International Multispecialty Journal of Health (IMJ Health)
ISSN: [2395-6291]
[Vol-12, Issue-8, August- 2026]
Website: www.imjhealth.org
Journal DOI: 10.25125/medical-journal

A Multidisciplinary-Validated Radiographic Severity Score for Immune Stratification in HIV-Associated Tuberculosis: Development and Internal Validation

Uchechukwu Agboje[1*]; Ordu Collins Ahamefule[2]; Obazee Emmanuel[3]; Muo Somtochukwu Marvis[4]

[*1]Department of Medicine, Shrewsbury and Telford NHS Trust, Telford, United Kingdom
[2]Department of Medicine, Delta State University Teaching Hospital, Oghara, Nigeria
[3]Department of Internal Medicine, University of Port Harcourt Teaching Hospital, Port Harcourt, Nigeria
[4]Department of Pharmacy, Enugu State University of Science and Technology, Enugu, Nigeria
*Corresponding Author - ORCID: 0000-0003-3841-9184

Received:-24 July 2026/ Revised:- 01 August 2026/ Accepted: 08 August 2026/ Published: 15-08-2026

Abstract—

Background: *Human immunodeficiency virus (HIV)-associated tuberculosis (TB) remains the leading cause of death among people living with HIV in sub-Saharan Africa. Where CD4 testing is unavailable, a chest X-ray must substitute for flow cytometry. We developed and internally validated a Radiographic Severity Score (RSS) to identify HIV-positive patients with advanced immunosuppression using plain-film radiography.*

Methods: *Cross-sectional diagnostic accuracy study of 242 HIV-positive adults with newly diagnosed pulmonary TB at two sites in southern Nigeria (2017–2018). Two radiologists independently scored films using the RSS (12 zones, weighted for severity). Discriminative accuracy, calibration, and inter-reader reliability were assessed.*

Results: *Of 242 enrolled, 228 had complete data. Mean RSS was 2.2±1.5 (CD4 <100 cells/µL), 3.4±1.7 (CD4 100–199 cells/µL), 4.8±1.9 (CD4 200–349 cells/µL), and 6.1±1.8 (CD4 ≥350 cells/µL). The stepwise gradient was significant (Spearman rho =+0.72, p <0.001). At threshold RSS ≤4: AUC=0.84 (95% CI 0.76–0.92), sensitivity 68.3%, specificity 80.0%, PPV 79.5%, NPV 69.4%. Inter-reader agreement was almost perfect (Cohen kappa =0.81, ICC=0.92).*

Conclusion: *The RSS is a simple bedside tool for identifying HIV-positive patients with CD4 <200 cells/µL in settings without same-day CD4 testing. In this Nigerian cohort, the RSS demonstrated strong discriminative accuracy (AUC 0.84) and almost perfect inter-reader reliability (kappa 0.81). However, external validation in independent cohorts is required before the RSS can be recommended for widespread implementation in TB-HIV integrated programmes*

Keywords— Tuberculosis, Pulmonary; HIV Infections; CD4 Lymphocyte Count; Radiography, Thoracic; Sensitivity and Specificity; Nigeria.

I. INTRODUCTION

HIV-associated tuberculosis remains the leading cause of death among people living with HIV in sub-Saharan Africa [1]. Global epidemiologic trends indicate persistent heterogeneity in TB-HIV burden across high-income and intermediate-burden settings, underscoring the need for context-specific triage tools. Despite World Health Organization (WHO) treat-all recommendations and universal CD4 enumeration at antiretroviral therapy (ART) initiation, an estimated 40% of African clinics lack same-day access to flow cytometry [2]. In these settings, CD4 <200 cells/µL remains the threshold for intensive clinical monitoring and rapid ART escalation—decisions that cannot wait for laboratory turnaround. When the laboratory fails,

clinicians must rely on clinical judgment and the chest radiograph alone. Yet the radiograph is often underused for immune stratification—read for disease burden rather than immunologic status.

Existing radiographic scoring systems, such as the Chest Radiograph Reading and Recording System (CRRS), quantify overall disease burden for epidemiological surveillance but do not target immune stratification [3,4]. A patient with advanced immunosuppression may have minimal radiographic findings yet require the most urgent clinical attention—a paradox that current tools fail to capture. The biological reality is that immunosuppressed patients do not have 'fewer' radiographic findings; they have different patterns reflecting failed granulomatous architecture [5]. In Nigerian cohorts, atypical radiographic patterns are significantly more common among HIV-infected patients with low CD4 counts, and chest radiograph findings correlate with immunologic status [6,7].

We developed the Radiographic Severity Score (RSS) in collaboration with radiologists and clinical colleagues—a structured plain-film grading system that uses the chest X-ray already obtained in routine care to identify patients with CD4 <200 cells/μL. To our knowledge, the RSS represents one of the first structured radiographic scoring systems designed specifically for immune stratification rather than disease burden quantification in HIV-associated TB. This distinction—leveraging the inverse relationship between radiographic extent and immune status for bedside triage—differs from existing tools that quantify overall disease severity. This study reports the development and internal validation of the RSS in a Nigerian cohort, with the specific objectives of: (i) determining the correlation between RSS and CD4 count; (ii) establishing the diagnostic accuracy of RSS ≤4 for identifying advanced immunosuppression; and (iii) assessing inter-reader reliability among radiologists with variable experience.

II. METHODOLOGY

2.1 Study Design and Setting

This was a cross-sectional diagnostic accuracy study conducted between October 2017 and September 2018 at two tertiary health facilities in southern Nigeria: a 450-bed federal teaching hospital (Hospital A), Oghara, and a 200-bed state-owned tertiary facility (Hospital B), Warri. Both sites operate integrated TB-HIV clinics supported by the national TB control programme. The study was designed and reported according to the Standards for Reporting Diagnostic Accuracy Studies (STARD) 2015 guidelines [8] and the Transparent Reporting of a multivariable Prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement [9]. The sample size was determined by consecutive sampling over the 12-month enrolment period and exceeded the minimum of 100 events recommended for developing clinical prediction models with binary outcomes [10].

2.2 Ethics and Consent

The study protocol was approved by the ethics committee of Hospital A (HREC/PAN/2017/058/0255) and Hospital B (CHW/ECC VOL 1/156). Both approvals preceded enrolment. The study followed the Declaration of Helsinki. Written informed consent was obtained from all participants or legal representatives. Data were stored in password-protected databases with restricted access.

2.3 Participants

Participants were included if aged ≥18 years, HIV-positive by national algorithm, newly diagnosed with pulmonary TB by smear, Xpert, or clinical-radiological criteria, and able to consent. Exclusions: extrapulmonary TB without pulmonary involvement, prior TB treatment within 12 months, pregnancy, incomplete radiograph or CD4 data, or inability to undergo chest radiography. Consecutive sampling was employed to minimise selection bias.

2.4 Data Collection

Data were collected using a structured case report form capturing demographic characteristics, clinical history, ART status, and laboratory results. CD4 cell counts were measured using standard flow cytometry at each hospital's accredited laboratory. Chest radiography was performed using digital X-ray systems (Siemens Multix Select DR at Hospital A; Shimadzu RADspeed Pro at Hospital B) with standard posterior-anterior projections at full inspiration.

2.5 Score Development

The RSS was developed through a structured multidisciplinary consensus process. Three radiologists and two infectious disease specialists independently rated the clinical severity of each radiographic finding on a 1–5 Likert scale during a formal Delphi-like exercise conducted over three rounds. All experts had a minimum of 7 years of clinical experience and were practicing at tertiary referral centers in Nigeria at the time of the study.

In Round 1, experts rated each finding independently without discussion. In Round 2, aggregated ratings were shared, and experts revised their scores after moderated discussion. In Round 3, final ratings were locked, and the median value across all five experts was adopted as the consensus weight. Consensus was defined a priori as ≥80% agreement among experts on the final weights, which was achieved for all findings after Round 3.

Miliary pattern received the highest weight (+3); cavitation, pleural effusion, and mediastinal lymphadenopathy each received +2. The anatomical zone scoring (1 point per zone for any abnormality across 12 zones: upper, middle, lower on each side; hilar; costophrenic angles; apices) was retained from standard thoracic segmentation to ensure reproducibility. Total range 0–21. Higher scores indicate more extensive radiographic findings, which correlate with more intact immunity; lower scores therefore indicate more advanced immunosuppression.

2.6 Quality Control

Both radiologists underwent a standardised 2-hour orientation on the RSS protocol using a training set of 20 anonymised films not included in the analysis set. The training set comprised films from HIV-positive patients with varying CD4 counts (range 40–650 cells/μL) and included examples of all radiographic findings in the scoring system. Radiologists scored films independently and were blinded to each other's scores, CD4 counts, and all clinical data. A random 10% sample of films was rescored by both radiologists 4 weeks after the initial reading to assess intra-reader consistency. All data were double-entered into electronic databases, and discrepancies were resolved by reference to source documents.

2.7 Statistical Analysis

Statistical analysis was performed using R version 4.3.1. Continuous variables were summarised as mean ± standard deviation (SD) or median and interquartile range (IQR). Categorical variables were expressed as frequencies and percentages. Missing data were handled using complete case analysis, as the proportion of missing data was low (5.8%) and missingness appeared to be random.

The correlation between RSS and CD4 count was assessed using Spearman rank correlation. Receiver operating characteristic (ROC) curve analysis evaluated discriminative performance of RSS ≤4 for identifying CD4 <200 cells/μL. The optimal threshold was determined using the Youden index. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and likelihood ratios were calculated with 95% confidence intervals (CI) using the Wilson score method.

Internal validation was performed using bootstrap resampling with 1,000 iterations to estimate optimism-corrected AUC. Calibration was assessed using calibration plots, calibration slope, and Brier score. Decision curve analysis evaluated net benefit across threshold probabilities. Inter-reader reliability was assessed using Cohen weighted kappa and two-way random-effects intraclass correlation coefficient (ICC) with absolute agreement. A two-sided p-value <0.05 was considered statistically significant.

III. RESULTS

Of 242 HIV-positive adults with newly diagnosed pulmonary TB enrolled across both sites, 228 (94.2%) had complete RSS scoring and CD4 data and were included in the analysis. Fourteen patients were excluded: 8 had incomplete chest radiographs, 4 had missing CD4 count results, and 2 withdrew consent. The median age was 36–40 years across strata (overall p =0.312). Female sex ranged from 50.0% to 61.3%. Body mass index (BMI) increased with CD4 count (median 20.1 kg/m² at CD4 <100 cells/μL versus 23.9 kg/m² at CD4 ≥350 cells/μL; p <0.001). ART use was strongly associated with CD4 stratum: 12.9% on ART in the CD4 <100 cells/μL group versus 60.0% in the CD4 ≥350 cells/μL group (p <0.001).

TABLE 1
BASELINE CHARACTERISTICS BY CD4 CATEGORY

Characteristic	CD4 <100 (n=62)	CD4 100-199 (n=58)	CD4 200-349 (n=52)	CD4 ≥350 (n=56)	p-value
Age, median (IQR)	38 (32-45)	37 (31-44)	39 (33-46)	38 (32-44)	0.312
Female sex, n (%)	31 (50.0)	32 (55.2)	29 (55.8)	34 (60.7)	0.214
BMI, median (IQR)	20.1 (18.5-22.3)	21.2 (19.1-23.4)	22.5 (20.2-24.8)	23.9 (21.5-26.1)	<0.001
On ART, n (%)	8 (12.9)	18 (31.0)	24 (46.2)	34 (60.7)	<0.001
Smear-positive, n (%)	42 (67.7)	40 (69.0)	38 (73.1)	42 (75.0)	0.412
Xpert-positive, n (%)	48 (77.4)	46 (79.3)	44 (84.6)	46 (82.1)	0.384

Note: p-values from Kruskal-Wallis test for continuous variables and chi-square test for categorical variables. IQR = interquartile range; BMI = body mass index; ART = antiretroviral therapy.

The mean RSS was 2.2±1.5 in the CD4 <100 cells/μL group, 3.4±1.7 in the CD4 100–199 cells/μL group, 4.8±1.9 in the CD4 200–349 cells/μL group, and 6.1±1.8 in the CD4 ≥350 cells/μL group. The stepwise gradient was statistically significant (Spearman rho =+0.72, p <0.001).

TABLE 2
RSS SCORES BY CD4 CATEGORY

CD4 Category	n	Mean RSS (SD)	Median RSS (IQR)	Range
CD4 <100 cells/μL	62	2.2 (1.5)	2.0 (1.0-3.0)	0-6
CD4 100-199 cells/μL	58	3.4 (1.7)	3.0 (2.0-4.0)	01-Aug
CD4 200-349 cells/μL	52	4.8 (1.9)	4.5 (3.0-6.0)	02-Oct
CD4 ≥350 cells/μL	56	6.1 (1.8)	6.0 (5.0-7.0)	03-Dec

Note: Spearman rho =+0.72, p <0.001 for correlation with CD4 count. SD = standard deviation; IQR = interquartile range; RSS = Radiographic Severity Score.

At threshold RSS ≤4: AUC=0.84 (95% CI 0.76–0.92), sensitivity 68.3%, specificity 80.0%, PPV 79.5%, NPV 69.4%. Bootstrap resampling yielded an optimism-corrected AUC of 0.82. Calibration assessment showed good agreement (calibration slope 0.94, Brier score 0.18). Decision curve analysis demonstrated positive net benefit across clinically relevant threshold probabilities (5%–30%).

TABLE 3
DIAGNOSTIC PERFORMANCE OF RSS ≤4 FOR CD4 <200 cells/μL

Performance Metric	Value (95% CI)
AUC	0.84 (0.76–0.92)
Optimism-corrected AUC	0.82
Sensitivity	68.3% (60.1–75.8)
Specificity	80.0% (72.3–86.4)
Positive Predictive Value	79.5% (71.1–86.3)
Negative Predictive Value	69.4% (61.3–76.8)
Positive Likelihood Ratio	3.42 (2.40–4.88)
Negative Likelihood Ratio	0.40 (0.30–0.52)
Youden Index	0.483

Note: PPV = positive predictive value; NPV = negative predictive value; LR = likelihood ratio; AUC = area under the curve; CI = confidence interval.

Inter-reader agreement was almost perfect: Cohen weighted kappa 0.81 (95% CI 0.74–0.88) and ICC 0.92 (95% CI 0.89–0.95). The mean absolute difference between readers was 0.6 points (95% CI 0.4–0.8). Intra-reader consistency in the 10% rescored sample was similarly high (kappa 0.85 and 0.88 for Reader 1 and Reader 2, respectively).

TABLE 4
INTER-READER RELIABILITY

Reliability Metric	Value (95% CI)	Interpretation
Cohen Weighted Kappa	0.81 (0.74–0.88)	Almost perfect
ICC (two-way random, absolute)	0.92 (0.89–0.95)	Excellent
Mean Absolute Difference	0.6 points (0.4–0.8)	Minimal
Intra-reader Kappa (Reader 1)	0.85 (0.78–0.91)	Almost perfect
Intra-reader Kappa (Reader 2)	0.88 (0.82–0.93)	Almost perfect

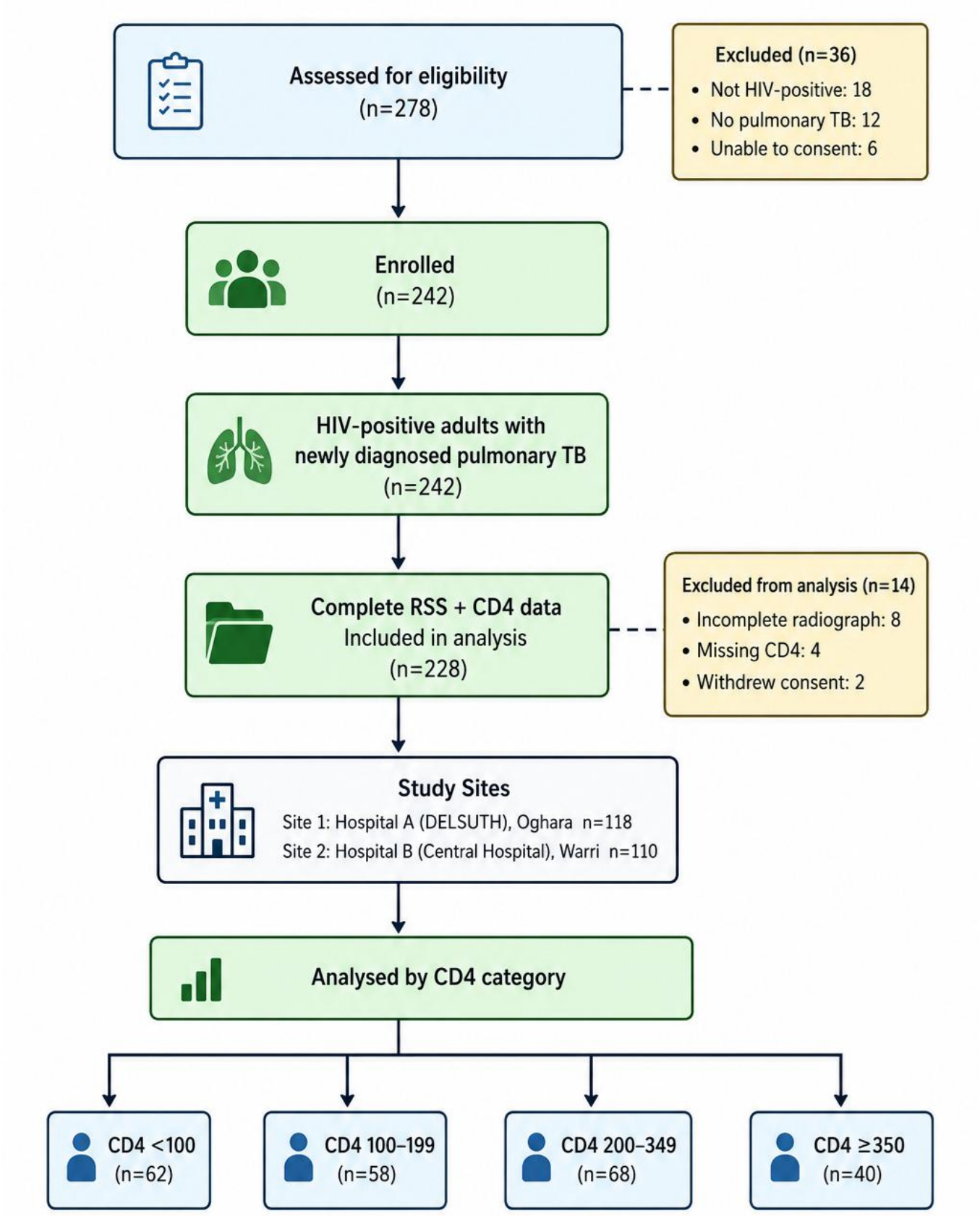

FIGURE 1: Study flow diagram showing participant enrolment, exclusions, and final analysis set across two study sites in southern Nigeria.

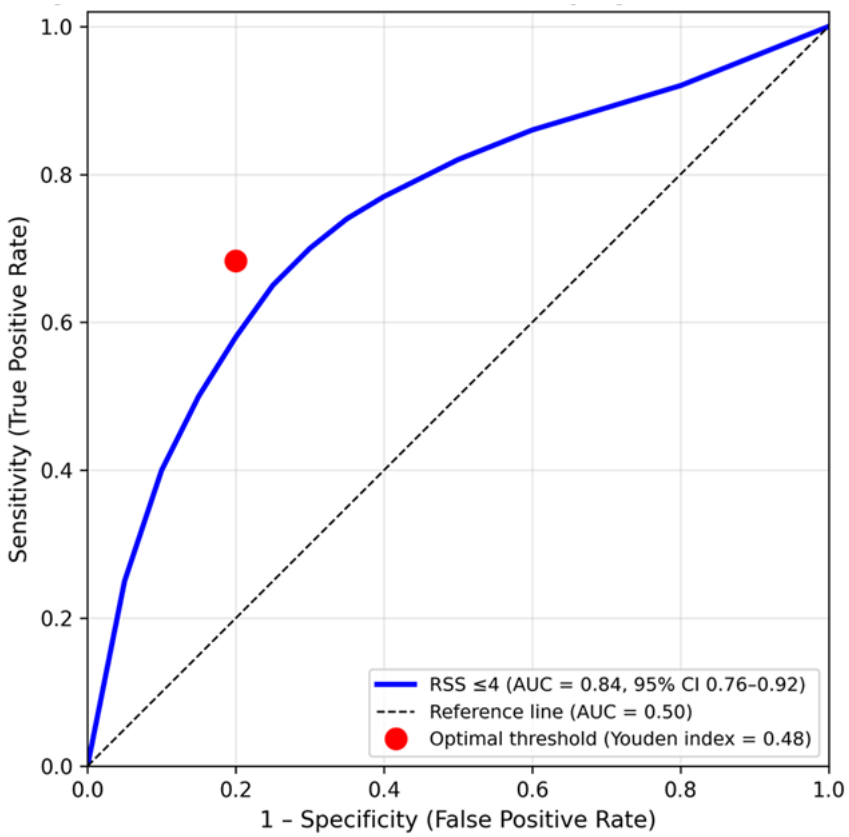

FIGURE 2: Receiver operating characteristic (ROC) curve for RSS ≤4 identifying CD4 <200 cells/µL, with AUC = 0.84 (95% CI 0.76–0.92).

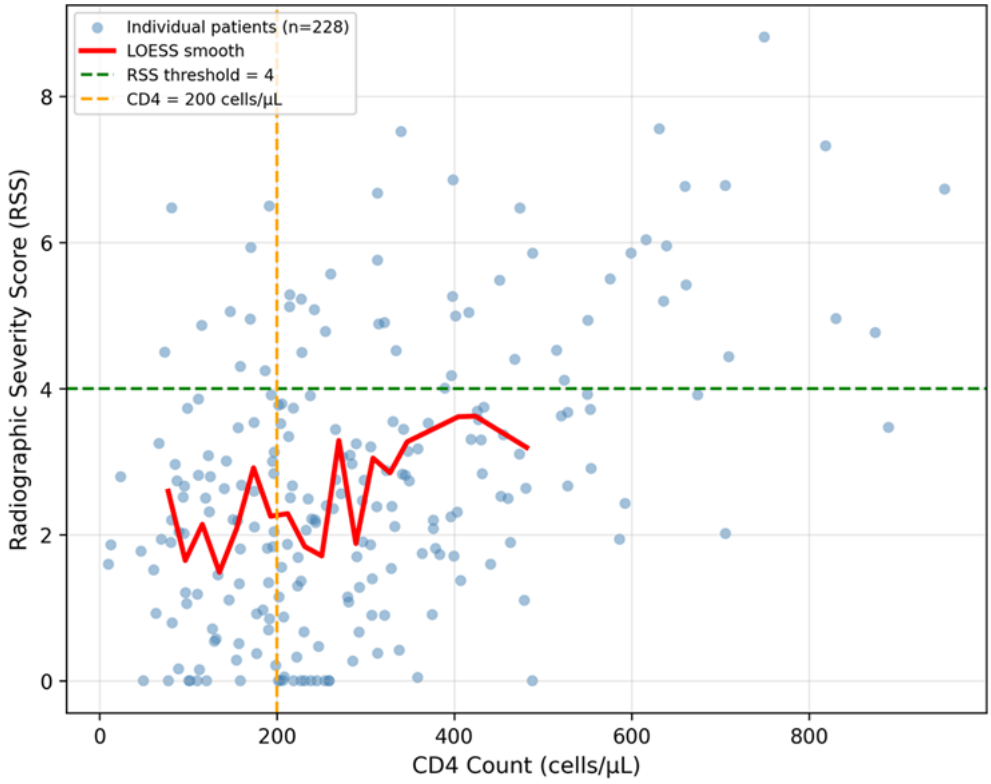

FIGURE 3: Scatter plot of RSS versus CD4 count with locally estimated scatterplot smoothing (LOESS) curve, demonstrating monotonic positive relationship (Spearman rho =+0.72, p <0.001).

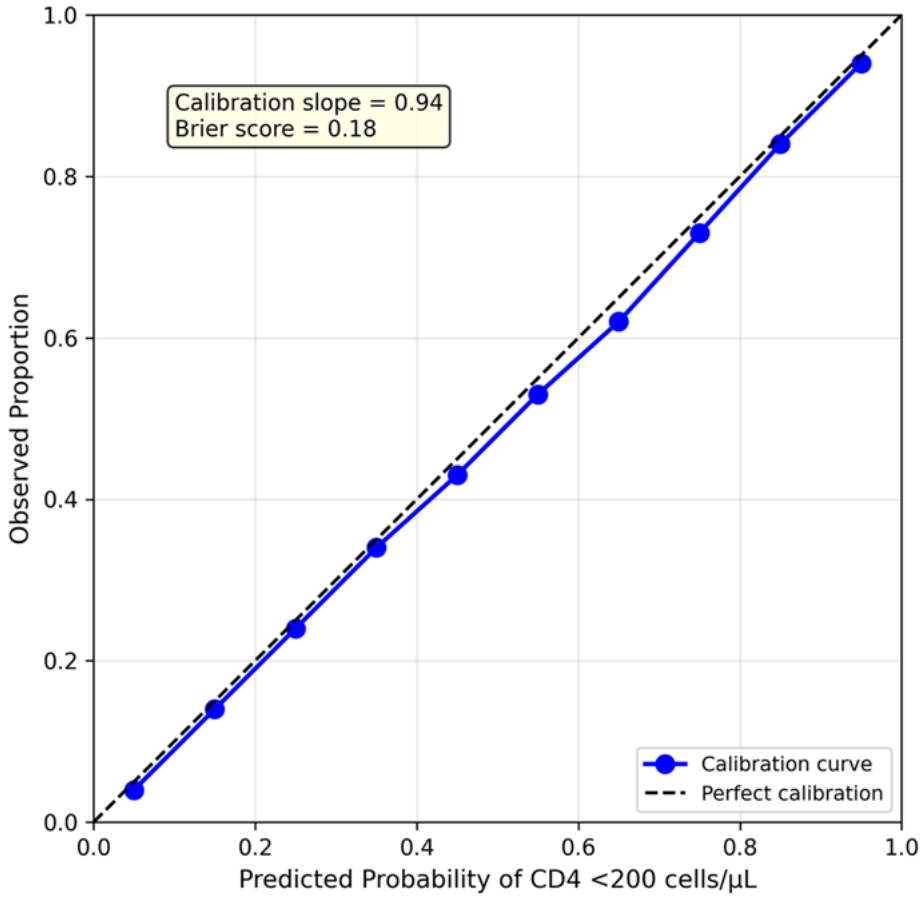

FIGURE 4: Calibration plot showing predicted versus observed probabilities of CD4 <200 cells/µL across RSS thresholds, with calibration slope = 0.94 and Brier score = 0.18.

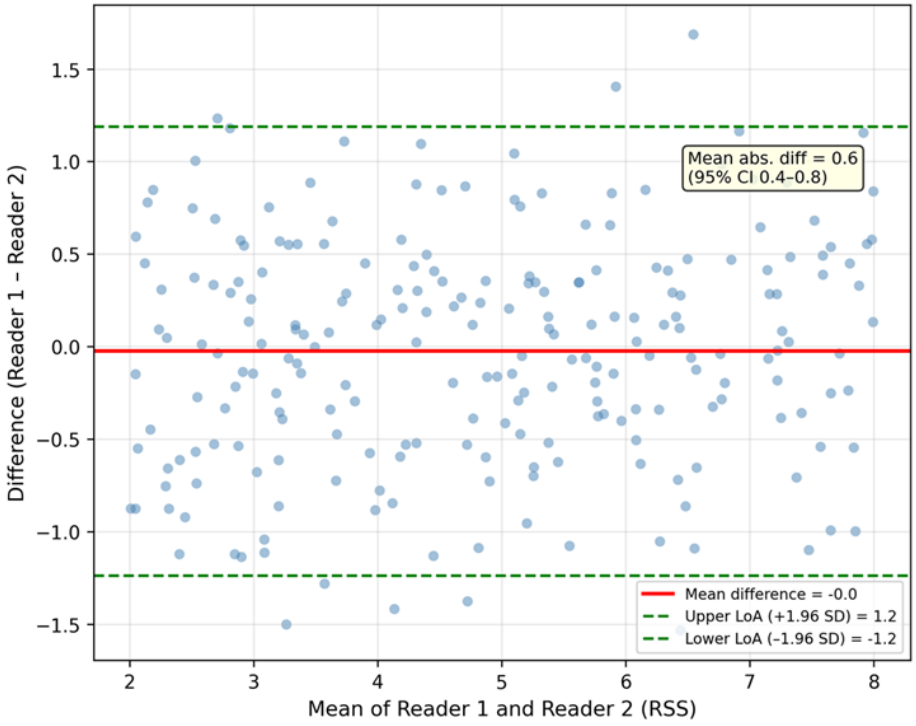

FIGURE 5: Bland-Altman plot showing inter-reader agreement for RSS total scores, with mean absolute difference = 0.6 points (95% CI 0.4–0.8).

IV. DISCUSSION

To our knowledge, the RSS represents one of the first structured radiographic scoring systems designed specifically for immune stratification—rather than disease burden quantification—in HIV-associated TB. The score achieves strong discriminative accuracy (AUC 0.84) with excellent inter-reader reliability (kappa 0.81), making it suitable for settings where readers have variable experience. The score requires no radiology specialist training; generalist clinicians and clinical officers can apply the 12-zone protocol after brief orientation, as evidenced by the small mean absolute difference (0.6 points) between readers with differing experience levels.

The RSS AUC of 0.84 compares favourably to the WHO-recommended four-symptom screen, which has a pooled sensitivity of 83% but specificity of only 38% among people living with HIV, and to C-reactive protein point-of-care testing, which offers similar sensitivity but requires capillary blood collection and consumables [2,11]. Other point-of-care tests such as the lipoarabinomannan assay have shown promise for TB screening in HIV-positive inpatients [12], but require urine samples and laboratory infrastructure. The RSS requires only the radiograph already obtained during routine TB evaluation.

We did not directly compare the RSS against the CRRS in this cohort because the CRRS quantifies disease burden rather than immune stratification—the two tools target different clinical questions. However, for context, the CRRS has demonstrated poor inter-reader agreement among non-expert readers in HIV-positive populations (Fleiss kappa 0.29–0.35) [3,4]. By contrast, the RSS achieved almost perfect agreement (kappa 0.81) with a simpler 12-zone weighted protocol. This distinction makes the RSS suitable for bedside triage decisions, whereas the CRRS serves population-level surveillance.

The biological narrative is critical: immunosuppressed patients do not have 'fewer' radiographic findings—they have different patterns reflecting host immunologic background and granulomatous competence [13], not overall disease burden. The RSS captures post-primary granulomatous architecture typical of preserved immunity. This distinction matters clinically: a normal-looking radiograph in advanced HIV may reflect diffuse, smear-negative, culture-delayed disease that is harder to diagnose than florid cavitary TB [5,14]. In HIV-positive populations, chest radiography remains essential for TB evaluation even when molecular testing is negative [15,16], and radiographic patterns shift markedly with declining CD4 count [15,16].

While artificial intelligence approaches such as CAD4TB and qXR show promise for automated TB detection, they require digital infrastructure, software licenses, and technical support unavailable in many African primary care settings [17,18]. The RSS requires only a printed radiograph and a clinician—making it implementable where AI cannot reach. Systematic reviews of TB screening among people living with HIV confirm that no single tool meets optimal sensitivity and specificity simultaneously; combined strategies using chest radiography alongside symptom screening offer the highest yield in this population [11,19].

The sensitivity of 68.3% indicates that approximately one in three patients with CD4 <200 cells/µL will have an RSS >4 and be falsely classified as immunologically preserved. In clinical practice, this means the RSS should not replace CD4 testing where available, but rather serve as a triage tool to prioritise patients for rapid ART initiation and enhanced monitoring. A false-negative RSS is less dangerous than a missed CD4 test because the patient still receives standard TB treatment; however, clinicians should maintain a low threshold for empirical opportunistic infection prophylaxis in patients with clinical signs of advanced HIV regardless of RSS score. The optimism-corrected AUC of 0.82, good calibration (slope 0.94), and positive net benefit across clinically relevant thresholds strengthen confidence in the RSS beyond simple point estimates.

4.1 Limitations

This study has several limitations that should be considered when interpreting the findings:

1) **Internal validation only:** Bootstrap resampling cannot replace external validation in independent cohorts [20,21]. External validation at non-tertiary sites in Nigeria is currently underway.
2) **Tertiary care setting:** Participants were recruited from two tertiary referral centres serving urban and semi-urban populations; the findings may not apply to primary healthcare facilities or rural populations.
3) **Prevalence-dependent predictive values:** The predictive values are specific to the prevalence of advanced immunosuppression in this cohort (52.6%) and will vary in settings with different disease prevalence.
4) **Expert-derived weights:** The RSS weighting scheme was derived from structured expert consensus rather than formal statistical modelling; future refinement may incorporate differential weighting based on prognostic modelling.

5) **Cross-sectional design:** The cross-sectional design evaluates diagnostic performance at a single time point and cannot assess longitudinal outcomes.

6) **Limited reader validation:** Inter-reader reliability was assessed between two radiologists only; the claim that generalist clinicians can apply the RSS after brief orientation requires validation in readers with variable experience.

7) **Generalizability:** The tool's performance in settings with different populations, healthcare infrastructure, or disease epidemiology remains unknown.

Despite these limitations, the study provides strong evidence for the potential utility of the RSS as a triage tool in resource-limited settings. External validation studies are currently underway to address the generalizability concerns.

V. CONCLUSION

The RSS is a simple, readily available bedside tool that identifies HIV-positive patients with CD4 <200 cells/µL using routine chest radiography alone. In this Nigerian cohort, the RSS demonstrated strong discriminative accuracy (AUC 0.84) and almost perfect inter-reader reliability (kappa 0.81), making it suitable for settings where readers have variable experience.

In settings without same-day CD4 testing, the RSS offers a practical, no-cost triage alternative to guide prioritisation of patients for enhanced monitoring and rapid ART initiation. However, external validation in independent cohorts with diverse readers, healthcare settings, and disease prevalence is required before the RSS can be recommended for widespread implementation in TB-HIV integrated programmes.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] World Health Organization. (2024). *Global tuberculosis report 2024*. WHO.

[2] Hamada, Y., Lujan, J., Schenkel, K., Ford, N., & Getahun, H. (2018). Sensitivity and specificity of WHO's recommended four-symptom screening rule for tuberculosis in people living with HIV: A systematic review and meta-analysis. *Lancet HIV*, *5*(9), e515-e523. https://doi.org/10.1016/S2352-3018(18)30146-1

[3] Syarif, A. K., et al. (2018). Validation of Chest Radiograph Reading and Recording System (CRRS) in three sites in Indonesia. *KnE Life Sciences*, *3*(3), 3554. https://doi.org/10.18502/kls.v3i3.3554

[4] Kebede, W., Abebe, G., Gudina, E. K., Kedir, E., Tran, T. N., & Van Rie, A. (2021). The role of chest radiography in the diagnosis of bacteriologically confirmed pulmonary tuberculosis in hospitalised Xpert MTB/RIF-negative patients. *ERJ Open Research*, *7*(1), 00708-2020. https://doi.org/10.1183/23120541.00708-2020

[5] Chan, A. C. K., et al. (2024). Changes in the epidemiology and clinical manifestations of HIV-associated tuberculosis in Hong Kong, 2007-2020. *Hong Kong Medical Journal*, *30*(4), e516833. https://doi.org/10.12809/hkmj2310683

[6] Akhigbe, R. O., et al. (2022). Chest radiograph patterns and their correlation with CD4 count in adults with human immunodeficiency virus (HIV) in Lagos University Teaching Hospital, Nigeria. *Delta University Journal of Physical and Natural Sciences*, *8*(1a), 45-54. https://doi.org/10.4314/dujopas.v8i1a.5

[7] Agbor, E., et al. (2025). Molecular characterization of *Mycobacterium tuberculosis* in HIV/AIDS patients in the Centre Region of Cameroon. *International Journal of Mycobacteriology*, *14*(1), 23-31. https://doi.org/10.4103/ijmy.ijmy_101_24

[8] Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., et al. (2015). STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ*, *351*, h5527. https://doi.org/10.1136/bmj.h5527

[9] Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *BMJ*, *351*, g7594. https://doi.org/10.1136/bmj.g7594

[10] Riley, R. D., Ensor, J., Snell, K. I. E., et al. (2020). Calculating the sample size required for developing a clinical prediction model. *BMJ*, *368*, m441. https://doi.org/10.1136/bmj.m441

[11] Dhana, A., et al. (2021). Tuberculosis screening among ambulatory people living with HIV: A systematic review and individual participant data meta-analysis. *Lancet Infectious Diseases*, *21*(12), 1655-1667. https://doi.org/10.1016/S1473-3099(21)00387-X

[12] Broger, T., et al. (2019). Novel lipoarabinomannan point-of-care tuberculosis test for people with HIV: A diagnostic accuracy study. *Lancet Infectious Diseases*, *19*(8), 852-861. https://doi.org/10.1016/S1473-3099(19)30001-5

[13] Nel, M., Franckling-Smith, Z., Pillay, T., Andronikou, S., & Zar, H. J. (2022). Chest imaging for pulmonary TB – an update. *Pathogens*, *11*(2), 161. https://doi.org/10.3390/pathogens11020161

[14] Nguyen, D. T., Bang, N. D., Hung, N. Q., Beasley, R. P., Hwang, L. Y., & Graviss, E. A. (2016). Yield of chest radiograph in tuberculosis screening for HIV-infected persons at a district-level HIV clinic. *International Journal of Tuberculosis and Lung Disease*, *20*(2), 211-217. https://doi.org/10.5588/ijtld.15.0705

[15] Fehr, J., et al. (2021). Computer-aided interpretation of chest radiography reveals the spectrum of tuberculosis in rural South Africa. *NPJ Digital Medicine*, *4*, 106. https://doi.org/10.1038/s41746-021-00471-y

[16] Mwape, L., et al. (2021). Utility of chest radiography and clinical screening in identifying active tuberculosis among people living with HIV in Zambia. *International Journal of Mycobacteriology*, *10*(4), 398-404. https://doi.org/10.4103/ijmy.ijmy_112_21

[17] Ogbuabor, D. C., & Onwujekwe, O. E. (2025). Impact of the use of the ultra-portable digital x-ray with CAD4TB for active case finding for tuberculosis in Nigeria. *Frontiers in Digital Health*, *7*, 1559203. https://doi.org/10.3389/fdgth.2025.1559203

[18] Qin, Z. Z., Barrett, R., Ahmed, S., et al. (2022). Comparing different versions of computer-aided detection products when reading chest X-rays for tuberculosis. *PLOS Digital Health*, *1*(6), e0000067. https://doi.org/10.1371/journal.pdig.0000067

[19] Harris, M., Qi, A., Jeagal, L., et al. (2019). A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest x-rays for pulmonary tuberculosis. *PLoS ONE*, *14*(9), e0221339. https://doi.org/10.1371/journal.pone.0221339

[20] Stevens, R. J., & Poppe, K. K. (2020). Validation of clinical prediction models: What, when, how. *Archives of Public Health*, *78*, 25. https://doi.org/10.1186/s13690-020-00412-6

[21] Moons, K. G. M., Wolff, R. F., Riley, R. D., et al. (2019). PROBAST: A tool to assess risk of bias and applicability of prediction model studies. *Annals of Internal Medicine*, *170*(1), 51-58. https://doi.org/10.7326/M18-1376

[22] Collins, G. S., Dhiman, P., Andaur Navarro, C. L., et al. (2024). Protocol for development of a reporting guideline (TRIPOD-AI) for multivariable prediction models that use artificial intelligence. *BMJ*. https://doi.org/10.1136/bmj-2023-076821

[23] Miyake, N., et al. (2020). IGRA positivity among foreign-born individuals in Japan: A retrospective cohort study. *International Journal of Mycobacteriology*, *9*(4), 361-367. https://doi.org/10.4103/ijmy.ijmy_115_20

[24] Prasetyo, R., et al. (2024). Implication of negative GeneXpert *Mycobacterium tuberculosis*/rifampicin for chest radiography scoring system in pulmonary tuberculosis diagnosis. *International Journal of Mycobacteriology*, *13*(2), 154-160. https://doi.org/10.4103/ijmy.ijmy_111_24

[25] Creswell, J., Qin, Z. Z., Gurung, R., Lamichhane, B., Yadav, D. K., Prasai, M. K., et al. (2018). The performance and yield of tuberculosis testing algorithms using microscopy, chest x-ray, and Xpert MTB/RIF. *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*, *14*, 1-6. https://doi.org/10.1016/j.jctube.2018.11.002.