NIHR | National Institute for Health and Care Research

PROSPERO
International prospective register of systematic reviews

# Artificial Intelligence-Augmented Meta-Analysis of Complete Blood Count (CBC) Markers for Irritable Bowel Syndrome (IBS) Diagnosis: A PRISMA-Compliant Protocol

*Naif Ali, Hana Ali*

**Citation**

Naif Ali, Hana Ali. Artificial Intelligence-Augmented Meta-Analysis of Complete Blood Count (CBC) Markers for Irritable Bowel Syndrome (IBS) Diagnosis: A PRISMA-Compliant Protocol. PROSPERO 2025 CRD420251069483. Available from https://www.crd.york.ac.uk/PROSPERO/view/CRD420251069483.

## REVIEW TITLE AND BASIC DETAILS

**Review title**

Artificial Intelligence-Augmented Meta-Analysis of Complete Blood Count (CBC) Markers for Irritable Bowel Syndrome (IBS) Diagnosis: A PRISMA-Compliant Protocol

**Condition or domain being studied**

*Irritable bowel syndrome; Complete Blood Count; Diagnostic Procedure; Machine learning*

Clarification on Healthcare Condition and Domain This systematic review and meta-analysis focus on Irritable Bowel Syndrome (IBS), a prevalent functional gastrointestinal disorder (FGID) affecting 10-15% of adults globally. IBS is diagnosed clinically using the Rome IV criteria, as no definitive biomarker exists. The study evaluates Complete Blood Count (CBC)-derived markers—specifically neutrophil-to-lymphocyte ratio (NLR), mean platelet volume (MPV), and platelet-to-lymphocyte ratio (PLR)—as potential diagnostic tools. These inexpensive, routinely measured biomarkers may address critical gaps in IBS diagnosis, particularly in resource-limited settings like Yemen, where advanced diagnostics are scarce. The review uniquely integrates artificial intelligence (AI) (XGBoost, Random Forest) to synthesize multi-parameter CBC data, enhancing diagnostic accuracy beyond traditional statistical pooling. By combining meta-analysis with interpretable machine learning (SHAP/LIME), we aim to provide a clinically actionable framework. Domain: This bridges clinical gastroenterology, diagnostic hematology, and AI-driven healthcare innovation. The outcomes could redefine IBS diagnostics by validating CBC markers and demonstrating AI's role in low-resource settings. Note: While PROSPERO currently lacks a dedicated "Diagnostic Test

Accuracy" category, this review's methodology aligns with PRISMA-DTA standards. We seek to establish CBC markers as supplementary tools for IBS subtyping and diagnosis.

### Rationale for the review

This systematic review and meta-analysis addresses the critical unmet need for objective biomarkers in Irritable Bowel Syndrome (IBS) diagnosis, which currently relies on subjective symptom-based criteria (Rome IV). Despite the widespread use of Complete Blood Count (CBC) tests in clinical practice, the diagnostic potential of CBC-derived markers (NLR, MPV, PLR) for IBS remains unvalidated through rigorous evidence synthesis. Our review will: 1. Provide the first PRISMA-compliant meta-analysis of CBC markers across IBS subtypes, quantifying their pooled diagnostic accuracy. 2. Integrate artificial intelligence (XGBoost, SHAP) to model multi-parameter CBC interactions, overcoming limitations of univariate analyses in prior studies. 3. Bridge a key gap in global health equity by validating low-cost CBC biomarkers for resource-limited settings, where advanced diagnostics are inaccessible. This work will shift IBS diagnosis from symptom-based to biomarker-augmented paradigms, offering clinicians an evidence-based, interpretable tool to reduce diagnostic delays and heterogeneity.

### Review objectives

1. To systematically evaluate the diagnostic accuracy of CBC-derived biomarkers (NLR, MPV, PLR) in differentiating IBS patients from healthy controls, using Rome IV criteria as the reference standard. 2. To assess variations in biomarker performance across IBS subtypes (IBS-D, IBS-C, IBS-M). 3. To develop and validate interpretable machine learning models integrating multiple CBC parameters for improved IBS diagnosis. 4. To establish pooled effect sizes (SMD) for NLR, MPV, and PLR through meta-analysis, with subgroup analyses by geographic region and measurement methods. This study aims to transform IBS diagnosis by validating accessible CBC biomarkers and demonstrating AI's role in enhancing clinical decision-making, particularly in resource-limited settings.


### Keywords

Irritable bowel syndrome; Biomarkers; Machine learning; Complete blood count; Diagnostic accuracy


### Country

Yemen; United Kingdom; United States of America; Saudi Arabia

## ELIGIBILITY CRITERIA

### Population

*Included*

Adults (≥18 years) diagnosed with Irritable Bowel Syndrome (IBS) according to Rome IV criteria, regardless of subtype (IBS-D, IBS-C, IBS-M). Studies must report at least two CBC-derived biomarkers (NLR, MPV, or PLR) with measurable outcomes (means/standard deviations or diagnostic accuracy metrics). Only case-control or cohort studies published in English will be included. .

*Excluded*
Excluded: - Pediatric populations (<18 years) - Studies of organic GI diseases (e.g., IBD, celiac disease) - Case reports, reviews, or editorials - Non-English publications - Studies with incomplete statistical data (e.g., missing SDs, unclear diagnostic thresholds)

### Intervention(s) or exposure(s)

*Included*
Studies measuring CBC parameters (NLR, MPV, PLR) as diagnostic exposures in IBS patients. Includes both observational (routine blood tests) and experimental (protocol-standardized CBC) measurements.

*Excluded*
Studies using modified/non-standard CBC techniques - Pharmacological/surgical interventions affecting CBC values - Non-blood-based biomarkers (e.g., fecal, genetic)

### Comparator(s) or control(s)

*Included*
*PICO tags selected: Active control*

Comparator groups must be clearly defined as either:- Healthy controls without GI symptoms, OR- Patients with non-IBS GI conditions (e.g., IBD, functional dyspepsia) confirmed by appropriate diagnostic testing.

*Excluded*
Studies without clearly defined control groups- Controls with systemic conditions affecting CBC (e.g., cancer, chronic infections)- Studies where control group received interventions altering CBC parameters

### Study design
Only nonrandomized study types will be included.

*Included*
Case-control and cohort studies meeting Rome IV criteria for IBS diagnosis, reporting ≥2 CBC parameters with means/SD.

*Excluded*
Reviews, case reports, editorials, studies with organic GI disease diagnoses, non-English publications, or incomplete statistical reporting.

### Context
This meta-analysis focuses on adult populations (≥18 years) diagnosed with Irritable Bowel Syndrome (IBS) according to Rome IV criteria, across diverse clinical and community settings globally. Special emphasis is placed on studies reporting Complete Blood Count (CBC) parameters (e.g., NLR, MPV, PLR) with standardized measurements. While no geographic restrictions apply, the study aims to address gaps in low-resource settings (e.g., Yemen) through transfer learning in AI modeling. Excluded are pediatric populations, organic GI disease cases, and studies lacking quantitative CBC data or diagnostic clarity.

## TIMELINE OF THE REVIEW

**Date of first submission to PROSPERO**
07 June 2025

**Review timeline**
Start date: 30 June 2025. End date: 31 December 2025.

**Date of registration in PROSPERO**
07 June 2025

## AVAILABILITY OF FULL PROTOCOL

**Availability of full protocol**
A full protocol has been written but is not available because:

*Protocol is undergoing institutional review prior to journal submission.*

## SEARCHING AND SCREENING

**Search for unpublished studies**
Only unpublished studies will be sought.

**Main bibliographic databases that will be searched**
The main databases to be searched are *CLIB - The Cochrane Library*, *Embase - Embase via Ovid*, *PubMed*, *SSCI - Social Science Citation Index* and *Scopus*.

**Search language restrictions**
The review will only include studies published in English.

**Search date restrictions**
There are no search date restrictions.

**Other methods of identifying studies**
No other methods will be used.

**Link to search strategy**
A full search strategy is available in the full protocol as described in the *Availability of full protocol* section

**Selection process**
Studies will be screened independently by at least two people (or person/machine combination) with a process to resolve differences.

**Other relevant information about searching and screening**
AI-assisted screening (spaCy NLP) will prioritize studies reporting CBC parameters. Dual independent reviewers will resolve conflicts via consensus, referencing predefined Rome IV criteria.

Non-English studies are excluded per protocol. Search alerts will monitor new publications until manuscript submission to ensure comprehensive evidence coverage.

## DATA COLLECTION PROCESS

### Data extraction from published articles and reports

Data will be extracted independently by at least two people (or person/machine combination) with a process to resolve differences.

Authors will not be contacted for further information.

### Study risk of bias or quality assessment

Risk of bias will be assessed using: *Newcastle-Ottawa*

Data will be assessed independently by at least two people (or person/machine combination) with a process to resolve differences.

Additional information will **not** be sought from study investigators if required information is unclear or unavailable in the study publications/reports.

### Reporting bias assessment

Funnel plots and Egger's test will assess publication bias. Sensitivity analyses will exclude studies with incomplete CBC data. Contour-enhanced funnel plots differentiate asymmetry due to bias vs. chance. Small-study effects will be quantified using trim-and-fill analysis if needed.

### Certainty assessment

Certainty of findings will not be assessed

## OUTCOMES TO BE ANALYSED

### Main outcomes

1. Neutrophil-to-Lymphocyte Ratio (NLR) - Definition: Ratio of absolute neutrophil count to absolute lymphocyte count. - Measurement: Standard CBC-derived values (reported as mean ± SD or median with IQR). - Time Point: Baseline measurements at diagnosis. - Effect Measure: Standardized mean difference (SMD) with 95% CI. 2. Mean Platelet Volume (MPV) - Definition: Average size of platelets measured in femtoliters (fL). - Measurement: Automated hematology analyzers (reported as mean ± SD). - Time Point: Baseline measurements at diagnosis. - Effect Measure: Standardized mean difference (SMD) with 95% CI. 3. Platelet-to-Lymphocyte Ratio (PLR) - Definition: Ratio of absolute platelet count to absolute lymphocyte count. - Measurement: CBC-derived values (reported as mean ± SD or median with IQR). - Time Point: Baseline measurements at diagnosis. - Effect Measure: Standardized mean difference (SMD) with 95% CI. 4. AI Model Diagnostic Performance - Definition: Predictive accuracy of machine learning models integrating CBC parameters. - Measurement: AUC-ROC, sensitivity, specificity, F1-score. - Time Point: Post-model validation. - Effect Measure: Pooled performance metrics with 95% CI. Note: All outcomes will be stratified by IBS subtype (IBS-D, IBS-C, IBS-M) and geographic region where data permit.

### Additional outcomes

1. White Blood Cell Count (WBC) - Definition: Total leukocyte count per microliter (μL) of blood. - Measurement: Standard CBC analysis (reported as mean ± SD or median with IQR). - Time Point: Baseline measurements at diagnosis. - Effect Measure: Standardized mean difference (SMD) with 95% CI. 2. Hemoglobin (Hb) Levels - Definition: Concentration of hemoglobin in grams per deciliter (g/dL). - Measurement: CBC-derived values (reported as mean ± SD). - Time Point: Baseline measurements at diagnosis. - Effect Measure: Mean difference (MD) with 95% CI. 3. C-reactive Protein (CRP) Levels (if reported in included studies) - Definition: Inflammatory marker measured in milligrams per liter (mg/L). - Measurement: High-sensitivity assays (reported as mean ± SD or median with IQR). - Time Point: Baseline measurements at diagnosis. - Effect Measure: Standardized mean difference (SMD) with 95% CI. 4. Subgroup-Specific Variability - Definition: Variability in CBC parameters across demographic subgroups (age, sex, BMI). - Measurement: Stratified data from included studies. - Time Point: Baseline measurements. - Effect Measure: Interaction p-values and subgroup SMDs. 5. Correlation with Symptom Severity (if reported) - Definition: Association between CBC markers and IBS symptom scores (e.g., IBS-SSS). - Measurement: Correlation coefficients or regression coefficients. - Time Point: As reported in studies. - Effect Measure: Pooled correlation coefficients (Fisher's Z-transformed). Note: These outcomes will be analyzed if sufficient data are available, though they are not primary to the study's objectives.

## PLANNED DATA SYNTHESIS

### Strategy for data synthesis

1. Effect Size Calculation: - For continuous outcomes (NLR, MPV, PLR, WBC, Hb), we will calculate standardized mean differences (SMDs) with 95% confidence intervals to account for variability in measurement scales across studies. - For diagnostic accuracy metrics (AUC-ROC, sensitivity), we will use bivariate random-effects models when sufficient studies are available, or univariate pooling if data is limited.2. Statistical Models: - Primary analysis will use random-effects models (DerSimonian-Laird method) to accommodate expected clinical and methodological heterogeneity. - Fixed-effects models will be employed as sensitivity analysis when $I^2 < 50\%$.3. Handling of Missing Data: - Where standard deviations are missing, we will: Contact study authors for original data Estimate SDs from p-values/confidence intervals if available Impute using median SDs from other included studies as last resort4. Subgroup Synthesis: - Stratified analyses will be conducted by: IBS subtypes (IBS-D, IBS-C, IBS-M) Geographic regions Measurement techniques (automated vs manual CBC analysis) - Between-subgroup differences will be tested using meta-regression5. Special Cases: - For studies reporting medians/IQR, we will use Wan et al. (2014) methods to estimate means/SDs - Skewed data will be log-transformed before pooling - Multiple arms from single studies will be properly weighted to avoid unit-of-analysis errors6. Software Implementation: - Primary synthesis in R using metafor and metamisc packages - Sensitivity analyses in RevMan 5.4 - All code will be made publicly availableValidation Measures:- Leave-one-out sensitivity analysis to assess robustness- Cumulative meta-analysis by publication year to detect temporal trends- Reporting bias assessment via contour-enhanced funnel plots and Egger's test

## CURRENT REVIEW STAGE

### Stage of the review at this submission

| Review stage | Started | Completed |
|---|---|---|
| Pilot work | | |
| Formal searching/study identification | | |
| Screening search results against inclusion criteria | | |
| Data extraction or receipt of IPD | | |
| Risk of bias/quality assessment | | |
| Data synthesis | | |

### Review status
The review is currently planned or ongoing.

### Publication of review results
Results of the review will be published in English and Arabic.

## REVIEW AFFILIATION, FUNDING AND PEER REVIEW

### Review team members
**Dr Naif Ali** (review guarantor and contact) ORCID: 0000-0002-2266-1569. University of Science and technology. Yemen.

No conflict of interest declared.

**Dr Hana Ali**. ORCID: 0009-0005-6144-4769. Aden Gulf International University. Yemen.

No conflict of interest declared.

### Named contact
**Dr Naif Ali** (n.taleb@ust.edu). ORCID: 0000-0002-2266-1569. University of Science and technology. Yemen.

### Review affiliation
University of Science and Technology, Aden, Yemen

### Funding source
Review has no specific/external funding but is supported by guarantor/review team (non-commercial) institutions.

### Peer review
There has been no peer review of this planned review.

## ADDITIONAL INFORMATION

### Additional information

1. Regional Focus: - While synthesizing global evidence, this review includes a unique implementation framework for low-resource settings (Yemen), addressing a gap in IBS diagnostics (Section 1.3). 2. AI Model Transparency: - All machine learning code and meta-analysis datasets will be made publicly available upon publication to ensure reproducibility (GitHub repository linked in final manuscript). 3. Ethical Compliance: - Secondary data analysis only (no patient contact), with IRB approval from University of Sciences and Technology, Yemen (Section 8). 4. Protocol Updates: - Any amendments post-PROSPERO registration will be documented and dated in the final publication.

### Review conflict of interest

Declared individual interests are recorded under team member details.. No additional interests are recorded for this review.

### Medical Subject Headings

Irritable Bowel Syndrome; Blood Cell Count; Neutrophils; B-Lymphocyte Subsets; Mean Platelet Volume; Artificial Intelligence; Machine Learning; Diagnostic Techniques and Procedures; Digestive System; Biomarkers; Yemen

## SIMILAR REVIEWS

### Check for similar records already in PROSPERO

*PROSPERO identified a number of existing PROSPERO records that were similar to this one (last check made on 7 June 2025). These are shown below along with the reasons given by that the review team for the reviews being different and/or proceeding.*

- Exploring the Relationship Between Psychological Factors and Subtypes of Irritable Bowel Syndrome (IBS): A Systematic Review and Meta-Analysis [published 30 March 2024] [CRD42024526397]. The review was judged **not to be similar**
- The awareness on microscopic colitis and irritable bowel syndrome overlap: a systematic review [published 13 October 2014] [CRD42014014195]. The review was judged **not to be similar**
- Comparative Efficacy of Different Cognitive Behavioral Therapies and Acupuncture for Irritable Bowel Syndrome: A Network Meta-Analysis [published 17 October 2024] [CRD42024598351]. The review was judged **not to be similar**
- Peripheral biomarkers of the irritable bowel syndrome: A systematic review and meta-analysis [published 13 May 2024] [CRD42024502876]. The review was judged **not to be similar**
- Diagnostic Potential of various laboratory tests for Irritable Bowel Syndrome: A Systematic Review [published 1 July 2023] [CRD42023438164]. The review was judged **not to be similar**
- Systematic Review on the Effects of Various Dietary Interventions on Symptoms of Irritable Bowel Syndrome [published 18 July 2024] [CRD42024565979]. The review was judged **not to be similar**
- Systematic Review and Meta-Analysis of Cytokine Imbalance in Diarrhea-Predominant Irritable Bowel Syndrome [published 21 June 2024] [CRD42024556941]. The review was judged **not to**

**be similar**

- Rifaximin versus Low-FODMAP Diet for Bloating in Irritable Bowel Syndrome: A Systematic Review and Network Meta-Analysis [published 23 April 2025] [CRD420251036670]. The review was judged **not to be similar**
- Prevalence of Eating Disorders Among Individuals With Irritable Bowel Syndrome: A Systematic Review and Meta-analysis [published 24 February 2025] [CRD420250652312]. The review was judged **not to be similar**

## PROSPERO version history

- Version 1.0, published 07 Jun 2025

## Disclaimer