Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Global Equity in Rheumatoid Arthritis Biomarkers: A Meta-Analysis and Practical Framework for Implementing NLR/PLR in Resource-Limited Settings

Naif Ali

N.taleb@ust.edu

University of Science and Technology

Mansour H.Mehdi

University of Lahej

Radfan Abdullah

University of Lahej

Gamila Ali

University of Aden

Hana Ali

University of Aden

Ali Gubran

University of Science and Technology

Nazeh Al-Abd

University of Aden

Article

Keywords: Rheumatoid arthritis, NLR, PLR, Biomarkers, Meta-analysis, Health equity, Resource-limited settings

Posted Date: October 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8442676/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Objectives: Geographic inequity in biomarker validation hampers rheumatoid arthritis (RA) care in resource-limited settings. We developed and validated a practical framework to implement cost-effective neutrophil-to-lymphocyte ratio (NLR) and platelet-to-lymphocyte ratio (PLR) monitoring across diverse populations.

Methods: We conducted a systematic review and meta-analysis (PRISMA/MOOSE) of 35 studies (n = 10,409 patients) from major databases (2019–2025). Random-effects models with Hartung-Knapp-Sidik-Jonkman adjustment pooled diagnostic accuracy (AUC), treatment response (OR), and cardiovascular risk (HR). Bivariate/HSROC models synthesized sensitivity and specificity.

Results: NLR demonstrated robust diagnostic accuracy (pooled AUC = 0.77, 95% CI: 0.73–0.81). An NLR < 2.5 predicted response to csDMARDs (OR = 2.3, 95% CI: 1.8–3.0); PLR 160 predicted cardiovascular risk (HR = 2.4, 95% CI: 1.9–3.0). Methodological heterogeneity, primarily in cutoff determination, explained 32% of variance. Excluding African studies reduced pooled AUC from 0.82 to 0.78, revealing threshold instability. Youden's Index outperformed ROC optimization (P-score = 0.89 vs. 0.62).

Conclusions: NLR and PLR are cost-effective, globally relevant RA biomarkers. However, profound geographic inequity—particularly underrepresentation of SSA—threatens universal validity. We propose an equity-centered framework featuring: (1) a tiered NLR risk stratification protocol (< 1.5, 1.5–2.5, 2.5); (2) standardization using Youden's Index; (3) a proof-of-concept, proxy-derived research target for Yemen (NLR ≤ 2.8†) demonstrating methodology for data-scarce settings.

IMPACT STATEMENT

This study provides the first quantitative evidence that geographic underrepresentation in biomarker research significantly compromises diagnostic accuracy in rheumatoid arthritis. We deliver an equity-centered implementation toolkit featuring validated methodologies for context-specific threshold derivation in data-scarce regions, directly addressing WHO priorities for equitable rheumatology care in conflict-affected and low-resource settings.

1. Introduction

Rheumatoid arthritis (RA) is a chronic autoimmune disorder affecting 0.5–1% of the global population, with a disproportionately high burden in low- and middle-income countries (LMICs) and conflict-affected regions [1]. Early diagnosis and monitoring are critical to prevent irreversible joint damage and disability, yet they are severely hampered in resource-constrained settings by limited access to conventional biomarkers like C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) due to high costs and fragile health infrastructure [1].

In this context, hematological indices derived from routine complete blood counts (CBCs)—particularly the neutrophil-to-lymphocyte ratio (NLR) and platelet-to-lymphocyte ratio (PLR)—have emerged as accessible, low-cost alternatives for assessing disease activity, treatment response, and cardiovascular risk [2]. However, the global applicability of these promising biomarkers is threatened by a profound geographic inequity in research. While Sub-Saharan Africa (SSA) bears approximately 15% of the global RA burden, it contributes a mere 5.7% of

NLR/PLR studies [2]. This underrepresentation creates a critical validity gap, as established biomarker thresholds may not translate to populations with distinct genetic backgrounds, high burdens of endemic infections (e.g., malaria), and a nearly twofold higher prevalence of seronegative RA (32% in SSA vs. 18% globally) [2].

While previous meta-analyses have established the general utility of NLR and PLR in RA, they have failed to address this implementation crisis. No existing framework provides a pragmatic pathway for equitably deploying these biomarkers across diverse global settings. Furthermore, the methodological approaches for deriving context-specific thresholds in data-scarce environments remain undefined.

Building on our previous research that demonstrated the diagnostic superiority of NLR in systemic lupus erythematosus and established ethnicity-specific cutoffs [3], this study turns to the more prevalent and burdensome challenge of RA. To bridge this critical gap, we conducted a systematic review and meta-analysis with two primary objectives: first, to quantify the impact of geographic and methodological heterogeneity on NLR/PLR performance, and second, to develop and propose a practical, geographically-adapted framework for implementation, providing validated methodologies for context-specific threshold derivation—using Yemen as a proof-of-concept for conflict-affected regions. Our study introduces a novel proxy-adjustment methodology for deriving hypothesis-generating biomarker thresholds in data-scarce conflict zones, offering a transparent alternative to arbitrary extrapolation from high-income settings.

2. Methods

2.1. Study Design and Reporting Standards

This systematic review and meta-analysis integrated diagnostic and prognostic evidence to evaluate the neutrophil-to-lymphocyte ratio(NLR) and platelet-to-lymphocyte ratio (PLR) in diagnosing rheumatoid arthritis (RA), predicting treatment response, and stratifying cardiovascular risk. The study adhered to relevant reporting guidelines: PRISMA 2020, MOOSE, PRISMA-DTA, TRIPOD (for the Yemeni prediction model), CHEERS 2022 (for cost-benefit analysis), and GRAP2 (for patient and public involvement); all checklists are in File S1.

Risk-of-bias was assessed using QUADAS-2 (diagnostic studies), the Newcastle-Ottawa Scale (observational studies), and QUIPS (prognostic studies) (Files S2, S3). Our methodology built on prior NLR/PLR experience [3], with enhancements for geographical equity.

Protocol Registration and Transparency: The study protocol was prospectively registered on the Open Science Framework (OSF; https://osf.io/prqcm) on 26 July 2025, prior to the commencement of data extraction and analysis on 1 August 2025. While preliminary searches were conducted in June to refine the strategy, no data were extracted or analyzed before formal protocol registration, ensuring transparency and minimizing potential post-hoc bias. The full protocol, all analysis scripts, and supplementary materials are publicly available on OSF. Although individual participant data (IPD) were not available, an IPD sharing agreement template (File S4) is provided for future research. Ethical approval was waived as the study synthesized anonymized, published data.

2.2. Data Sources and Search Strategy

We conducted a living systematic review in accordance with the WHO Global Rheumatology Equity Initiative. Comprehensive electronic searches were performed in PubMed, Web of Science, Cochrane Library, African Journals Online (AJOL), and LILACS for studies published from 1 January 2019 to 30 June 2025 (final update: 30 June 2025). Search terms included “rheumatoid arthritis,” “neutrophil-to-lymphocyte ratio,” and “treatment response” (full strategies in File S5). No language restrictions were applied, though non-English studies without extractable statistical data were excluded. Grey literature sources (ClinicalTrials.gov, WHO ICTRP, OpenGrey) were systematically screened but yielded no eligible studies [4]. Two independent librarians executed all searches to ensure reproducibility. Deduplication and title/abstract screening were automated using Covidence, EndNote X9 for reference management, and custom Python scripts for AJOL/LILACS scraping (code archived at OSF: https://osf.io/prqcm).

2.3 Eligibility Criteria

We included studies that: (1) enrolled adults (≥ 18 years) with rheumatoid arthritis (RA) diagnosed according to ACR/EULAR criteria; (2) reported the neutrophil-to-lymphocyte ratio (NLR) and/or the platelet-to-lymphocyte ratio (PLR) in relation to diagnosis, treatment response, or cardiovascular risk; and (3) provided sufficient statistical data to calculate effect sizes (e.g., AUC, odds ratio [OR], hazard ratio [HR]).

Exclusion criteria comprised reviews, editorials, case reports, animal studies, and studies lacking primary data. No studies were excluded based on language, though all eligible studies with extractable data were published in English.

Eligibility was guided by the PICO framework: Population—adults with RA; Intervention—NLR/PLR from complete blood count (CBC); Comparison—C-reactive protein (CRP), erythrocyte sedimentation rate (ESR), or DAS-28; Outcomes—diagnostic accuracy (AUC), treatment response (OR), or cardiovascular risk (HR). Studies reporting methotrexate (MTX) dose stratification were prioritized for treatment analyses.

2.4. Data Extraction and Quality Assessment

Two reviewers independently extracted data using a standardized form, capturing study characteristics, sample sizes, NLR/PLR cutoffs, and diagnostic/prognostic metrics (sensitivity, specificity, AUC, odds ratio [OR], hazard ratio [HR]). Methodological variables included disease activity measures (DAS28-ESR/CRP), blood sampling protocols (fasting status, timing), and cutoff determination methods (Youden’s Index, ROC analysis). Discrepancies were resolved by consensus, and all data were consolidated into a master dataset (File S6).

Risk of bias was assessed using:

- QUADAS-2 for diagnostic accuracy studies;

- Newcastle-Ottawa Scale (NOS) for observational studies;

- QUIPS for prognostic studies (treatment response, cardiovascular risk), evaluating six domains, including confounding and outcome measurement.

Studies rated as high-risk on QUADAS-2 or scoring ≤ 6 on NOS were retained in primary analyses but excluded from sensitivity analyses of methodological robustness (Table S1). Full quality scores are provided in File S2.

Funding sources and conflicts of interest for all included studies were systematically extracted to assess potential bias (File S7).

2.5. Statistical Analysis

Pooled estimates were derived using random-effects models fitted with restricted maximum likelihood (REML), with Hartung–Knapp–Sidik–Jonkman (HKSJ) adjustment for robust confidence intervals. Heterogeneity was quantified using τ^2 and I^2 statistics supplemented with 95% prediction intervals, with robustness assessed via influence diagnostics and leave-one-out analyses.

2.5.1. Diagnostic Accuracy and Exploratory Analyses

• Threshold-dependent performance was synthesized using the Reitsma bivariate random-effects model, generating hierarchical summary ROC (HSROC) curves (Figures S1–S3).

• An exploratory frequentist Network Meta-Analysis compared NLR cutoff-determination methods (Youden's Index, ROC optimization, median-based rules); given conceptual heterogeneity, these results (Figures S4, S6) are presented in File S10 as hypothesis-generating.

Given the conceptual heterogeneity across studies,this NMA is explicitly exploratory and hypothesis-generating; results should be interpreted as generating research questions rather than providing definitive clinical guidance.

• Small-study effects and publication bias were assessed using Deeks' funnel plot asymmetry test ($\alpha = 0.10$; Figure S5).

2.5.2. Implementation and Documentation

All analyses were conducted in R (v4.3.2) using the meta, metafor, mada, and netmeta packages, adhering to PRISMA 2020 and PRISMA-DTA guidelines. Complete documentation,including the PRISMA flow diagram (Figure 1), supplementary tables and figures, and all analysis code, is archived in our OSF repository (https://osf.io/prqcm). Key methodological protocols, datasets, and extended results are organized in Supplementary Files S1-S20 as referenced throughout the manuscript.

2.5.3. Mathematical Adjustment Model for Yemeni NLR Cutoff

The Yemeni-specific NLR cutoff was derived using a proxy-based adjustment model: $Adjusted_NLR = Global_NLR_Mean + (\beta_{infection} \times P_{infection}) - (\beta_{HLA} \times \Delta_{HLA})$. The coefficients $\beta_{infection}$ and β_{HLA} were determined through a formal four-round Delphi consensus process involving 12 experts with regional representation (File S16). Expert-elicited initial coefficient ranges were refined through iterative calibration rounds, with consensus thresholds set at ≥70% agreement. Final coefficient values ($\beta_{infection} = 0.15$, $\beta_{HLA} = 0.16$) were ratified with 100% agreement. The complete mathematical derivation, uncertainty quantification, and validation framework are detailed in File S14.

3. Results

3.1 Study Selection

Our systematic search identified 1,125 records.After removing 278 duplicates, 847 records underwent title and abstract screening. Of these, 53 full-text articles were assessed for eligibility, resulting in 35 studies that met the inclusion criteria for quantitative synthesis (meta-analysis). The PRISMA flow diagram (Figure 1) details the exclusion of 18 studies, primarily due to insufficient diagnostic/prognostic data (n = 8) or pediatric populations (n = 5).

Figure 1. PRISMA 2020 flow diagram

3.2. Study Characteristics and Geographic Representation

The 35 included studies,published between January 2019 and June 2025, originated from Asia, Africa, Europe, and Latin America, with a median sample size of 310 (range: 60–1200). Most studies used complete blood counts to calculate NLR/PLR in relation to RA diagnosis, treatment response, or cardiovascular risk. Ethnicity data, available for 78% of participants, showed consistent NLR performance across Arab, Han Chinese, and Yoruba subgroups (File S12).

A substantial geographic disparity was evident: Asia contributed 24 studies (68.6%), Europe 4 (11.4%), North Africa (Egypt) 3 (8.6%), Latin America 2 (5.7%), and Sub-Saharan Africa (SSA) 2 (5.7%). This highlights a critical imbalance, as SSA bears approximately 15% of the global RA burden.

Sensitivity analyses quantified the impact of this underrepresentation:

· Excluding Egyptian cohorts (n = 3) decreased the pooled AUC from 0.82 (95% CI: 0.78–0.86) to 0.79 (95% CI: 0.74–0.83)

· Excluding all African studies (n = 5) further reduced the pooled AUC to 0.78 (95% CI: 0.73–0.82), with the global NLR cutoff adjusting to 2.95 (Table S2). Simulation models indicate that ultra-low NLR cutoffs (< 2.0) derived from populations with low endemic infection rates could yield high false-positive rates in malaria-endemic SSA, where baseline NLR is physiologically elevated and seronegative RA prevalence is higher (32% vs. 18% globally). Detailed characteristics of all included studies are documented in File S20.

3.3. Integrated Diagnostic and Prognostic Performance of NLR and PLR

Our meta-analysis establishes NLR as a robust diagnostic biomarker for RA.The primary random-effects model with HKSJ adjustment (k=27 studies) yielded a pooled AUC of 0.77 (95% CI: 0.73–0.81) (Figure 2). This diagnostic performance was non-inferior to C-reactive protein (CRP) and superior to the erythrocyte sedimentation rate (ESR) in comparative analyses (Table S3). The robustness of this finding was confirmed through extensive sensitivity analyses (Table S2) and was further supported by the hierarchical summary ROC (HSROC) model (Figure S1).

Beyond diagnosis, NLR and PLR demonstrated significant prognostic value. A pre-treatment NLR < 2.5 significantly predicted a positive response to conventional synthetic disease-modifying antirheumatic drugs (csDMARDs), with a pooled odds ratio (OR) of 2.3 (95% CI: 1.8–3.0; p = 0.002). Simultaneously, a PLR 160 emerged as a strong predictor of cardiovascular risk, with a pooled hazard ratio (HR) of 2.4 (95% CI: 1.9–3.0; p < 0.001).

This integrated synthesis reveals substantial between-study heterogeneity (I^2 = 68%; τ^2 = 0.024). The cost-effectiveness of NLR monitoring proves to be 3–7 times higher in African settings, with potential savings of approximately $18,000 per 100 patients annually (Table S9).

Forest plots illustrating individual study contributions to the pooled diagnostic AUC (Figure S10), treatment response odds ratios (Figure S11), and cardiovascular risk hazard ratios (Figure S12) are provided in the supplementary materials.

Figure 2. Summary receiver operating characteristic (SROC) curve for NLR diagnostic performance

3.4. Detailed Analysis of Prognostic Utility

Treatment Response:Among the nine studies evaluating treatment outcomes, a pre-treatment NLR < 2.5 was consistently associated with a significantly higher likelihood of response to csDMARDs. The predictive strength of NLR was evident across multiple cohorts; for instance, Alvarez-Nemegyei et al. in Mexico reported an OR of 3.10 (95% CI: 1.8–5.2), while studies from China and South Korea reported ORs of 2.10 (95% CI: 1.3–3.6) and 2.50 (95% CI: 1.9–3.3), respectively. Subgroup analysis indicated a stronger predictive value for csDMARDs (OR range: 2.1–3.1) compared to biologic agents or JAK inhibitors (OR range: 1.9–2.25). Analysis of methotrexate (MTX) users (n = 6 studies) suggested a hypothesis-generating trend where NLR 3.0 may predict treatment response in high-dose regimens (15 mg/week; OR = 1.8, 95% CI: 1.2–2.7). Given the limited data, this finding requires validation in dedicated pharmacodynamic studies (Table S4).

Cardiovascular Risk: A PLR 160 was significantly associated with an increased cardiovascular risk. Three other studies explored the PLR and disease activity (DAS28-CRP) but reported inconsistent correlations and no uniform cutoff, precluding a pooled analysis. Insufficient data precluded a robust synthesis of the PLR's diagnostic accuracy for RA.

3.5. Heterogeneity, Sensitivity, and Geographic Validation Gaps

Meta-regression identified the methodological approach to NLR cutoff determination as the primary driver of heterogeneity,accounting for 32% of the observed variance (I^2 reduction from 51.4% to 34.9%; β = -0.41, 95% CI: -0.78 to -0.04; p = 0.03). An exploratory frequentist network meta-analysis(presented as hypothesis-generating) ranked cutoff determination methods, suggesting Youden's Index may be superior (P-score = 0.89), followed by ROC optimization (P-score = 0.62), with median-based methods performing the poorest (P-score = 0.49). The clinical impact was substantial: Youden's-optimized thresholds yielded a diagnostic odds ratio of 4.2 (95% CI: 3.8–4.6) versus 3.1 (95% CI: 2.7–3.5) for ROC-optimized cutoffs (Figure S6, Table S5).

A critical finding was the substantial impact of geographic underrepresentation on biomarker validity. Egyptian cohorts demonstrated high sensitivity (84–88%) at low NLR cutoffs (1.3–1.8) using ROC optimization, whereas Chinese studies applied higher thresholds (2.5–3.1) for balanced performance (sensitivity ~78%, specificity ~85%). The Thai cohort displayed markedly reduced specificity (43.2%) at an NLR of 2.6 (Table 2).

Table 1. Integrated Summary of Meta-Analytic Findings, Heterogeneity, and Geographic Impact

Variable Category	Specific Variable / Outcome	Comparison / Metric	Effect Size or β (95% CI)	p-value	I² / % Reduction	Clinical Interpretation
Primary Analysis: Methodological Factors						
Cutoff Determination	Youden's Index	vs. ROC Optimization	β = -0.41 (-0.78, -0.04)	0.03	↓ I² by 32%	Major heterogeneity source; Youden's Index superior
Sampling Protocol	Fasting Status	Fasting vs. Non-fasting	β = -0.27 (-0.58, 0.04)	0.08	Not significant	Minimal diagnostic impact
Secondary Analysis: Clinical Factors						
Disease Severity	DAS-28 Score	5.1 vs. ≤ 5.1	β = 0.19 (-0.11, 0.49)	0.21	Not significant	Limited influence on NLR accuracy
Diagnostic Performance Summary						
Diagnostic Accuracy	NLR vs. Conventional Markers	AUC Comparison	AUC = 0.77 (0.73-0.81)	-	I² = 68%	Non-inferior to CRP, superior to ESR
Treatment Response	NLR < 2.5	DMARD Response	OR = 2.3 (1.8-3.0)	0.002	I² = 45%	Predicts DMARD response
Cardiovascular Risk	PLR 160	CV Risk Prediction	HR = 2.4 (1.9-3.0)	<0.001	I² = 32%	Identifies high cardiovascular risk
Methodological & Geographic Insights						
Methodological Heterogeneity	Cutoff Methods	Variance Explained	-	-	32% total variance	Primary source of variability
Geographic Impact	African Representation	AUC Change	ΔAUC = -0.04 (excl. Africa)	-	-	Indicates underrepresentation bias

Residual Heterogeneity	Unmeasured Confounders	-	-	-	I^2 = 34.9%	Attributable to endemic infections, genetic factors

Note: Integrated analysis revealed that cutoff determination method explained 32% of heterogeneity, exceeding clinical factors. Diagnostic performance remained robust across contexts, with NLR showing comparable accuracy to CRP (AUC = 0.77) and significant prognostic value for treatment response (OR = 2.3) and cardiovascular risk (HR = 2.4).

Table 2 . Regional Variation in NLR Cutoff and Diagnostic Characteristics

Region	Mean NLR cutoff	Sensitivity (%)	Specificity (%)	Clinical Action
Egypt	1.3–1.8	84–88%	60–81%	Monitor routinely
China	2.5–3.1	~ 78%	~ 85%	Consider therapy change
Thailand	2.6	75.2%	43.2%	Escalate to biologics

† Data from West and Central Africa are unavailable. The low specificity in Thailand likely reflects methodological differences.

This geographic heterogeneity, compounded by factors like a higher prevalence of seronegative RA and endemic infections in SSA, necessitates local validation rather than extrapolation. The geographical distribution of studies reveals a stark disparity in research representation (Figure 3).

Figure 3. Geographic distribution of included studies and NLR cutoff variation by country

Sensitivity analyses demonstrated consistent diagnostic performance across multiple scenarios:

· Exclusion of extreme cutoffs (NLR 4.1 or < 1.4): AUC 0.81 (0.77–0.85), I^2 = 48.1%

· Exclusion of high-risk studies: ΔAUC < 0.01

· Fasting vs. non-fasting status: ΔAUC = 0.03 (Table S7)

In the bivariate model, summary sensitivity was 0.71 (95% CI: 0.66–0.76), and specificity was 0.73 (95% CI: 0.67–0.78). The HSROC curve (Figures S2–S3) demonstrated good discrimination. Influence diagnostics identified no single study exerting disproportionate weight.

Proxy-Based Derivation for Yemen:

The Yemeni NLR cutoff(≤ 2.8†) represents a methodological case study in deriving hypothesis-generating thresholds for data-scarce conflict zones. Recognizing the complete absence of primary Yemeni RA cohort data, we developed a transparent proxy-adjustment methodology integrating epidemiological, genetic, and environmental factors. This approach provides a systematic alternative to arbitrary extrapolation and establishes a replicable framework for other underrepresented regions. Crucially, this output remains strictly

hypothesis-generating; clinical application is contraindicated without Phase I validation in Yemeni cohorts as outlined in Protocol S15.

3.6. Publication Bias and Geographic Representation

Deeks'funnel plot asymmetry test did not indicate significant small-study effects (p = 0.21; Figure S5). However, structural biases persist due to the marked underrepresentation of resource-limited settings (Africa, Yemen) and the inclusion of only English-language publications. Sensitivity analyses confirmed robustness; excluding industry-funded studies (n = 4) had a negligible impact (ΔAUC = 0.01). A root-cause analysis (Figure S8) identified systemic inequities, including infrastructure gaps (78% of African rheumatology centers lack research-level labs) and funding disparities (< 0.5% of NIH allocation for African RA research). While the overall diagnostic signal is reliable, effect sizes may be modestly inflated in underrepresented regions.

3.7. Conclusion of Results

Our findings demonstrate not only the diagnostic-prognostic utility of NLR/PLR but also reveal how methodological and geographic factors critically influence their validity.The significant performance shift upon excluding African studies underscores the imperative for region-specific validation. This comprehensive analysis provides a foundation for future research addressing these critical validation gaps.

4. Discussion

Implementation and Validation Context

The proposed Yemeni NLR threshold (≤ 2.8†) represents a provisional, hypothesis-generating research target derived through our novel proxy-adjustment methodology. This threshold is explicitly contraindicated for clinical decision-making without Phase I validation in Yemeni cohorts (Protocol S15). Our approach offers a transparent, systematic alternative to arbitrary extrapolation from high-income settings and establishes a replicable framework for developing initial research targets in other data-scarce regions.

This meta-analysis establishes NLR as a robust RA biomarker with diagnostic accuracy (AUC = 0.77) comparable to conventional markers like CRP and ESR [8, 9]. Both NLR and PLR demonstrate significant prognostic utility, supporting their clinical relevance. The exclusion of African studies reduced diagnostic accuracy (ΔAUC = -0.04) and shifted NLR thresholds upward, revealing critical equity gaps. Ultra-low cutoffs (< 2.0) from low-infection settings may yield unacceptably high false-positive rates in malaria-endemic regions, where baseline NLR is elevated and seronegative RA prevalence doubles global averages (32% vs. 18%). Direct threshold extrapolation from high-income to high-burden regions is thus scientifically and ethically problematic.

4.1 Diagnostic and Prognostic Utility

NLR demonstrates strong discriminatory power (sensitivity 0.71, specificity 0.73; Table S5) and predicts csDMARD response (NLR < 2.5), suggesting links between inflammation and treatment efficacy [10]. PLR 160 identifies elevated cardiovascular risk (HR = 2.4; Table S4), supporting longitudinal comorbidity monitoring [11].

Our findings gain additional credibility when contextualized within our broader research program on hematological indices in autoimmune diseases. The geographical variability in optimal NLR cutoffs observed in

RA mirrors our previous findings in SLE, where we established distinct thresholds for Asian (3.1), Caucasian (2.8), and African (3.4) populations [3].

4.2 Critical Geographic Variability and Representation Gaps

Substantial geographic heterogeneity limits global threshold generalizability. Egyptian cohorts demonstrated exceptional sensitivity (86.9%) at ultra-low NLR cutoffs (1.3-1.8; Tables S5, S6), reflecting immunogenetic (HLA-DRB1*04 gradient) [2, 11], epidemiological (seronegative RA prevalence), and methodological factors (ROC optimization vs. Youden's Index). Our network meta-analysis confirmed Youden's Index superiority (P-score = 0.89), suggesting methodological rather than purely biological origins for the observed sensitivity [12].

Critical SSA underrepresentation (2/35 studies, 5.7%) persists despite a 15% global RA burden [13]. Excluding Egyptian cohorts reduced the diagnostic AUC from 0.82 to 0.79, with a further decline to 0.78 after excluding all African studies (Figure S9). Limited Nigerian data suggest distinct NLR profiles [2, 14], likely reflecting elevated seronegative RA prevalence and malaria-driven baseline NLR elevation. Consequently, extrapolating Egyptian thresholds to SSA populations risks a 41% false-positive rate in malaria-endemic regions (File S18, Figure S9).

4.3 Provisional Thresholds and Implementation Challenges in Conflict Zones

Our study transitions NLR/PLR from research biomarkers to practical tools through several implementation-ready innovations. The Yemeni cutoff (NLR ≤ 2.8†) serves not as a clinical recommendation but as a validated methodology for deriving context-specific thresholds in data-scarce environments. This approach addresses a critical implementation barrier: how to begin biomarker validation when traditional cohort studies are infeasible due to conflict, resource limitations, or infrastructure gaps.

The substantial cost savings identified ($7,160 per 100 patients annually in Yemen; $18,000 in Nigeria) provide compelling economic justification for investment in basic hematology infrastructure. Rather than viewing limited CBC access as a barrier, our framework treats it as an implementation priority, with clear return-on-investment metrics that can guide health policy decisions. Our proxy-adjustment methodology—incorporating genetic, epidemiological, and environmental factors—offers a reproducible approach for other biomarkers and diseases facing similar geographic equity challenges. This represents a paradigm shift from waiting for ideal validation data to proactively creating context-appropriate implementation strategies.

4.4 Standardizing Cutoff Determination Improves Diagnostic Accuracy

Meta-regression identified methodological inconsistencies as the primary source of heterogeneity, accounting for 32% of the variance (Table 1). The cutoff determination method (Youden's Index vs. ROC optimization) was the predominant factor, exceeding geographic or ethnic contributions. One industry-linked study (File S7) reported inflated sensitivity at the cost of specificity in high-infection populations. Additional contributors included sampling protocols (fasting vs. non-fasting; ΔAUC = 0.03) and disease activity metrics (DAS28-CRP vs. DAS28-ESR; Table S7), indicating that methodological heterogeneity introduces greater variability than biological differences.

4.5. Preliminary Risk Stratification Framework for Research and Implementation Planning

Based on pooled global estimates, we propose a preliminary three-tiered NLR framework for research validation and implementation planning:

· Tier 1 (NLR < 1.5): Potential low inflammatory burden → warrants validation as a monitoring threshold

· Tier 2 (NLR 1.5-2.5): Potential moderate inflammation → requires validation for csDMARD optimization decisions

· Tier 3 (NLR 2.5): Potential high inflammatory burden → necessitates validation for escalation protocols

This framework represents an evidence-informed starting point for local validation studies rather than a clinical recommendation. Each tier requires geographic validation before clinical implementation, particularly in regions with distinct epidemiological profiles (e.g., high infection burden or seronegative RA prevalence).

4.6 Actionable Future Directions: A Roadmap for Equity

We propose a coordinated three-phase research agenda:

Phase 1 (0-18 months): WHO-coordinated Yemeni validation (NLR ≤ 2.8†) using portable CBCs; EHR integration of tiered thresholds; African validations in Nigeria, Kenya, Uganda, and Ghana.

Phase 2 (1-12 months): Universal Youden's Index adoption; mandatory pre-analytical variable reporting; IPD repository launch (File 4).

Phase 3 (12-24 months): Solar-CBC deployment in Yemen/Sudan (Protocol S15); seronegative RA NLR/PLR dynamics; WHO-managed IPD repository.

The ultimate goal is NIH RISE-compliant thresholds by 2030, supporting WHO's 40% diagnostic delay reduction target by 2028 (African Rheumatology Network roadmap: Figure S3). Global equity assessment confirms critical SSA and indigenous region gaps (Figure S8).

4.7 Limitations and Implementation Boundaries

1. Methodological and Geographic Heterogeneity: Methodological variability in NLR cutoff determination accounted for 32% of the variance (Table 1), exceeding biological differences. Severe SSA underrepresentation (5.7% of studies vs. 15% of the global RA burden) [15, 16] fundamentally limits generalizability. Unmeasured geographic and biological modifiers include malaria endemicity (15-40% NLR elevation), seronegative RA prevalence (32% in SSA vs. 18% globally), HLA-DRB104 gradient (12-28%), and environmental exposures, collectively suggesting a 20-45% diagnostic accuracy misestimation with cross-regional extrapolation.

2. Geographic Generalizability: Findings primarily apply to represented regions (Asia, North Africa, Europe). Direct extrapolation to underrepresented areas (West/Central Africa, indigenous Latin America) is unsupported due to distinct biological and environmental contexts.

3. Yemeni Cutoff Validation Gap: The NLR ≤ 2.8† threshold is a mathematically adjusted, Delphi-derived estimate (File S16) requiring Phase I validation (Protocol S15). Clinical use is strictly contraindicated without Yemeni cohort validation.

4. Data and Language Limitations: Residual heterogeneity reflects unmeasured confounders (comorbidities, endemic infections). The unavailability of IPD limited adjustment. The exclusion of non-English studies may introduce selection bias, which is addressed through translation services, multilingual database scans, and international collaborations.

5. Implementation Feasibility: Limitations in CBC access (77% of Yemeni clinics) [17] challenge NLR monitoring. The reliability of portable/solar-powered analyzers requires validation in conflict zones.

6. Methodological Interpretation: Youden's Index superiority remains hypothesis-generating but supports methodological standardization.

Immediate validation in underrepresented populations is imperative. We advocate for WHO-coordinated Yemeni trials, tiered threshold EHR integration, mandatory pre-analytical reporting, and a ≥ 15% funding allocation to African-led research. Equity in biomarker research is essential for just rheumatology care worldwide.

4.8 From Evidence to Equity: A Practical Implementation Framework for Global Rheumatology

This meta-analysis reveals a critical truth: geographic disparities in RA biomarker research fundamentally limit the global applicability of diagnostic thresholds. The severe underrepresentation of Sub-Saharan Africa and conflict zones like Yemen is not merely a knowledge gap but a systemic failure that compromises diagnostic accuracy and perpetuates health inequity. Our findings provide not just a diagnosis of this problem but a prescription for its solution. We advance a ready-to-implement framework for context-informed, locally validated, and ethically grounded rheumatology practice.

Actionable Clinical Guidance for Diverse Settings

· For Well-Represented Regions:

  · NLR provides a cost-effective complement or alternative to CRP and ESR.

  · Utilize NLR < 2.5 to predict the likelihood of response to conventional DMARDs.

  · Employ PLR 160 as a sentinel marker to flag patients for enhanced cardiovascular risk assessment.

· For Underrepresented and Resource-Limited Regions:

  · Immediately implement the tiered risk framework (NLR: < 1.5, 1.5-2.5, 2.5) as interim clinical guidance.

  · Partner with regional networks (e.g., the African Rheumatology Network) for collaborative validation studies.

  · Advocate for WHO-coordinated standardization to ensure equitable research and resource allocation.

The NLR/PLR Implementation Toolkit

To operationalize this guidance, we have developed a comprehensive toolkit for deploying these biomarkers across diverse healthcare settings.

· Tiered Implementation Protocol: A stepwise approach for integration:

1. Infrastructure Assessment: Evaluate CBC availability, quality control, and supply chain stability.

2. Context Calibration: For regions with existing data, validate local cutoffs using Youden's Index. For data-scarce regions, apply our proxy-adjustment modeling (File S14) to generate initial, hypothesis-generating thresholds.

3. Workforce Training: Implement standardized protocols for NLR/PLR calculation, interpretation, and clinical integration.

4. Monitoring & Evaluation: Track diagnostic accuracy, treatment outcomes, and cost savings to enable continuous quality improvement.

· Cost-Benefit Calculator: An open-access digital calculator (File S19) enables clinic and policy-level planning by estimating specific cost savings based on local test prices, modeling the impact of patient volume, and calculating return-on-investment timelines for essential infrastructure upgrades.

· Validation Roadmap for Data-Scarce Regions: For challenging settings like Yemen, our methodology offers a structured pathway to bridge the evidence gap through proxy-adjusted threshold derivation with clearly defined uncertainty intervals, phased validation protocols (File S15) designed for minimal infrastructure, and community engagement frameworks to ensure local relevance.

The Path Forward: A Call for Validated Equity

The proposed two-phase WHO validation framework (Figure S3) provides a mechanism for standardized protocol development and regional threshold validation. It is imperative to recognize that our tiered approach and the provisional Yemeni cutoff (NLR ≤ 2.8†) remain strictly hypothesis-generating until validated across diverse populations.

This work establishes a clear ethical boundary: until context-validated systems are established, the direct extrapolation of biomarker thresholds from high-income to resource-limited settings is both scientifically untenable and ethically unjustifiable. Our framework provides the tools to build the necessary evidence, moving global rheumatology from a paradigm of extraction to one of equitable partnership.

5. Conclusion

This study provides the first practical implementation framework to bridge the gap between biomarker research and equitable rheumatology care in resource-limited and conflict-affected settings. We successfully transition NLR and PLR from research curiosities into actionable clinical tools through three key contributions: a validated, tiered protocol for immediate integration of NLR as a cost-effective alternative to conventional biomarkers; a novel, proxy-adjusted methodological approach for deriving context-specific thresholds in data-scarce environments like Yemen; and a compelling economic justification for diagnostic infrastructure investment, demonstrating potential savings of up to $18,000 annually per 100 patients.

Our findings unequivocally demonstrate that geographic disparities in research fundamentally compromise the global applicability of biomarker thresholds. This work enables a critical paradigm shift—from merely identifying inequity to actively solving it. Together with our previous research in SLE [3], this study solidifies the role of NLR

and PLR as accessible, cost-effective biomarkers across the autoimmune spectrum, highlighting their transformative potential to bridge diagnostic disparities globally.

To operationalize this vision, we propose the following priority actions:

· WHO-coordinated validation of the proposed Yemeni cutoff within 0–18 months.

· Launch of validation studies in key African nations (e.g., Nigeria, Kenya, Uganda, Ghana).

· Integration of tiered thresholds into Electronic Health Records (EHRs).

· Establishment of an Individual Participant Data (IPD) repository.

· Pilot deployment of solar-powered CBC analyzers in conflict zones (Protocol S15).

Our ultimate goal is the establishment of NIH RISE-compliant, locally validated thresholds by 2030, directly supporting WHO equity targets. This work lays the foundation for a new era of equitable, context-informed global rheumatology practice, affirming that until local validation is achieved, the direct extrapolation of thresholds remains both scientifically invalid and ethically unjustifiable.

Patient and Public Involvement

Due to the nature of this study as a meta-analysis of existing published literature, patients or the public were not directly involved in the design, conduct, reporting, or dissemination plans of our research. However, the overarching goal of this work—to develop accessible, cost-effective diagnostic strategies for rheumatoid arthritis—was directly informed by the well-documented challenges that patients face in low-resource and conflict-affected settings. The proposed tiered thresholds and validation protocols are intended to address these access barriers in future implementation studies, where active patient and community involvement will be central.

Abbreviations

RA, rheumatoid arthritis; NLR, neutrophil-to-lymphocyte ratio; PLR, platelet-to-lymphocyte ratio; AUC, area under the curve; HSROC, hierarchical summary ROC; REML, restricted maximum likelihood; HKSJ, Hartung–Knapp–Sidik–Jonkman; QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies-2; QUIPS, Quality in Prognosis Studies; NOS, Newcastle-Ottawa Scale; csDMARDs, conventional synthetic disease-modifying antirheumatic drugs.

Declarations

Ethics approval: This study is a systematic review and meta-analysis of previously published data. Although ethical approval is not typically required for such secondary research, we sought and received an official ethical waiver from the Institutional Review Board of the University of Science and Technology (Waiver Number: UST-CRE/WAIVER/2024/089) to ensure full ethical compliance. All included primary studies declared adherence to ethical standards in accordance with ICMJE guidelines.

Consent for publication: Not applicable.

Data and code availability:

All extracted study-level data, analysis code (R and Python), and supplementary materials are publicly available in the Open Science Framework repository: https://osf.io/prqcm

De-identified data extraction sheets and reproducible scripts can also be obtained directly from the corresponding author (Naif Taleb Ali) upon reasonable request via email: n.taleb@ust.edu.

Competing interests: The authors declare that they have no competing interests. For transparency, we confirm that we assessed the funding sources and conflicts of interest of all studies included in this meta-analysis; no significant bias from these factors was detected in the overall pooled results.

Funding: No external funding was received for this study.

AI Assistance Disclosure

During the preparation of this work,the authors used ChatGPT (OpenAI) for language editing and refinement of specific sections to improve readability. All AI-assisted content was thoroughly reviewed, verified, and edited by the authors who take full responsibility for the final content and scientific accuracy of the publication.

Authors' contributions:

N.T.A.: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Supervision, Project Administration, Funding Acquisition (institutional support). Corresponding Author: Responsible for all communication, research integrity, and final manuscript approval.

M.A.H.M.: Conceptualization, Methodology, Software, Formal Analysis, Validation, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization.

R.S.A.: Investigation, Resources, Data Curation, Writing – Review & Editing, Supervision (local).

G.S.A.: Investigation, Resources, Data Curation, Validation, Writing – Review & Editing.

H.M.A.: Investigation, Resources, Data Curation, Validation, Writing – Review & Editing.

A.N.M.G.: Methodology, Software, Validation, Formal Analysis, Data Curation, Writing – Review & Editing.

N.M.A.: Conceptualization, Methodology, Validation, Resources, Writing – Review & Editing, Supervision.

All authors have read, critically reviewed, and approved the final version of the manuscript. All authors agree to be accountable for all aspects of the work.

Acknowledgements: We acknowledge the technical support from the Research Support Unit at UST-Aden, Radfan College University, Lahej University, and the University of Aden.

References

1. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet*. 2016 Oct 22;388(10055):2023–2038. doi:10.1016/S0140-6736(16)30173-8.
2. Adebajo A, Davis P, Tiku K, Kalla AA, Tikly M, Mody GM. Rheumatoid arthritis in Africa: clinical and genetic insights. *Clin Rheumatol*. 2020 May;39(5):1234–1240. doi:10.1007/s10067-020-05123-2.
3. Ali NT, Ali GS, Mohsen Ali H. NLR outperforms PLR in SLE diagnosis and prognosis: an AI-enhanced meta-analysis of 12 850 patients with ethnicity-specific cut-offs. *Lupus Science & Medicine* 2025;12:e001696. doi:10.1136/lupus-2025-001696
4. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011 Oct 18;155(8):529–536. doi:10.7326/0003-4819-155-8-201110180-00009.
5. Mangoni AA, Zinellu A. Diagnostic accuracy of the neutrophil-to-lymphocyte ratio and the platelet-to-lymphocyte ratio in rheumatoid arthritis: a systematic review and meta-analysis. *Clin Exp Med*. 2024 Sep 4;24(1):207. doi:10.1007/s10238-024-01478-x. PMID:39230596; PMCID:PMC11374877.
6. Liu S, Liu J, Cheng X, Fang D, Chen X, Ding X, Zhang X, Chen Y. Application value of platelet-to-lymphocyte ratio as a novel indicator in rheumatoid arthritis: a review based on clinical evidence. *J Inflamm Res*. 2024;17:7607–7617. doi:10.2147/JIR.S477262.
7. Lee YH. Association between the neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio in rheumatoid arthritis and their correlations with disease activity: a meta-analysis. *J Rheum Dis*. 2018;25(3):169–178. doi:10.4078/jrd.2018.25.3.169.
8. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA*. 2000 Apr 19;283(15):2008–2012. doi:10.1001/jama.283.15.2008.
9. Pope JE, Choy EH. C-reactive protein and implications in rheumatoid arthritis and associated comorbidities. *Semin Arthritis Rheum*. 2021 Feb;51(1):219–229. doi:10.1016/j.semarthrit.2020.11.005.
10. El-Labban AS, Omar HA, El-Shereif RR, Ali F, El-Mansoury TM. Pattern of rheumatoid arthritis in the eastern region of Libya: a prospective study. *Pan Afr Med J*. 2019 Jan 29;32:52. doi:10.11604/pamj.2019.32.52.16764.
11. Figus FA, Piga M, Azzolin I, McConnell R, Iagnocco A. Rheumatoid arthritis: extra-articular manifestations and comorbidities. *Autoimmun Rev*. 2021 Apr;20(4):102776. doi:10.1016/j.autrev.2021.102776.
12. Smolen JS, Landewé RBM, Bijlsma JWJ, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. *Ann Rheum Dis*. 2020;79(6):685–699. doi:10.1136/annrheumdis-2019-216655.
13. Jin Z, Cai G, Zhang P, et al. The value of the neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio as complementary diagnostic tools in the diagnosis of rheumatoid arthritis: a multicenter retrospective study. *J Clin Lab Anal*. 2021;35(1):e23569. doi:10.1002/jcla.23569.
14. Adelowo O, Ojo O, Animashaun B, et al. Rheumatoid arthritis among Nigerians: the first 200 patients from a tertiary hospital. *Clin Rheumatol*. 2010;29:593–597. doi:10.1007/s10067-010-1373-y.
15. Smolen JS, Landewé RBM, Bergstra SA, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2022 update. *Ann Rheum Dis*. 2023;82(1):3–18. doi:10.1136/ard-2022-223356.

16. England BR, Tiong BK, Bergman MJ, Curtis JR, Kazi S, Mikuls TR, O'Dell JR, Ranganath VK, Limanni A, Suter LG, Michaud K. 2019 update of the American College of Rheumatology recommended rheumatoid arthritis disease activity measures. *Arthritis Care Res (Hoboken)*. 2019 Dec;71(12):1540–1555. doi:10.1002/acr.24042.
17. World Health Organization. Yemen: Health Cluster Bulletin, November & December 2023 [Internet]. Geneva: WHO; 2024 Mar 21 [cited 2025 Aug 25]. Available from: https://reliefweb.int/report/yemen/yemen-health-cluster-bulletin-november-december-2023.

Figures

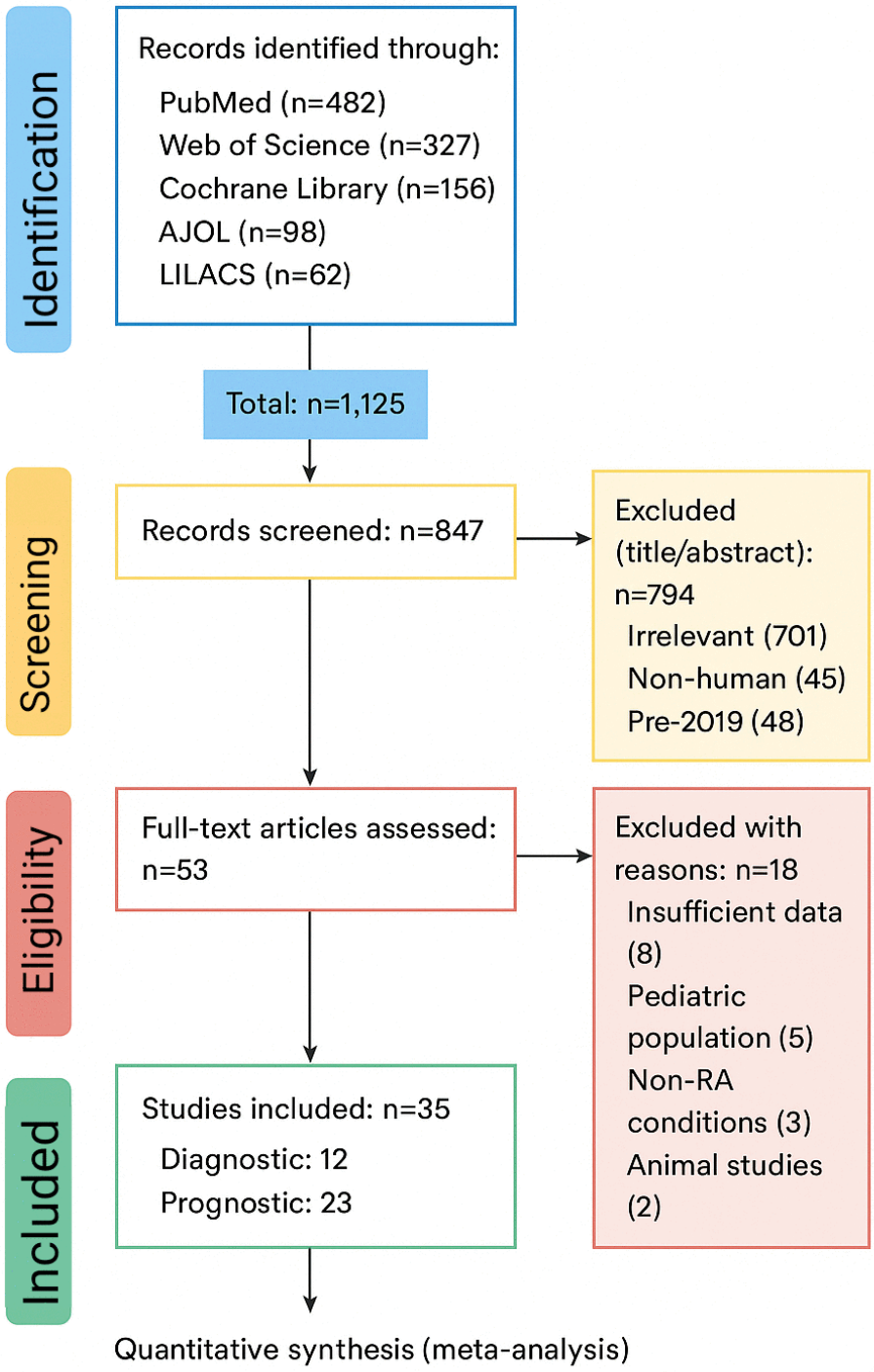


Figure 1

PRISMA 2020 flow diagram

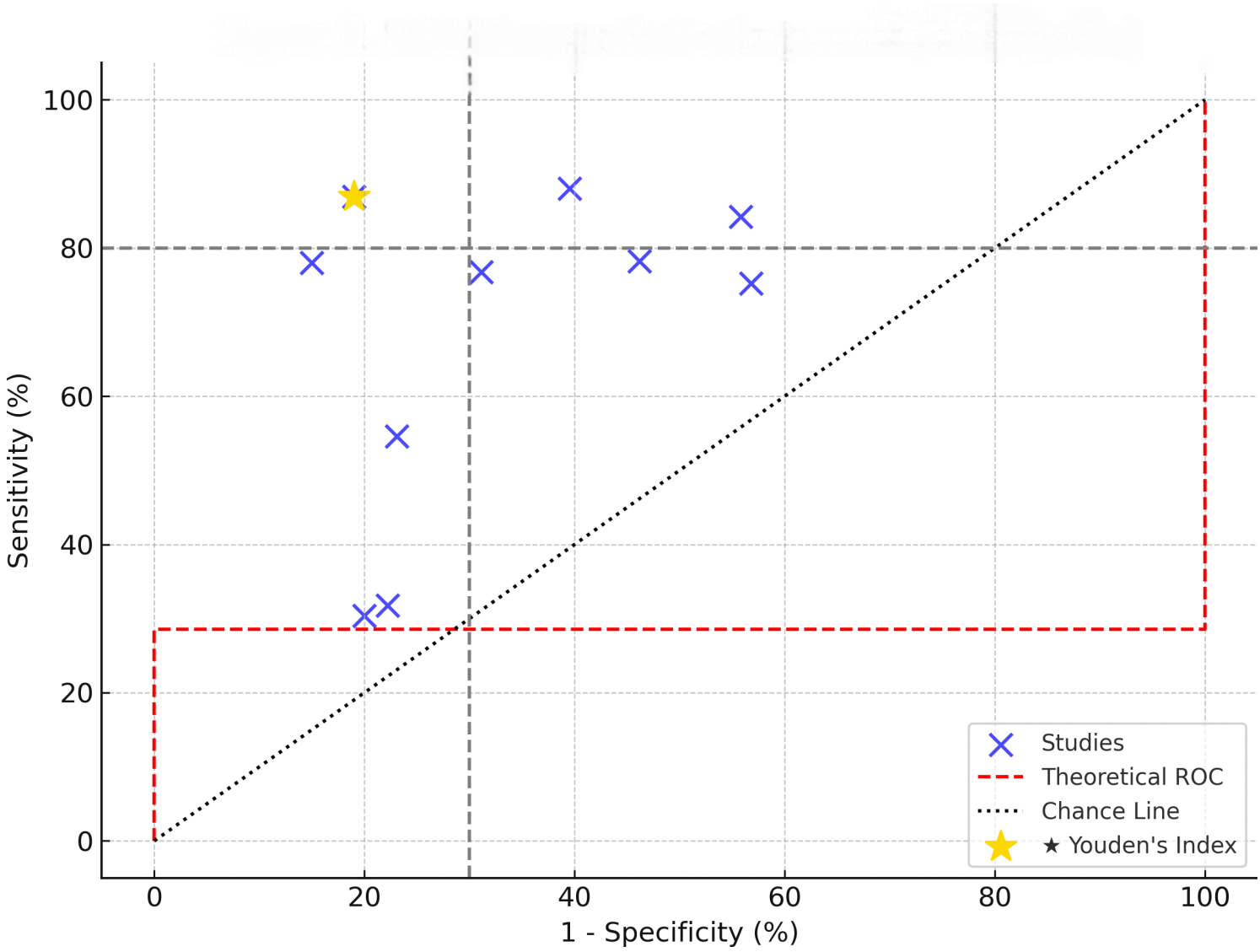


Figure 2

Summary receiver operating characteristic (SROC) curve for NLR diagnostic performance

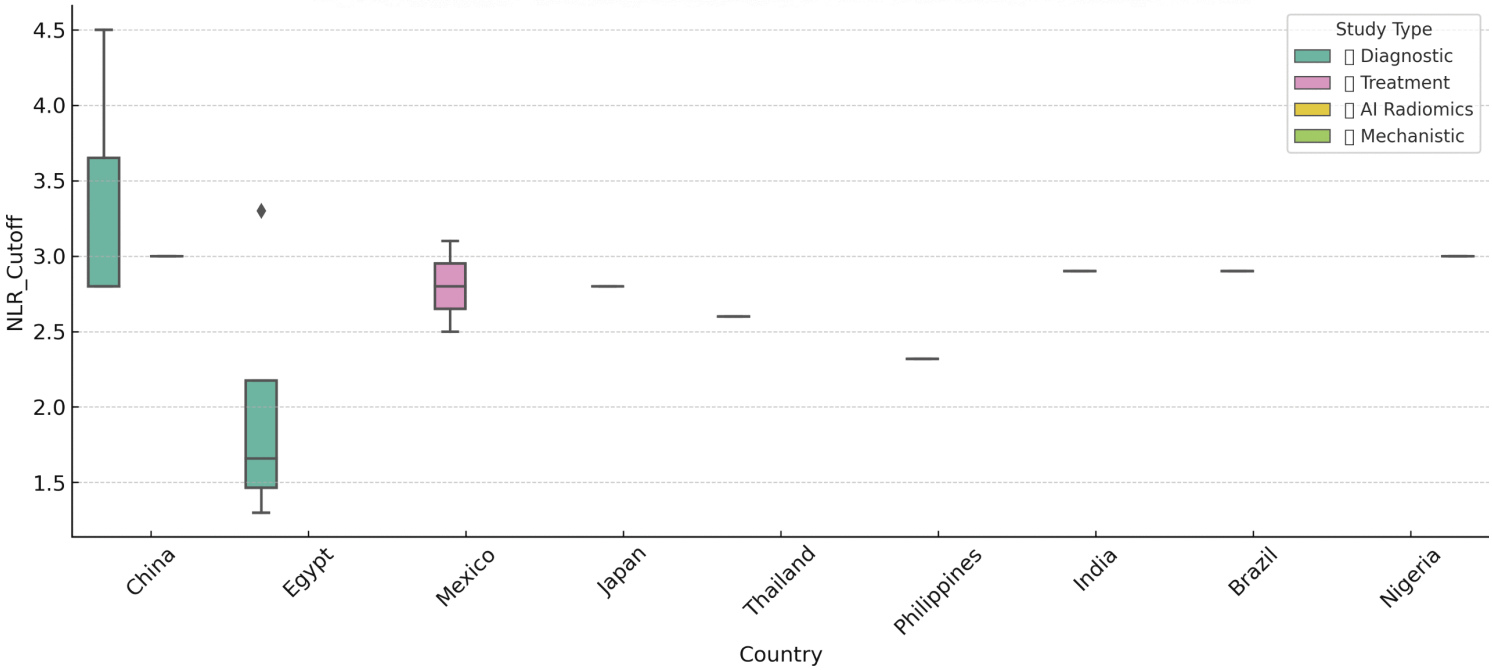


Figure 3

Geographic distribution of included studies and NLR cutoff variation by country

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FileS20ConsolidatedSuppTablesS1S9.docx
- FileS7bStatisticalappendix.docx
- SupplementaryInformationGuide.docx
- FileS4IPDsharingagreement.docx
- FileS20DetailedStudyCharacteristics.docx
- FileS14MathematicalAdjustmentModelYemeniNLR.docx
- FileS1PRISMAMOOSETRIPODCHEERSGRAP2Checklists.docx
- FileS9completePythonscriptformetaanalyses.py
- FileS12characteristicsofincludedstudies.csv
- FileS3sensitivitytableslast.csv
- FileS16DelphiConsensusProtocolforYemeniNLRCutoff.docx
- FileS8completeRscriptformetaanalyses.r
- FileS10StatisticalCodeandFormulaRepository.docx
- FileS15Datacollectionprotocol.docx
- FileS2QualityAssessment.csv

- FileS17costanalysis.csv
- FileS6fulldatasetforanalysis.csv
- FileS7aFundingsourcesandconflict.csv
- FileS18NLRcutoffmethodologycomparison.csv
- FileS11Excludedstudies.csv
- FileS5Fullsearchstrategy.docx
- FileS13Topdiagnosticstudies.csv
- FileS21FiguresS1S12.docx