0.282 to 1.146
Diagnostic odds ratio:	0.595	0.251 to 1.408	0.792	0.315 to 1.989	2.629	0.748 to 9.239

Likelihood ratios are used for assessing the value of performing a diagnostic test. They use the sensitivity and specificity of the test to determine whether a test result usefully changes the probability that a condition (such as a disease state) exists (positive) or the reverse (negative). The diagnostic odds ratio is a measure of the effectiveness of a diagnostic test.

Table 8: Summary diagnostic test values of having positive TB test Vs. all of 4 signs and symptoms during inter- examiner calibration.

Table 8: *Summary Diagnostic Test Values of Having Positive TB Test Vs all of 4 Signs and Symptoms During Inter- Examiner Calibration*

	ZN	95% CI	FM	95% CI	Gx	95% CI
Sensitivity:	0.4444	0.1888 to 0.7333	0.5000	0.2152 to 0.7848	0.7143	0.3589 to 0.9178
Specificity:	0.4167	0.2447 to 0.6117	0.4615	0.2876 to 0.6454	0.3333	0.1206 to 0.6458
Positive likelihood ratio:	0.762	0.341 to 1.704	0.929	0.426 to 2.024	1.071	0.555 to 2.069
Negative likelihood ratio:	1.333	0.629 to 2.828	1.083	0.483 to 2.43	0.857	0.193 to 3.81
Diagnostic odds ratio:	0.571	0.122 to 2.679	0.857	0.176 to 4.186	1.250	0.146 to 10.699

There was no significant difference between the different TB tests in terms of the diagnostic values *(Table 7)*. However Gx demonstrated consistency than the other two tests except for the negative Likelihood ratio. No significant difference between the different TB tests in terms of the diagnostic values similar to values in *table 7* except the specificity values shown in table 8. However Gx demonstrated consistently higher values than FM and Zn except for the negative

Likelihood Ratio

Discussion

The results of this study show that except in the ZN+ve/HIV+ve, FM+ve/HIV+ve, Gx+ve/HIV+ve Gx+ve/HIV-ve categories, at least all four signs and symptoms of TB had concordant sensitivity results that were significantly higher than those of any combinations of at least 3, least 2 and least 1 cardinal sign and symptom being present respectively in the intra-examiner calibration findings. These results correspond to findings of a meta-analysis study [4] where the use of four cardinal TB signs and symptoms for clinical diagnosis found an overall sensitivity of 78.9% (95% confidence interval [CI] 58.3%– 90.9%).

In the study sensitivity increased to 90.1% (95% CI 76.3%–96.2%) among participants selected from clinical settings and to 88.0% (95% CI 76.1%–94.4%) among those who were not previously screened for TB. Early tuberculosis is likely associated with greater diagnostic uncertainty and risk of differential diagnosis misclassification. Participants in this study were people with symptoms or signs suggestive of TB not likely to have been screened previously visiting the hospital for relief of signs and symptoms. This indicates that these clinical features initiate the health seeking behavior bringing them to the health care facilities where TB diagnosis is made. The sensitivity occurred within the 95% CI found in the meta-analysis study showing both the external validity and reliability of the examiners in this study [4].

Gx+ve/HIV+ve and Gx+ve/HIV-ve sensitivity results in this study also show that these clinical signs and symptoms reliably diagnose TB. The highest priority for TB control is the identification and cure of infectious cases, i.e. patients with sputum smear-positive PTB. Clinical findings (all 4 signs and symptoms) have low specificity due to their inability to have accurate differential diagnosis results. Therefore all patients (regardless of HIV status) with clinical features suggestive of PTB must submit sputum for diagnostic sputum smear microscopy. Findings in a paper in this issue show that Gx has good sensitivity in diagnosing TB when culture is used as a gold standard. This paper also showed low sensitivity for ZN and FM with respect to culture indicating that even in HIV positive patients validity of the two tests is questionable [21]. The inability to differentiate between clinically diagnosed TB cases has major implications in TB control programmes [7,8,20].

Finding from these two studies differ from a calibration study between clinician diagnosis of TB from Post mortems (PM) and verbal autopsy (VA) from

the community. Post mortems (PM) were done at a public referral hospital. Culture and GeneXpert tests complemented the PM results. They found sensitivity of 27% [95% CI 10%, 57%] that had concordant results with the clinical VA interpretation [19].

The sensitivity of having at least 3 signs and symptoms in the meta-analysis study varied from 74.0% (95%C=51.7–88.3) to 63.1% (95% CI=39.3–81.9). These were however significantly higher than those found in the intra-examiner calibration in this study except in ZN+ve/HIV+ve, FM+ve/HIV+ve and Gxve/HIV-ve categories. The meta-analysis study compared a referent group with CD4 count of >= 200 and those with <200 the odds ratio for sensitivity was 6.38(95% CI=2.87–14.17). These findings are reflected in the two HIV+ve categories in this study which may have included immunosuppressed patients [4].

Although weight loss and fever are more common in HIV-positive PTB patients than in those who are HIV-negative this was not consistent in this study [8]. The inter-examiner sensitivity(table 3), intra and inter-examiner Kappa results from this study(tables 5 and 6) indicate the need for calibration of clinicians through operational research studies to capture the effect of clinician variability on clinical study findings and clinical management of TB cases. The variations between clinicians using a similar clinical form may reflect qualitative differences in interpreting these findings as present or absent. These variations are also reflected in the variations between the different TB diagnostics since significant difference from zero kappa score is not observed in all the categories. There are also variations depending on the cardinal TB sign and symptom recorded.

Study Limitations

The results were easily influenced by the sample size. The inter-examiner 57 patients result showing greater variability than the intra-examiner 155 patients variations. There is need to continue the study in order to achieve adequate sample size to be more definitive about these observed differences.

Conclusion

Excellent kappa can be achieved in low resource settings by clinicians using all four cardinal signs and symptoms of TB. There was possibility of using the clinical diagnosis where laboratory diagnosis was not possible but the specificity was low. Good clinical differential diagnosis of other conditions will improve the specificity.

Acknowledgement

We wish to thank, study sites personnel from Malindi, Lamu, Busia, Kitale, Nyahururu, Wajir, Machakos, Narok and Kisii county hospitals for specimens and data collection, persons with symptoms or signs suggestive of TB who provided informed consent and participated in the study.
We thank;

1. Center for Respiratory Diseases Research (CRDR) TB laboratory.
2. The Tuberculosis Central Reference Laboratory staff for performing the routine Mycobacteriology testing.
3. Director Kenya Medical Research Institute (KEMRI) for the support and permission to perform the study

References

1. **Reld M.J. Shah N. S**. Approaches to tuberculosis screening and diagnosis in people with HIV in resource limited settings. *Lancet Infect Dis* 2009;9:173-84.

2. **World Health Organization.** Improving the diagnosis and treatment of smear - negative pulmonary and extra-pulmonary tuberculosis among adults and adolescents. recommendations for HIV- prevalent and resource-constrained settings. WHO/HTM/2007.379 & WHO/HIV/2007.1.Geneva WHO 2007. http://whqlibdoc.who.int/hq/2007/WHO_HTM_TB_2007.379_Eng.pdf.Accessed 19 July 2011

3. **Margaret A. Fisch l, M.D. George L. Daikos, M.D. Raj B. Uttamchandani M.D. Rita B., M.D. Poblete M,D., Jose N. Moreno M.D., Ricardo R. Reyes M.D. Ahmad M. Boota M.D. Lisa M., Thompson B.S. B.M.E. Timothy J. Clearly PhD. Sandra A. Oldham M.D. Mario J. Sandra M.D**. and **Shenghan Lai, PhD.** Clinical Presentation and outcome of patients with HIV Infection and Tuberculosis caused by Multiple-Drug-resistant *Bacilli*. *Ann Intern 1992*;117(**3**):184-190. dol:10.7326/0003-4819-1 17-3-184

4. **Getahun H.** Meta-analysis to inform the development of a standardized approach for TB screening in HIV -Infected patients. In:Program and abstracts of the 40th Union World Conference on Lung Health Cancun.

5. **American Thoracic Society** and **Centres for Disease control and Prevention**. Diagnostic standards and classification of tuberculosis in adults and children. *Am J Respir Crit Care Med*. 2000;161 (4pt);1376-1395.

6. TB elimination: the difference between latent TB and active TB disease. **Centres for Disease Control and Prevention** Website. htt://cdc.qov/pubs/tb fact sheets/LTBI and Active TB. pdf. Updated October 7. 2008. *Accessed January* 28, 2009.

7. **World Health Organization.** Treatment of tuberculosis. *Guidelines for national programmes 3rd ed.* WHO/CDC/TB 2003.313 Geneva.

8. **WHO**,2003.http://www.who.int/tb/publications/TB_treatmentguidelines/en/index.html. Accessed 22 July2011.

9. **Mexico:** International Union Against Tuberculosis and Lung Disease, 2009. *Symposium no. 2.*

10. **Paton N.I, Chua Y.K, Earnest A, Chee C. B.** Randomized controlled trial of nutritional supplementation in patients with newly diagnosed tuberculosis and wasting. *Am J ClinNutr. 2004*; 80: 450–465.

11. **Chien J.W., Johnson J. L.** Paradoxical reactions in HIV and pulmonary TB. *Chest.1998*; 114 :933 -6 .

12. **French MA.** HIV/AIDS: immunereconstitution inammatory syndrome: areappraisal. *Clin Infect Dis. 2009*; 48:101 -7 .

13. **Meintjes G, Lawn S.D, Scano F.,Maartens G., French M.A., Worodria W., et al**. International Network for the Study of HIVassociated IRIS. Tuberculosis-associated immune-reconstitution inammatory syndrome:case denitions for use in resource-limited settings. *Lancet Infect Dis. 200*8; 8:516 -2 3.

14. **Breen R.A, Smith C.J, Cropley I, Johnson M.A, Lipman M.C.** Does immunereconstitution syndrome. Promote activetuberculosis in patients receiving highly active antiretroviral therapy? *AIDS . 2005*; 19:1201 – 06.

15. **Tarantola G.J., Becker I.M.,** and **H. Gremillion** The reproducibility of centric relation : a clinical approach *JADA , 1997* ; 128 (9): 1245 - 1251;.*2004*;80:450–465.

16. **Altman G.,** "Practical statistics for medical research," in Inter-Rater Agreement, pp. 403 – 404, 1991.

17. **The World Health Organization.** Guidelines for intensified tuberculosis case finding and isoniazid preventive therapy for people living with HIV in resource-constrained settings. Geneva: WHO 2010. http://whqlibdoc.who.int/publications/2011/9789241500708_eng.pdf. Accessed 7 December 2010.

18. **Donner A., Rotondi M. A.** Sample Size Requirements for Interval Estimation of the Kappa Statistic for Inter-observer Agreement-studies with a Binary Outcome and Multiple Raters.2010 ; 6:31.

19. **Gacheri S. Kipruto H. Amukoye E. et al** Performance of Clinicians in Identifying Tuberculosis as cause of Death using Verbal Autopsy Questionnaires in Siaya County, Kenya. *Afr. J. Health Sci. 2014* ; 27(**3**):232 - 8.

2O. **Kenya TB treatment guidelines**. http://www.nltp.co.ke/docs/TB_Treatment_GUIDELINES_2013.pdf .

21. **Githui W.A, Orina F., Mwangi M., Kiptoo M., Ogaro T , Wanzala P., Kariuki J.N., Sang W. K.,** Performance of Ziehl - Neelsen Microscopy, Light Emitting Diode–F M andXpert MTB/RIF in the diagnosis of Tuberculosis in TB suspects from *EAPHLN Project sites in Kenya. 2014*; 27(**4**):433-449.