The Research of Medical Science Review
ISSN: 3007-1208 & 3007-1216 Volume 3, Issue 3, 2025

VALIDITY ANALYSIS OF COMPUTER-AIDED DIAGNOSIS (CAD4TB) VS. GENEXPERT FOR TUBERCULOSIS DETECTION IN VERBAL SCREENED PRESUMPTIVE PULMONARY TB PATIENTS IN QUETTA, BALUCHISTAN

Zafar Iqbal Khosti[1], Dr Samia Ahmed Khan*[2], Shagufta Naseer[3], Sher Afghan[4], Syeda Farhat Bukhari[5], Mohsin Jamil Siddiqui[6]

[1]*Health Department Govt of Baluchistan, Pakistan*
[*2]*Sindh Healthcare Commission, Pakistan,*
[3]*Humanities & Social Sciences Dept, Bahira University Karachi Campus, Pakistan,*
[4]*TB Control Program Baluchistan, Pakistan,*
[5]*Hamdard University Dental Hospital Karachi,*
[6]*Sindh Healthcare Commission,*

[1]zaffar_khosti@ymail.com, [*2]khansamia804@gmail.com, [3]shaguftanaseer.bukc@bahria.edu.pk,,
[4]drsherafganraisani@gmail.com, [5]f_bukhari@hotmail.com, [6]siddiquicuba@gmail.com,

DOI: https://doi.org/10.5281/zenodo.14965802

Keywords
Tuberculosis, Computer Aided Diagnosis, GeneXpert

Article History
Received on 25 January 2025
Accepted on 25 February 2025
Published on 04 March 2025

Corresponding Author: *
Dr Samia Ahmed Khan

Abstract

Background: *Choosing diagnostic tools for serious diseases like tuberculosis requires careful consideration of parameters like sensitivity, specificity, and predictive values. Currently, both Computer-Aided Diagnosis for TB (CAD4TB) and GeneXpert are used in tandem for tuberculosis detection, but their relative effectiveness has not been conclusively established in the literature.*

Objective: *This study aims to evaluate the validity of CAD4TB by comparing it against GeneXpert, considered the gold standard.*

Methods: *A cross-sectional validation analysis was conducted using secondary data obtained from the Provincial TB Control Program, sourced from public sector healthcare institutes in Quetta, where both CAD4TB and GeneXpert facilities were available. Validity was calculated, yielding sensitivity (89.4%), specificity (10.2%), positive predictive value (14.7%), and negative predictive value (84.9%). Significant associations were further evaluated for trend association with the Moses Extreme Reactions test. Sensitivity and specificity were established, and Receiver Operating Curve analysis confirmed validity. Statistical significance was set at $p < 0.05$, and SPSS v 25.0 was used for data processing and analysis.*

Results: *Out of 573 presumptive patients, 514 were included in the final analysis (response rate: 89.7%). The majority were over 54 years old (260, 50.5%), and the cohort was predominantly male (282, 54.8%). The mean CAD4TB score was 80.15 ± 12.93, while GeneXpert detected 84.8% of cases. Significant associations were found between CAD4TB and Xpert scores, with moderately strong relationship interpreted. Moses Extreme Reaction test confirmed a significant relationship favoring GeneXpert. CAD4TB reported an area under*

the curve value of 0.497, indicating poor discrimination. This reinforced the preference for GeneXpert over CAD4TB.

Conclusion: *This study demonstrates that GeneXpert provides a more sensitive and specific categorization of patients compared to CAD4TB.*

INTRODUCTION

Tuberculosis (TB) remains a significant global health issue, causing substantial mortality and morbidity (1). As of 2019, the World Health Organization (WHO) reported nearly 10 million TB cases, surpassing HIV/AIDS as the primary cause of death from a single infectious agent (2-3). The delay and expense associated with the gold standard sputum culture test present obstacles to timely diagnosis, emphasizing the importance of early detection due to the highly contagious nature of the disease (4-6).

TB diagnosis typically involves methods like sputum smear microscopy, Mantoux test, and, for confirmation, GeneXpert or Culture and Sensitivity tests(7-8). While smear microscopy has been instrumental, its accuracy falters in HIV-positive patients, children, and those with low bacterial load(9-10). GeneXpert revolutionized TB diagnosis, offering rapid results and detecting rifampicin resistance. Its effectiveness and efficiency are evident when compared to traditional culture methods. CAD4TB, operating on digital chest radiographs, automatically assesses TB presence based on a scoring system (11). The latest version employs deep learning technology for faster results, providing invaluable support, particularly in resource-limited settings. This advancement has significant potential in expediting TB diagnosis.

Despite progress, challenges persist in CAD systems, including data collection, processing, and assessment (12). Standardized assessment measures, essential for FDA certification, are lacking. Additionally, acceptance and adoption by healthcare providers remain a concern, often due to insufficient training. CAD4TB, in conjunction with digital chest radiographs, plays a crucial role in TB diagnosis, particularly in areas with high disease prevalence. This technology enhances case finding and aids in the early identification of presumptive TB cases. GeneXpert, a cartridge-based nucleic acid amplification test, swiftly detects Mycobacterium Tuberculosis and identifies rifampicin resistance. Its effectiveness and efficiency have made it a cornerstone of TB diagnostics, offering advantages over traditional methods. GeneXpert, with its molecular-level detection capabilities, offers high sensitivity and specificity, earning WHO's endorsement (13). However, its cost and complexity in resource-limited settings present challenges. CAD4TB, utilizing digital chest radiographs, complements GeneXpert by serving as an effective pre-screening tool, especially in areas lacking radiology expertise.

Both CAD4TB and Gene Xpert are commonly employed in clinical practice. However, considering the nature of the disease and its pathophysiology, a more sensitive and cost-effective method is warranted. This study aims to validate CAD4TB by comparing it to GeneXpert, determining the superior testing method for diagnosing TB in high-burden countries like Pakistan. While several methods exist to diagnose Tuberculosis, such as CXR, sputum smear microscopy, culture tests, Mantoux test, and GeneXpert, they all require time and expertise. The invention of CAD4TB offers a quicker, expert-independent method to automatically detect TB in chest X-rays. This study aims to validate CAD4TB's potential to replace GeneXpert for diagnosing Tuberculosis patients. Despite longstanding efforts to eradicate.

Materials and Methods:

This study was conducted in Public Sector Healthcare Institutes of Quetta, where both CAD4TB and Gene Xpert facilities are available. The research design employed for this investigation was a validation analysis using a cross-sectional study approach. The sample size was determined to be n=514, and a non-probability purposive sampling technique was utilized to identify cases that were verbally suspected of having TB and subsequently underwent CAD4TB and Gene Xpert testing. The CAD4TB and Gene Xpert data of presumptive patients were gathered in collaboration with the Indus hospital network and the Provincial TB

Control Program in Baluchistan. The study encompassed participants aged 14 years and older who were verbally suspected of having TB and had undergone CAD4TB, as well as provided a sputum sample for Gene Xpert testing. Excluded from the study were individuals who were already diagnosed with TB, known cases of TB, those who failed to provide a suitable sputum sample for Gene Xpert, or those who provided a sample that was mixed with gastric contents or saliva. Additionally, pregnant women were not included in the study. The scores obtained from the computer-aided diagnosis were then compared with the results of the Gene Xpert test, which served as the "reference standard" or "gold standard". Sensitivity was calculated, followed by specificity, and subsequently, the Positive Predictive Value (PPV) and Negative Predictive Value (NPV) of CAD4TB were determined in comparison to Gene Xpert. For data analysis, version 23 of SPSS was utilized. Ethical approval was obtained from the Program Manager of the Provincial TB Control Program, Baluchistan, prior to conducting the research. The study was carried out after receiving clearance from the Institutional Review Board of AFPGMI.

Results:

Elaboration of study characteristics: A total of 514 patients were included in the final analysis after confirming the accuracy of the data before coding and integration into SPSS v 23 (Table 4.1). Among these, the majority (260, 50.5%) were aged 54 years or older, with a mean age of 52.79 ± 16.43 years. The cohort was predominantly male (282, 54.8%), and the majority sought care at Sandeman Provincial Hospital, Quetta for routine treatment. The mean CAD4TB score was 80.15 ± 12.93, while GeneXpert (GXP) successfully detected TB in 84.8% of cases (Table 4.1).

Validity Analysis of CAD4TB

Figure 4.1 and 4.2 present the graphical representation of CAD4TB and GeneXpert results. The validity analysis utilized a 2x2 comparison model, and the results are outlined in Table 4.2. Based on the observed values, we applied established formulas to calculate sensitivity 89.4%, specificity 10.2%, as well as positive 14.7% and negative 89.4% predictive values.

Comparison of CAD4TB and GXP Scores

A Chi-Square test was employed for an initial comparison and to establish a generalized relationship between CAD4TB and GeneXpert. The analysis revealed a significant association, with a Phi value (φc) of 0.275, indicating a "moderately strong relationship" and suggesting the potential for further nuanced analysis.

Assessment of Relationship Trend

To evaluate the trend of association, the Moses Extreme Reaction test was applied. A significant relationship (p=0.047) was observed in the control group span (CAD4TB and GXP detected), favouring GXP over CAD4TB (Table 4.4 and Table 4.5). The trimmed control group span (CAD4TB and GXP none detected) showed no significant difference, further supporting the superiority of GXP.

Confirmatory Analysis by Receiver Operating Curve (ROC)

The ROC analysis was employed to assess the diagnostic accuracy. CAD4TB reported an Area Under the Curve (AUC) of 0.497, indicating poor performance compared to GeneXpert. This confirmed the advantage of GXP in terms of sensitivity and specificity (Figure 4.3).

Confirmatory Analysis of Positive and Negative Predictive Values

The positive predictive value was 15% and the negative predictive value was 85%. Cross tabulation reconfirmed these values (Table 4.7). GXP detected 15.2% within the CAD4TB score, while 84.8% of cases were undetected. This reiterated the earlier findings, supporting the preference for GXP over CAD4TB.

Discussion

In clinical practice, the choice of diagnostic tool is influenced by factors like availability and affordability (14). While CAD4TB is a cost-effective option for TB diagnosis, GeneXpert's high sensitivity provides a notable advantage. The selection process is complex and multifaceted. Notably, a comprehensive comparative analysis between GeneXpert and CAD4TB remains scarce in existing literature. This study aimed to address this gap by comparing the outcomes of CAD4TB and GeneXpert and determining the superior diagnostic choice. The initial assessment revealed a significant association between CAD4TB and GeneXpert, with a moderately strong relationship, as supported by Cramer's V value. This indicates a noteworthy relationship between the two variables (15-16) . The observed effect size further underscores this relationship, aligning with previous findings(17).

Following effect size assessment, we delved into trend analysis.

The Extreme Reactions method demonstrated a significant relationship favoring GeneXpert over CAD4TB in the observed control group span. Conversely, the trimmed control group span did not show a significant difference, reinforcing GeneXpert's superiority. This aligns with the findings of Rahman et al (18), emphasizing GeneXpert's higher specificity. However, despite considerable efforts, Tuberculosis remains a major global health concern. Affordable, practical, and efficient diagnostic solutions remain a significant challenge. Comparing our sensitivity results (89.5%) to the World Health Organization's estimate for GeneXpert (92.5%), discrepancies may arise from sample size and study contexts. Using GeneXpert as the reference standard in our study favoured its performance. Future research should explore machine learning systems for digital CXR in a diverse population.

Advancements in digital technology and molecular methods have improved TB diagnosis. This study demonstrates GeneXpert's superiority in sensitivity and specificity over CAD4TB. GeneXpert is recommended, especially in high-burden areas, while CAD4TB can serve as a valuable pre-screening tool. Larger studies should validate these findings.

Given the limited scope of our study, a broader investigation across various healthcare facilities is advised. Additionally, raising the CAD4TB cut-off value to 90 for comparison with GeneXpert is proposed. Further evaluation of diagnostic measures, particularly in cases of GeneXpert-negative TB with positive cultures, is essential for accurate labelling. To generalize our findings, a province or nationwide study is recommended. Additionally, evaluating multiple diagnostic measures for specific evidence is crucial. Cost-effectiveness should be a key consideration, and a cost utility analysis for both CAD4TB and GeneXpert is suggested.

Conclusion

This study shows that GeneXpert offers a more accurate and specific classification of patients than CAD4TB. Consequently, GeneXpert should be the preferred method, particularly in regions with high disease prevalence

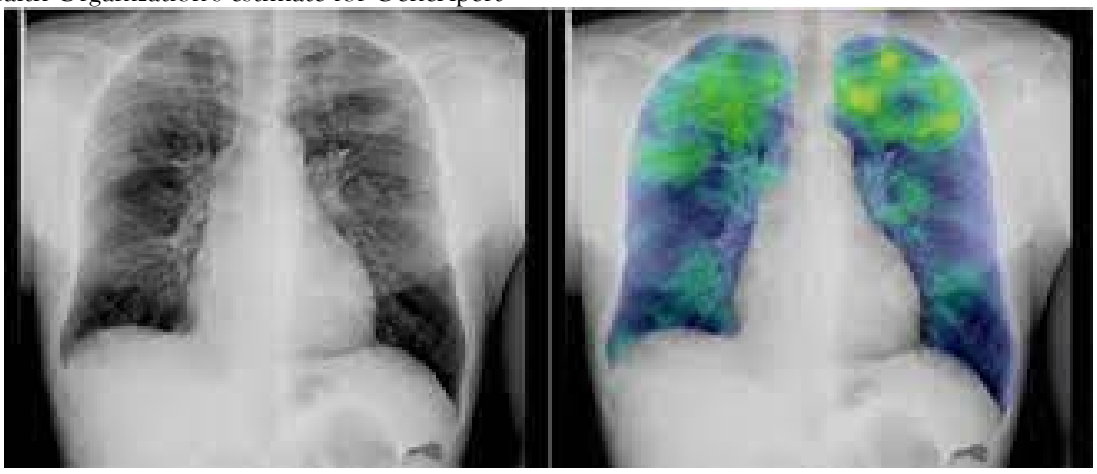

Figure 01: Normal chest X-ray vs CAD4TB CXR

Figure 02: GeneXpert Machine of 16 slots in BSL lab

Table 4.1 : Demographic characteristic of the study respondents

Characteristics	Frequency	Percentage (%)
Age group (52.79±16.43)		
15-24	28	5.4
25-34	40	7.7
35-44	79	15.3
45-54	107	20.8
> 54	260	50.5
Gender		
Male	282	54.8
Female	232	45.2
Sample Collection Sites		
Sandeman Provincial Hospital, Quetta	448	87.3
Community of the Quetta city	40	7.7
Rural Health Camp, Quetta City	26	5.0

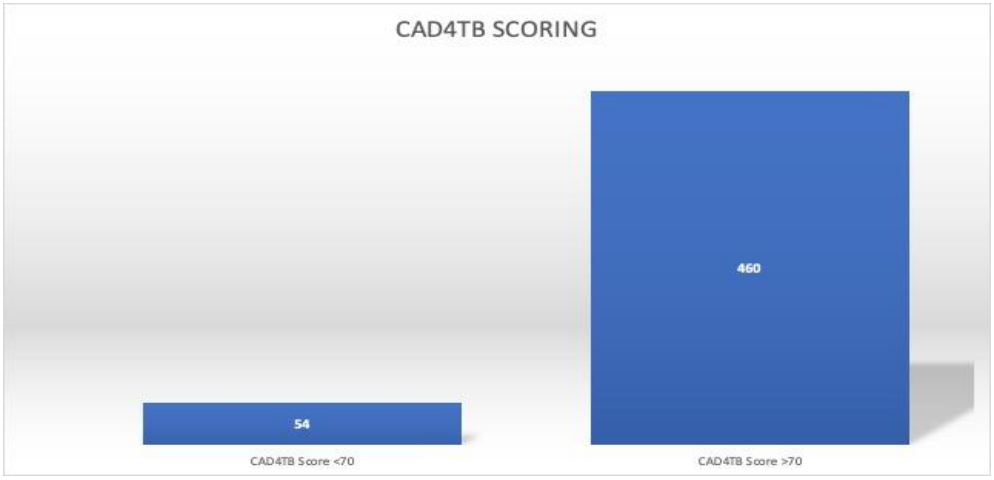

Figure 4.1: CAD4TB Scoring of 514 presumptive TB Patients

Figure 4.2: Results of GeneXpert of 514 TB presumptive patients

Table 4.2 : 2 X 2 Comparison Table

	GXP Detected	GXP Not Detected	
CAD4TB Score >70	A = 68	B = 393	**A+B = 461**
CAD4TB Score <70	C = 08	D = 45	**C+D = 53**
	A+C=76	**B+D=438**	**A+B+C+D = 514**

Table 4.3 : Correlation Values and interpretation.

Phi and Cramer'V Values	Values Lying in the range means
(0.00)	"Having No Relationship"
(0.00 - 0.15)	"Results are Not generally useful"
(0.10 - 0.20)	"Weak relationship"
(0.20 - 0.25)	"Moderate relationship"
(0.25 - 0.30)	"Moderately Strong relationship"
(0.30 - 0.35)	"Strong relationship"
(0.35 - 0.40)	"Very Strong relationship"
(0.40 - 0.45)	"Worrisomely Strong relationship"
(0.45 - 0.99)	"Redundant relationship"
(1.00)	"Perfect Relationship"

Table 4.4 : Assessment of Trend of Relationship

Frequencies		
	GXP results	N
CAD4TB Score	Detected (Observed Control group)	78
	Not detected (Trimmed Control Group)	436
	Total	514

Table 4.5 : Test Statistics (trend analysis)

		CAD4TB Score
Observed span		484
	Sig., (1-tailed)	***0.047***
Trimmed Span		458
	Sig., (1-tailed)	0.376
Outliers Trimmed from each End		3

The Moses Extreme Reaction test with GXP rated as grouping variable

Table 4.6 : Parameters for Area under the Curve Assessment

Test Result Variable: CAD4TB SCORE				
Area Under the Curve	Standerd. Error[a]	Asymptotic Significance.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
0.497	0.036	0.937	0.427	0.567

[a]***Under the nonparametric assumption***
[b]***Null hypothesis: true area = 0.5***

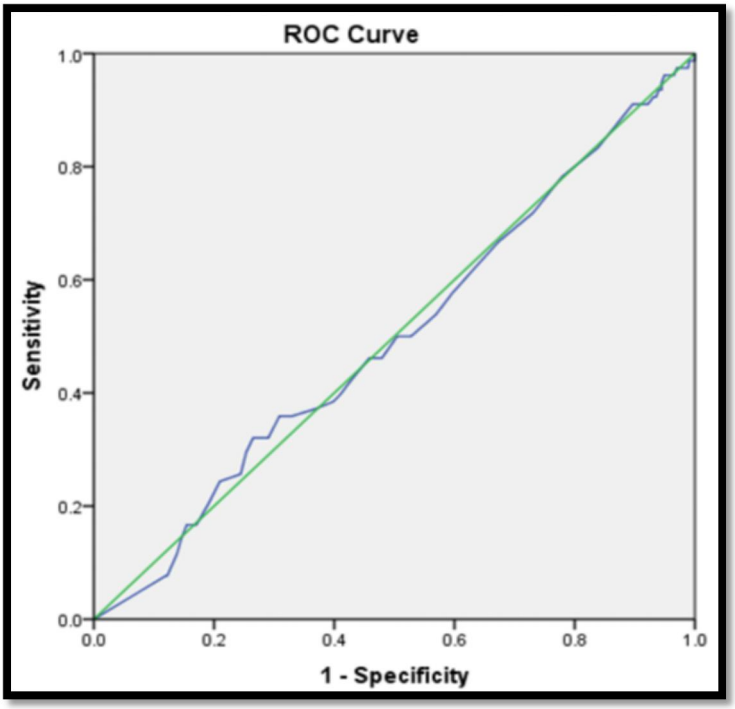

Figure 4.3: The ROC Curve: Diagonal segments are produced by ties

Table 4.7 : CAD4TB Score X GXP Results (Cross Tabulation)

			GXP results		Total
			detected	not detected	
CAD4TB score	< 70	Count	7	45	52
		% within CAD4TB score	13.5%	86.5%	100.0%
		% within GXP results	9.0%	10.4%	10.2%
		% of Total	1.4%	8.8%	***10.2%***

	> 70	Count	71	389	460
		% within CAD4TB score	15.4%	84.6%	100.0%
		% within GXP results	91.0%	89.6%	89.8%
		% of Total	13.9%	76.0%	***89.8%***
Total		Count	78	436	514
		% within CAD4TB score	15.2%	84.8%	100.0%
		% within GXP results	100.0%	100.0%	100.0%
		% of Total	***15.2%***	***84.8%***	***100.0%***

REFERENCE:

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish journal of emergency medicine*, *18*(3), 91-93.

Bujang, M. A., & Adnan, T. H. (2016). Requirements for minimum sample size for sensitivity and specificity analysis. *Journal of clinical and diagnostic research*, *10*(10), YE01.

Chudi, I. P. (2010). Healthcare problems in developing countries. *Medical Practice and Reviews*, *1*(1), 9-11.

Drezner, Z., Turel, O., & Zerom, D. (2010). A modified Kolmogorov–Smirnov test for normality. *Communications in Statistics–Simulation and Computation*, *39*(4), 693-704.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. sage.

Graham, S., Das Gupta, K., Hidvegi R, J., Hanson, R., Kosiuk, J., Al Zahrani, K., & Menzies, D. (2002). Chest radiograph abnormalities associated with tuberculosis: reproducibility and yield of active cases. *The International journal of tuberculosis and lung disease*, *6*(2), 137-142.

Hajian-Tilaki, K. (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian journal of internal medicine*, *4*(2), 627.

Henostroza, G., Harris, J. B., Kancheya, N., Nhandu, V., Musopole, R., Krüüner, A., Chileshe, C., Dunn, I. J., & Reid, S. E. (2016). Chest radiograph reading and recording system: evaluation in frontline clinicians in Zambia. *BMC infectious diseases*, *16*(1), 1-8.

Iademarco, M. F., O'Grady, J., & Lönnroth, K. (2012). Chest radiography for tuberculosis screening is back on the agenda. *International Journal of Tuberculosis and Lung Disease*, *16*(11), 1421-1422.

Jaeger, S., Karargyris, A., Candemir, S., Siegelman, J., Folio, L., Antani, S., & Thoma, G. (2013). Automatic screening for tuberculosis in chest radiographs: a survey. *Quantitative imaging in medicine and surgery*, *3*(2), 89-99.

Kumar, R., & Indrayan, A. (2011). Receiver operating characteristic (ROC) curve for medical researchers. *Indian pediatrics*, *48*(4), 277-287.

Maduskar, P., Muyoyeta, M., Ayles, H., Hogeweg, L., Peters-Bax, L., & Van Ginneken, B. (2013). Detection of tuberculosis using digital chest radiography: automated reading vs. interpretation by clinical officers. *The International journal of tuberculosis and lung disease*, *17*(12), 1613-1620.

Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology*, *5*(9), 1315-1316.

Murphy, K., Habib, S. S., Zaidi, S. M. A., Khowaja, S., Khan, A., Melendez, J., Scholten, E. T., Amad, F., Schalekamp, S., & Verhagen, M. (2020). Computer aided detection of tuberculosis on chest radiographs: An evaluation of the CAD4TB v6 system. *Scientific reports*, *10*(1), 1-11.

Muyoyeta, M., Maduskar, P., Moyo, M., Kasese, N., Milimo, D., Spooner, R., Kapata, N., Hogeweg, L., van Ginneken, B., & Ayles, H. (2014). The sensitivity and specificity of using a computer aided diagnosis program for automatically scoring chest X-rays of presumptive TB patients compared with Xpert MTB/RIF in Lusaka Zambia. *PloS one*, *9*(4), e93757.

Ntinginya, E., Squire, S., Millington, K., Mtafya, B., Saathoff, E., Heinrich, N., Rojas-Ponce, G., Kowuor, D., Maboko, L., & Reither, K. (2012). Performance of the Xpert®

MTB/RIF assay in an active case-finding strategy: a pilot study from Tanzania. *The International journal of tuberculosis and lung disease*, *16*(11), 1468-1470.

O'Grady, J., Bates, M., Chilukutu, L., Mzyece, J., Cheelo, B., Chilufya, M., Mukonda, L., Mumba, M., Tembo, J., & Chomba, M. (2012). Evaluation of the Xpert MTB/RIF assay at a tertiary care referral hospital in a setting where tuberculosis and HIV infection are highly endemic. *Clinical infectious diseases*, *55*(9), 1171-1178.

Pai, M., & Palamountain, K. (2012). New tuberculosis technologies: challenges for retooling and scale-up [State of the art series. New tools. Number 4 in the series]. *The International journal of tuberculosis and lung disease*, *16*(10), 1281-1290.

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge.

Rahman, M. T., Codlin, A. J., Rahman, M. M., Nahar, A., Reja, M., Islam, T., Qin, Z. Z., Khan, M. A. S., Banu, S., & Creswell, J. (2017). An evaluation of automated chest radiography reading software for tuberculosis screening among public-and private-sector patients. *European Respiratory Journal*, *49*(5), 1602159.

Sachs, L. (2012). *Applied statistics: a handbook of techniques*. Springer Science & Business Media.

Story, A., Aldridge, R. W., Abubakar, I., Stagg, H., Lipman, M., Watson, J., & Hayward, A. (2012). Active case finding for pulmonary tuberculosis using mobile digital chest radiography: an observational study. *The International journal of tuberculosis and lung disease*, *16*(11), 1461-1467.

Theron, G., Pooran, A., Peter, J., van Zyl-Smit, R., Mishra, H. K., Meldau, R., Calligaro, G., Allwood, B., Sharma, S. K., & Dawson, R. (2012). Do adjunct tuberculosis tests, when combined with Xpert MTB/RIF, improve accuracy and the cost of diagnosis in a resource-poor setting? *European Respiratory Journal*, *40*(1), 161-168.

Van't Hoog, A., Meme, H., Van Deutekom, H., Mithika, A., Olunga, C., Onyino, F., & Borgdorff, M. (2011). High sensitivity of chest radiograph reading by clinical officers in a tuberculosis prevalence survey. *The International journal of tuberculosis and lung disease*, *15*(10), 1308-1314.

World Health Organization. (2011). *Rapid implementation of the Xpert MTB/RIF diagnostic test: technical and operational'How-to'; practical considerations* (9241501561).

World Health Organization. (2014). *Xpert MTB/RIF implementation manual: technical and operational 'how-to'; practical considerations*. World Health Organization,. https://apps.who.int/iris/bitstream/handle/10665/112469/9789244506707_rus.pdf

Yin, L., Wu, Q., & Hong, D. (2016). Statistical methods and software package for medical trend analysis in health rate review process. *Journal of Health and Medical Information*, *7*(219), 2.

Zhu, W., Zeng, N., & Wang, N. (2010). Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. *NESUG proceedings: health care and life sciences, Baltimore, Maryland*, *19*, 67.