

# Machine learning-based prediction of inflammation adjusted iron deficiency anaemia using blood cell indices

Ravindranadh Palika,<sup>1</sup> Teena Dasi,<sup>1</sup> Devraj J.Parasannanavar,<sup>1</sup> Santosh Kumar Banjara,<sup>1</sup> Bharati Kulkarni<sup>1</sup> & Raghu Pullakhandam<sup>1</sup>

<sup>1</sup>ICMR-National Institute of Nutrition, Hyderabad, Telangana, India

Received June 11, 2025; Accepted December 21, 2025; Published April 27, 2026

**Background and objectives:** Diagnosis of iron deficiency anaemia (IDA), where both anaemia and iron deficiency coexist, is essential to improve the precision of iron treatment. The serum ferritin, a marker of iron deficiency, is an acute phase protein, requires adjustment for inflammation prior to using it for diagnosis of IDA. The blood cell indices (referred as complete blood counts, CBC) help in identifying the underlying cause of anaemia and presence of inflammation, both of which could be utilised for differential diagnosis of IDA. We conducted this study to evaluate machine learning (ML) models for predicting inflammation-adjusted IDA from CBC data.

**Methods:** We utilised CBC data from National Health and Nutrition Examination Survey (NHANES 2017-2023) of women of reproductive age (15-50 yr age, n=3604) and evaluated the performance of ML models to predict the inflammation adjusted IDA, in terms of sensitivity and specificity. We validated the performance of the optimised random forest (RF)- ML model, on women of reproductive age data set from India (n=381) in predicting the IDA, and haemoglobin response to iron treatment.

**Results:** The optimised RF model predicted the inflammation adjusted IDA, solely based on CBC, with sensitivity and specificity of ~95% either on the test or validation data. In those predicted to have IDA ( $2.3 \pm 1.6$  g/dL, n=283), the haemoglobin increments due to iron therapy was markedly higher compared to those predicted to have non-IDA ( $0.31 \pm 0.67$  g/dL, n=98).

**Interpretation and conclusions:** These findings suggest that it is possible to predict inflammation adjusted IDA with high sensitivity and specificity with ML models using CBC data.

**Keywords** Anaemia; Blood cell indices; Complete blood counts; Iron deficiency; Machine learning; Random forest; Serum ferritin; Women of reproductive age