OPEN ACCESS Journal of Interdisciplinary Research in Allied Health and Pharmacy
ISSN (Print): 3106-020X (Online): 3106-0218
Journal homepage: https://jirap.ubas.edu.pk/index.php/jirap/index

ORIGNAL ARTICLE

# Revolutionizing Thalassaemia Detection: The Emerging Role of Artificial Intelligence and Machine Learning

**Syeda Nayab Bukhari[1], Hafsa[1], Nishat Arshad[1], Alvia Batool[1], Ayesha Aftab[1]**

[1]Institute of Chemistry, University of Sargodha, Sargodha, Punjab, Pakistan.

**Correspondence**
Syeda Nayab Bukhari
Email: nayab4376@gmail.com

**Conflict of Interest**
All the authors have no conflict of interest

**Reference**
Rashid.F., Riaz.A., Riaz.A., (2025 Revolutionizing Thalassaemia Detection: The Emerging Role of Artificial Intelligence and Machine Learning, Journal of Interdisciplinary Research in Allied Health Sciences.

**Published**: 31st December 2025

**Abstract**

Thalassaemia is a serious genetic disorder that causes severe anaemia, resulting in serious health and economic problems on the entire globe, particularly in areas with restricted access to health care. Thalassaemia and iron deficiency anaemia (IDA) are clinically similar because of overlapping haematological characteristics. Conventional diagnostic tools comprising of molecular test, conventional index and complete blood count (CBC) parameters may provide pertinent information but are not always available, are not affordable and take a long turnaround time. Over the last few decades, the field of medicine diagnostics has encountered a disruptive technology in the form of Artificial Intelligence (AI) since it is both cheap and quick and incredibly precise in treatment. The AI like Decision Tree, Support Vector Machine, and Neural Networks can be able to determine the traits of thalassaemia using the normal haematological data. Additional methods proven to be more credible in diagnosis, like deep learning and sophisticated algorithms, XGBoost and Convolutional Neural Networks (CNN) can be further tested to introduce additional diagnostic reliability. It has also been found in comparative studies that AI-based models tend to be more sensitive and specific than standard indices, that's why these technologies are currently being utilized as genetic counseling and screening tools. Despite the problems of data quality, bias of the model, and considerations of ethics, AI can serve as an important supportive resource in the early diagnosis, accurate differentiation, and better management of thalassaemia in low resource healthcare facilities. The combination of AI and molecular testing is promising for global thalassaemia control and personalized medication. The review provides a unique, clinically oriented, synthesis of AI-based thalassaemia diagnostics through the combination of performance comparison, methodological criticism, and an implementation framework suggested in low-resource healthcare environments.

**Keywords:** Thalassaemia, iron deficiency anaemia, artificial intelligence, machine learning, β globin chain

## 1. Introduction

Thalassaemia is a broad group of autosomal recessive genetic diseases. In 2021, the worldwide number of thalassaemia cases was 1,310,407. The number of deaths remained at 11,087, with an age-standardized mortality rates (ASMR) of 0.15 per 100,000 persons [1] . It inhibits the synthesis of a single or more haemoglobin globin chains and is the most widespread monogenic disorder in the world with an estimated population of 1-5 percent of the world population having a hereditary thalassaemia mutation [2]. The disease is specifically prevalent in areas of Africa, Middle East, Mediterranean region and Southeast as well as South Asia. Migration has, however, made it widespread in North America and Northern Europe. This expanding global distribution has created significant inequities in access to screening, diagnostic, and management services, while simultaneously placing new pressures on already strained healthcare systems. [3].

Thalassaemia is caused by mutation which affects production of globin chain. It leads to the imbalanced production of the α and β globin chains resulting in chronic anaemia and poor erythropoiesis (red blood cell breakdown). More than 200 different mutations have been identified in the 11HBB globin gene on

chromosome 11. Most of these point mutations involve alteration of a single nucleotide. The clinical manifestations may vary tremendously in patients with high-risk transfusion-dependent anaemia up to an asymptomatic carrier state, depending on the type and the extent of genetic defect.

The two primary forms of thalassaemia include α-thalassaemia and β-thalassaemia resulting from insufficient or minimal production of alpha and beta-globin chains, respectively. It requires lifelong transfusion therapy and presents ranging from moderate hypochromic microcytosis to severe anaemia. Untreated cases of β-Thalassaemia major, the most serious condition typically diagnosed in childhood, may result in growth retardation, skeletal defects, and organ dysfunction caused by iron overload and chronic hypoxia. Resulting from development of iron chelation therapy and regular transfusion programs, thalassaemia major no longer presents an fatal childhood disease but a manageable chronic disease [4].

However, patients are susceptible to liver disease, osteoporosis, endocrine disorders, and heart failure. Gene therapy emerged as a promising field of therapy, while treatment regimens including allogeneic hematopoietic stem cell transplantation (HSCT) are applied in specific patients.

Prevention still remains the best course of action despite these medical advancements. Prenatal diagnosis, carrier identification, and premarital screening are essential for reducing prevalence of the disease [5]. However, there are insufficient screening and molecular diagnostic facilities in many highly affected areas.

Thalassaemia diagnosis is done using a combination of haematological, biochemical, and molecular testing. The initial screening step mostly involves a Complete Blood Count (CBC). Thalassaemia carriers have small, pale red blood cells with reduced haemoglobin content. This is reflected by low Mean Corpuscular Volume (MCV) values of approximately 60–70 fL and low Mean Corpuscular Haemoglobin (MCH) values of around 19–23 pg. Quantitative analysis of haemoglobin fractions is possible by performing haemoglobin electrophoresis and High-Performance Liquid Chromatography (HPLC) [6].

High HbF and HbA2 are used to identify β-thalassaemia carriers. As the two conditions present with small and pale red blood cells, serum ferritin testing helps in the exclusion of iron deficiency anaemia (IDA), which is a major differential diagnosis. Molecular confirmation is achieved by PCR based testing, such as ARMS-PCR, allele-specific PCR, and sequencing to identify harmful mutations.Molecular testing is considered the gold standard, but it is expensive, time-consuming, and is often not available in areas with limited resources [7]. This causes a treatment dilemma as although the two conditions have similar blood profiles, there is a significant difference in the treatment of IDA and thalassaemia. Misdiagnosis may lead to the probability of severe illness in children because a carrier may get unnecessary iron treatment or miss out the treatment of thalassaemia. Over the past years, medical diagnosis has been transformed by AI, which is able to learn and identify hidden patterns using big data [8]. In the field of haemotology, AI can distinguish thalassaemia and iron deficiency anaemia by analyzing routine CBC data without necessarily using special molecular analysis. Artificial Neural Networks (ANN), k-Nearest Neighbors (k-NN), and Decision Trees are examples of machine learning (ML) models that have been trained to identify subtle variations in the properties of red blood cells that cannot be seen by the humans [9].

An example can be the sensitivity of AI-generated diagnostic indices such as the Matos and Carvalho Index, which have shown up to 99% sensitivity when compared to several more traditional formulaic indices such as the Mentzer index or the Green and King Index. Neural networks are the most accurate machine learning method owing to their capability to reproduce complex, nonlinear interactions between two or more haematological parameters at the same time [10]. The benefits of the AI implementation in healthcare practice are as follow: it provides the opportunity to apply latest CBC data in order to provide non-invasive, fast, and inexpensive screening. In broader terms, it eliminate the cost of high follow-up genetic testing. It helps the medical practitioners to prioritize patients who require further testing. AI is a useful, pre-screening tool that enhances the efficiency of laboratory work and improves access to early detection, particularly in low-resource settings, instead of molecular diagnostics replacement [11].

## 2. Conventional Diagnostic Approaches

### 2.1 Haematological parameters (CBC, MCV, MCH, RDW)

Anaemia and thalassaemia are initially diagnosed with simple haematological tests including Complete Blood Count (CBC). Mean corpuscular volume (MCV), which gauges the typical size of red blood cells, is one of the key haematological parameters used in screening as MCV readings are usually low in thalassaemia and iron deficiency anaemia (IDA),

which is indicative of microcytosis. The average quantity of haemoglobin per red cell is indicated by mean corpuscular haemoglobin (MCH) which is reduced in thalassaemia and IDA, indicating hypochromia [12]. Red blood cell size variation is measured by the Red Cell Distribution Width (RDW). Because red blood cells change size as new cells are formed in iron-deficient environments, RDW is typically high in IDA. In thalassaemia, RDW is typically normal or slightly raised because of consistently small cells. Since the microcytic, hypochromic profiles of thalassaemia trait and IDA are identical, these parameters offer a first hint but are insufficient to distinguish between the two disorders [13].

### 2.2 Traditional indices (Mentzer, Shine & Lal, Green & King, England & Fraser, etc.)

Based on CBC characteristics, a number of mathematical diagnostic indices have been established to distinguish between IDA and thalassaemia. These indices integrate two or more haematological variables using straightforward formulas

**Table 1: Commonly used indices**

| Sr# | Index Name | Formula | Diagnostic Principle |
|---|---|---|---|
| **01** | **Mentzer Index** | MCV / RBC | < 13 → β-Thalassaemia trait; > 13 → Iron Deficiency anaemia (IDA) |
| **02** | **Shine & Lal Index** | $(MCV^2 \times MCH) / 100$ | Lower values suggest Thalassaemia; higher values suggest IDA |
| **03** | **Green & King Index** | $(MCV^2 \times RDW) / (Hb \times 100)$ | Higher values indicate IDA; lower values indicate Thalassaemia |
| **04** | **England & Fraser Index** | MCV – RBC – (5 × Hb) – 3.4 | Negative value → Thalassaemia; Positive value → IDA |

Although these indices are inexpensive and easy to calculate, their accuracy varies across populations because of differences in genetics, nutritional status, and reference ranges. Thus, while they serve as useful screening tools, they cannot replace confirmatory testing.

### 2.3 Genetic and molecular testing (gold standard, limitations in cost and access)

The most reliable method for diagnosing thalassaemia is still molecular diagnosis. Abnormal haemoglobin fractions can be found using methods like high-performance liquid chromatography (HPLC) and haemoglobin electrophoresis. Certain gene deletions or mutations in the HBA (α-globin) or HBB (β-globin) genes can be found using PCR-based techniques [16]. These tests do have certain limits, though. They need certain equipment and are costly. In areas with limited resources, availability is restricted. Their long turnaround times make them impractical for extensive screening. Consequently, alternative diagnostic methods, including models that rely on artificial intelligence (AI), which are capable of analysing regular blood samples and offer quick and low-cost screening prior to molecular confirmation are gaining popuarity.

## 3. Artificial Intelligence in Medical Diagnosis

### 3.1 Concept of AI and Machine Learning in Medicine

Artificial intelligence (AI) is the set of computational systems that are capable of performing tasks such as pattern recognition and decision-making that would otherwise require human intellect [17]. Machine learning (ML), which is a vital aspect of AI, allows algorithms for training by discovering patterns in data without human instructions and improve their performance. To offer precise diagnostic data, AI and ML are applied to the field of medicine to process complicated data, including laboratory outcomes and medical imaging. AI has a potential to identify genetic defects that are associated with hematologic diseases, including thalassaemia [20]. This approach is particularly advantageous in regions with a high prevalence of such infections where testing access is limited . Neural networks is a common machine learning method that processes several data inputs based on a simulated structure of biological neurons [22]. Ultimately, AI enhances the accuracy of medical diagnoses by integrating a large number of data sources and reducing the subjectivity of their interpretation.

### 3.2 How AI “Learns” from Blood Test Data

Annotated blood test datasets, where the outcome of tests, including the presence of thalassaemia, is confirmed genetically, are initially fed into AI models. haemoglobin (Hb), mean corpuscular volume (MCV), and red blood cell count (RBC) are examples of input features used by supervised

learning systems to predict diagnostic labels [23]. For example, thalassaemia carriers are frequently identified by low MCV and increased RBC, which models learn through recurrent training. Class imbalances are addressed by methods such as the Synthetic Minority Oversampling Technique (SMOTE), which creates samples for underrepresented groups. In order to reduce prediction errors during training, models adjust parameters using boosting in ensemble methods or back propagation in neural networks, while cross-validation techniques like 10-fold splits, assess generalizability by testing on unseen data sections. Prior to feature extraction and classification in image-based applications, pre processing divides electrophoresis lanes [24]. This learning process make models achieve accuracy, which often exceeds 95 percent with validated cohorts.

### 3.3 Benefits of AI In Compare and Contrast to Traditional Indices

Because of similarity, even traditional indicators such as Mentzer index (MCV/RBC), which relies on simple thresholds, often fail to discriminate between thalassaemia and iron deficiency anaemia. An example of this is the use of AI models, which integrate multiple parameters simultaneously with greater accuracy of 99.5 as compared to 93% with traditional methods. They are cost-effective and do not require any additional equipment; hence, a given environment may readily handle them since it only requires a few seconds to process data and provide findings. In normalcy cases, where traditional measures such as Green and King are not doing well, AI reduces false negative results. As an example, age- and red cell distribution width (RDW)-added models are more specific than those using mean cellular haemoglobin, by up to 22% . Interpretability tools, like SHAP values, are used to build clinical trust, as it offers an explanation of feature contributions. The AI helps with early intervention, screening carriers with the number of misdiagnosis decreased by 15-20% with the large population [25]. Finally, cross-center studies indicate that the capacity of AI to be able to adjust to a diversity of individuals works better than rigid classical formulas.

## 4. Artificial Intelligence in Detection of Thalassaemia

### 4.1 Machine Learning Models

Machine learning models, trained on thousands of datasets, can classify thalassaemia subtypes based on using blood parameters or pictures and achieve sensitivity of over 90%. The evaluation of these models is through metrics such as area under the curve (AUC) and is divided into rules based on deep learning models. Large, balanced datasets achieve better performance, and imbalances are dealt with by preprocessing.

In order to classify thalassaemia, **decision trees** create hierarchical rules by dividing data according to thresholds such as MCV <80 fL. By exposing important characteristics like RBC count, they provide clearity. In one investigation, a C4.5 decision tree was able to distinguish between beta-Thalassaemia minor and iron insufficiency with 99.1% accuracy . Similar to ensemble approaches, boosted variants achieve 96% carrier detection. One of the drawbacks is overfitting, which can be decreased via trimming.

Assuming independence among features such as Hb and MCV, **Naïve Bayes classifiers** employed probabilistic methods. For quick screening, these are excellent at managing partial data. For carriers of beta-thalassaemia, reported accuracy is 96%. In noisy datasets, this model performs better than trees, but if assumptions are incorrect, it could perform worse.

In **K-Nearest Neighbors** method, labels are given based on how close they are to k comparable instances in feature space, usually 5–10 . It effectively catches nonlinear patterns in short datasets, with a 98% sensitivity for alpha-thalassaemia characteristics. The choice of distance metrics, such as Euclidean, affects performance for CBC values. Implementation is simplified without a formal training step, but computation time increases as data size increases [25].

**Support Vector Machines** utilize hyperplanes to distinguish between classes in high-dimensional space, minimizing the risk of overfitting. With 99% specificity, electrophoresis image analysis is accurately done using this approach [26]. Kernel functions facilitate nonlinear classification, achieving 95% accuracy in mixed anaemias . With huge imbalance, this approach is suitable for varied datasets but faces scalability issues with volume.

**Artificial Neural Networks (ANNs)** comprise interconnected nodes processing inputs through layers to capture complex patterns. Deep variants like convolutional neural networks (CNNs) classify electrophoresis images with 95.8% accuracy. A deep neural network using RBC indices reached 89.7% overall. Ensemble boosts like XGBoost achieve 99.3% on private alpha-thalassaemia sets. CatBoost models hit 85% for genotyping, with SHAP highlighting MCV importance. ANN performance is further improved by unique framework that uses

medical imaging to diagnose thalassaemia, which achieves 98% accuracy in automated systems. These require large data but excel in nonlinear tasks [27].

**Table 2: Performance Metrics of Machine Learning Models for thalassaemia Detection**

| Model Type | Key Features | Accuracy (%) | Sensitivity (%) | Specificity (%) | AUC | Dataset Size | Citation |
|---|---|---|---|---|---|---|---|
| **Decision Trees** | MCV, RBC, RDW | 99.1 | 99.1 | 99.4 | 0.992 | 396 | (18) |
| **Naïve Bayes** | Hb, MCV, MCH | 96 | 98.4 | 98.7 | N/A | Various | (19) |
| **K- Nearest Neighbors** | RBC, RDW, Hb | 98 | 98.4 | 98.7 | N/A | Mixed | (20) |
| **Support Vector Machines** | MCV, MCH, Images | 95-99 | 83 | 95 | N/A | 200 Images | (20) |
| **ANNs(CNN, XGBoost, CatBoost)** | All CBC, Age, Sex | 85-99.3 | 85-99.3 | 81-99.3 | 0.84-0.96 | 8693-31,311 | (25) |

### 4.2 AI-Based Diagnostic Indices

#### Development of New Indices (e.g., Matos & Carvalho Index)

The index creation with the help of AI is based on optimization of hematologic formulas. Matos and Carvalho Index (MCI) = 1.91(RBC) + 0.44(MCHC) with a cutoff of 23.85 used to identify carriers with a sensitivity of 99.3%. It was based on 106 patients and tuned with ROC analysis on 227 patients and uses ROC analysis. MultiThal-Classifier (M-THAL) is an XGBoost-based classifier with sensitivity of 90% when used to identify normocytic, microcytic thalassaemia, iron deficiency and normals. It has 14 CBC features, with SHAP ranking features such as MCV. The index, TVGH-NYCU, which is driven by ML, has an AUC of 0.76 in adults. These indices take advantage of regular tests, and do not rely on genetics [28] .

#### Traditional Indices Comparisons

Conventional indices such as Mentzer have 94% sensitivity but have overlap issues. Green & King capture 97% AUC but fail to capture normocytic cases. The sensitivity of AI indices such as MCI is enhanced to 99%, and the false positives are reduced by 15%. GBoost models achieved an AUC of 99%, which is much higher than the 75% observed in SCS BTT. CatBoost is able to predict beta-thalassaemia with 80% accuracy in multi-class, which is better than formulas of 0% sensitivity [28].

**Table 3: Comparative Performance of AI-Based and Traditional Diagnostic Indices for thalassaemia**

| Index Type | Sensitivity (%) | Specificity (%) | AUC | Limitations | Citations |
|---|---|---|---|---|---|
| Mentzer (traditional) | 94 | 75 | 0.93 | Overlaps with IDA | (18) |
| Green and King (traditional) | 97 | 91 | 0.97 | Normocytic misses | (20) |
| M-Thal (Al) | 90 | 98 | 0.94 | Computation intensive | (25) |
| SCS BTT (traditional) | 64 | 79 | N/A | Low accuracy | (26) |

### 4.3 Deep Learning Approaches

Deep learning (DL), a subfield of machine learning, which uses multi-layer neural networks to automatically learn hierarchical feature representations from raw data. Unlike traditional machine learning, where human experts manually design or select features, DL models can internally optimize which features matter most. In haemotology and diagnostics, DL has been increasingly applied, from analyzing molecular data to automating cytomorphology and interpreting lab results [29].

DL is particularly promising for thalassaemia diagnosis because it can integrate **morphological information** (e.g. images of red blood cells) with **quantitative laboratory data** (e.g. CBC parameters) as learning nonlinear interactions is difficult with human-designed indices.

#### Convolutional Neural Networks (CNN)

Deep convolutional neural networks have been applied to thalassaemia screening by learning complex, non-linear patterns in standard blood count data (and sometimes images) that traditional formulaes cannot capture. In practice, CNNs often far outperform simple indices. For example, Nasir et al. (2025) applied a CNN to β-thalassaemia data and reported ~98.1% accuracy on a private clinical dataset. Similarly, another study trained a CNN on 288 Sri Lankan α-thalassemic cases and achieved ~85% accuracy (with AUC≈0.95) , notably higher than any classical index on the same data. These models automatically combine multiple CBC features (MCV, MCH, RBC, etc.) through learned filters, enabling robust classification of carriers vs healthy controls. In summary, CNNs have demonstrated very high sensitivity and specificity in recent studies, often reducing missed cases relative to older methods [30].

#### XGBoost

Extreme Gradient Boosting (XGBoost) is an ensemble tree-based algorithm well-suited to tabular haemotology data. It builds many decision trees sequentially, capturing feature interactions that simpler models might miss. In thalassaemia screening, XGBoost has achieved near-perfect accuracy in several reports. For instance, Nasir et al. found that XGBoost reached **99.34%** accuracy for α-thalassaemia on their private dataset. Another study combining CBC and HPLC features reported XGBoost training accuracy of ~99.5% (with >99% test accuracy) [31], outperforming SVM. These results suggest XGBoost can read subtle patterns (e.g. in MCV, MCH, HbA2 levels) with remarkable precision. In practice, XGBoost is often as accurate or more accurate than deep neural networks in this domain, while requiring less parameter tuning for smaller datasets. Its strengths are high predictive power on structured data and robustness to collinear features; its drawback is that it still needs adequate data and careful tuning to avoid overfitting.

#### Public vs. Private Datasets

Data source has a major impact on AI performance. Publicly available thalassaemia datasets are typically small (often <1000 samples) and may lack detailed features, whereas private clinical cohorts can be much larger and richer. Nasir et al. explicitly compared models on public vs. private data and found far higher accuracy on the private set . They attributed this to greater sample size and higher-quality measurements in the private data. For example, their CNN/XGBoost models achieved ~98–99% accuracy on a large hospital dataset, whereas training on smaller published datasets yielded weaker results. Likewise, Christensen et al. trained a CNN on only 288 patient records from Sri Lanka and achieved "only" ~85% accuracy, far below the ~98% seen with larger data. These examples underscore that modern AI models benefit immensely from large, diverse cohorts. In practice, a deep learning model trained on one hospital's dataset may not generalize to another region unless trained or validated; hence, data-sharing and multi-center studies are needed to achieve robust, widely applicable algorithms [32].

### 5. Comparative Analysis of Studies

Recent literature shows that AI-based classifiers consistently outperform traditional haematological indices in thalassaemia screening. Traditional formulae like the Mentzer or Shine–Lal index yield modest sensitivity/specificity (e.g. Mentzer index ~70–95% sensitivity ), leading to frequent misclassification. In contrast, ML and DL models typically achieve much higher accuracy. Numerous studies found many modern algorithms with accuracy in the mid-90% or higher. For example, in multiple recent reports, AI systems reached >98% accuracy. One CNN-based model (using image and CBC features) achieved **99%** accuracy with 100% sensitivity. In practice, ML/DL methods often detect virtually all carrier cases that simpler indices would miss.

Key comparative findings include:

**Accuracy and AUC:** AI models (SVM, RF, XGBoost, CNN, etc.) generally report **90–99%** accuracy. For instance, Nasir et al. (2025) report 98.10% accuracy (CNN) and 99.34% (XGBoost) on

large datasets. In comparison, classical indices typically generated accuracies around 75–90% (and often a trade-off between sensitivity and specificity). Several studies have reported that machine learning approaches can achieve remarkably high accuracy in detecting β-thalassaemia traits, often reaching or even exceeding 95%. In some cases, these advanced models demonstrate near-perfect performance, underscoring their strong capability to distinguish thalassaemia from other anaemias and their claim as powerful tools in clinical diagnostics. In summary, ML and DL approaches significantly reduce both false negatives and false positives relative to traditional scoring.

**Table 4: study results**

| Model | Dataset / Focus | Reported Accuracy | Source |
|---|---|---|---|
| CNN & XGBoost | α- & β-thalassaemia; public vs private datasets | 98.10 % (CNN), 99.34 % (XGBoost) | [30] |
| CNN | 288 α-thalassaemia cases (small dataset) | 85 % | [30] |
| CNN + PCA | β-thalassaemia | 96 % | [33] |
| Federated Learning | 5,066 β-thalassaemia patients | 92.38 % | [34] |
| Ensemble (multiple classifiers) | α + β detection (5,066 cases) | 93 % | [35] |
| Hybrid CNN (images + CBC) | Blood-smear + CBC data | 99 % (100 % sensitivity) | [36] |

**Strengths:** For AI classification, multiple features are simultaneously integrated, capturing subtle patterns. They can be trained to improve overall accuracy, sensitivity and specificity. Many studies report high sensitivity (often >95%) e.g. the CNN image-based model achieved 100% sensitivity. Ensemble methods like XGBoost combine weak learners to reach >99% accuracy [33], robustly distinguishing α and β variants. ML/DL can also be updated with new data, making them adaptable to different populations. Overall, these methods dramatically reduce misdiagnosis and the risk of giving inappropriate iron therapy to unrecognized thalassaemia carriers.

**Weaknesses:** The main limitations of these studies are **data quality and size**. Many models were trained on relatively small or homogeneous datasets, raising overfitting concerns. For example, deep neural net achieved 96% accuracy but the authors noted its small training set and overfitting risk [34]. Several studies could not distinguish thalassaemia subtypes due to incomplete labeling. Furthermore, while neural nets (CNN/ANN) often perform best, they are “black boxes” with low interpretability; simpler indices, though less accurate, are transparent. Traditional indices still have the advantage of requiring no training and being extremely easy to compute, but they sacrifice accuracy (e.g. Mentzer’s index had only ~70% sensitivity ). In summary, AI methods are powerful but depend on large, representative data and careful validation. When data are limited, even high-performing models can degrade; this is why some studies report lower accuracy (e.g. 85% in the small Sri Lankan study ).

Studies have shown that while traditional indices demonstrate high diagnostic efficiency for β-thalassaemia trait detection, modern machine learning models often achieve even greater accuracy. Conventional approaches like the Mentzer index provide reliable results, yet AI-based models consistently surpass them in precision and overall performance, highlighting the growing potential of intelligent algorithms in improving diagnostic outcomes. . In practice, this means AI approaches correctly classify far more carriers. In summary, across multiple recent high-impact studies, **AI/ML methods consistently outperform traditional hematologic indices** in accuracy and sensitivity, at the cost of requiring more data and computational effort. These methods’ strengths (high accuracy, flexibility, ability to handle complex feature sets) make them promising for screening programs, while their weaknesses (data-dependence, explainability)

are being addressed through larger studies and interpretable models.

## 6. Clinical applications and implications

Early diagnosis, carrier screening, and optimized monitoring are central to reducing morbidity and preventing new affected births. Artificial intelligence (AI) and machine learning (ML) offer promising enhancements to conventional approaches. Below is a structured discussion of clinical applications, implications, and the challenges/limitations of AI in thalassaemia care.

### 6.1 AI in Routine Screening Programme

Routine screening programme for thalassaemia carriers or mild cases historically rely on hematologic indices (CBC, RBC indices), iron studies, haemoglobin electrophoresis / HPLC, and, where available, DNA testing. AI/ML tools can augment or partially automate the early triage stage to flag suspected carriers or disease.

Several studies have demonstrated that ML models using routine CBC features (e.g., MCV, MCH, RDW, RBC count) can distinguish β-thalassaemia trait from IDA with high accuracy. For instance, a study of 396 individuals (216 IDA, 180 β-thalassaemia minor) showed that ANN and decision tree models outperformed discriminant indices in differentiation using CBC alone [35]. Another more recent review collated many ML models achieving sensitivities and specificities over 90% in distinguishing thalassaemia carriers from non-carriers or IDA controls. In addition, MultiThal, a multiclass machine learning classifier, has been proposed for more refined diagnosis and subtype classification of thalassaemia using CBC data [36]. A further model, designed for pregnant women, used a clinical indicator–based ML algorithm to predict thalassaemia risk and achieved good discrimination performance.

Beyond tabular indices, image-based AI has been applied. Deep learning on Hb electrophoresis strip images has been used for automatic detection of abnormal bands. For instance in one study, using 524 electrophoresis images, CNN models (e.g. InceptionV3) achieved detection accuracy ~95.8% for distinguishing thalassaemia vs normal patterns [37]. Thus, AI may accelerate and standardize interpretation of electrophoresis, reducing human error and burden.

In α-thalassaemia, classification using ML models has been explored. A study in Sri Lanka used CBC and gender features and applied ML to distinguish silent and non-carrier states of α-thalassaemia trait. Another experiment explored hybrid deep networks (transfer learning with feature fusion) achieving precision > 94% for thalassaemia detection [38].

Thus, in screening programmes, AI holds the potential to filter large numbers of low-risk individuals and flag high-risk ones for further confirmatory workup, increasing throughput and reducing cost burdens.

### 6.2 Avoiding Misdiagnosis and Unnecessary Iron Therapy

A clinical issue is misdiagnosis of thalassaemia with IDA which might result in unneccesary iron administration causing Iron overload, oxidative damage. The AI methods assist in decreasing the misclassification through the addition of multidimensional characteristics other than the straightforward indices. As an example, standard indices (e.g. Mentzer, RDW*MCH) do not work well in overlapping phenotypes. By comparison, multi-classifier models that combine several CBC characteristics have demonstrated a high diagnostic sensitivity and reduced false positive results [38]. One of them is the study named “Predicting thalassaemia Using Feature Selection Techniques” in which nine classification algorithms, as well as several feature selection methods, were evaluated; the authors established the high discrimination of carrier and non-carrier conditions and shed light on the shortcomings of prior single-metric indices. In a multiclass classification study, the models were trained to identify both α and β-thalassaemia (major/minor) on the same model this was to ensure that mixed phenotypes are not misdiagnosed. Therefore, AI can serve as a protective measure that will decrease under- and over-treatment.

### 6.3 Genetic Counselling and Carrier Detection by AI

The control strategies include genetic counselling and carrier detection. These can be improved with the help of AI and bioinformatics tools in the following ways:

**Carrier identification on sequencing data**: AI can help variant calling, prediction of pathogenicity, structural variant detection, as well as genotype structural variant correlation. Indicatively, more recent molecular reviews have started covering the incorporation of AI into gene editing, version interpretation, and epigenetic insights into thalassaemia scenarios [39].

**Risk estimation and phenotypic prediction**: Given a genotype (or pair of variants), AI models may help predict disease severity (trait, intermedia, major) by incorporating modifier genes, regulatory variants or epigenetic features.

**Counselling support tools**: Mobile or web-based platforms embedding AI modules can deliver personalized counselling.
By integrating genotype, phenotype, and demographic data, AI may support more specific carrier detection and assist in counselling couples with more precise risk estimates.

### 6.4 AI in Monitoring Complications and Treatment Response

Once a diagnosis is established, thalassaemia patients require ongoing surveillance and management. AI can add value in:

- **Predicting transfusion requirements or trends:** Time-series ML models may forecast when a patient will require increased transfusion intensity, enabling earlier intervention.
- **Estimating iron overload non-invasively:** MRI T2 images are standard for quantifying cardiac/liver iron, but computational methods may enhance analysis. For instance, CHMMOTv1 dataset of cardiac and hepatic MRI images in thalassaemia provides a resource for training AI models to estimate organ iron burden . AI models may integrate lab trends, imaging, and demographics to better personalize chelation therapy.
- **Predicting complications:** Endocrinopathies, cardiac disease, liver fibrosis often develop gradually. Risk stratification AI models could flag patients at higher risk requiring closer follow-up.
- **Assessing treatment response:** For chelators, gene therapy, or novel agents, AI models can analyze composite biomarkers to distinguish responders vs nonresponders earlier, potentially guiding therapy adjustments.

### 6.5 Clinical Implications

**Earlier and more accurate diagnosis:** AI tools can shorten the diagnostic delay, reducing end-organ damage and morbidity.
**Reduced misdiagnosis and harm:** Better discrimination avoids unnecessary iron therapy or delayed diagnosis.
**Personalization:** AI enables stratified risk and monitoring plans.
**Resource efficiency:** Automating parts of interpretation reduces the burden on specialist clinicians and laboratories.

#### 6.5.1 Economic Implications

AI use as a filter can save us on the number of confirmatory tests (HPLC, DNA sequencing) that are costly to carry out.
Preventing the unnecessary iron therapy or overtreatment helps decrease direct and indirect healthcare expenses.
Laboratory automation (smear reading, interpretation of electrophoresis) saves on labor time and turnaround.
These expenses comprise model development of AI, infrastructure (computing, storage), validation, as well as regulatory compliance, which should be traded off against long-term savings. Cost-benefit models in most contexts are necessary to gain adoption particularly in low and middle income nations with constrained budgets.

## 7. Challenges and Limitations

**Quality of Data and Sample Diversity**
Most AI research is retrospective, small, single-centre and based on quite homogeneous populations (ethnicity, mutation spectra). This restricts the general approach and external validity. Missing values, measurement error (different CBC machines, reagents), or labeling noise can lead to datasets with missing values, variability in measures, or which worsen the robustness of the models. The models cannot effectively deal with edge cases because the rare genotypes or mixed phenotypes (coexistent IDA + trait) are not well represented in the model.
**Overfitting and Bias**
The deep models or complex ML are prone to overfitting, particularly when the sample sizes are small; the performance can deteriorate when using external data. Its introduction may occur because of stratification of populations (populations trained on a single ethnicity may not work well on a different population), selection bias (between hospital and community samples), and prevalence bias (lop-sided classes). The absence of transparency (black boxes) can cause a decrease in clinical trust or the presence of hidden biases.
**Integration into Healthcare Systems**
AI tools must interoperate with laboratory information systems (LIS), EHRs, and hospital workflows. Without seamless integration, uptake is limited. Clinician acceptance requires that AI outputs be interpretable, reliable, and delivered in actionable form at the point of care. Maintenance, updates, retraining and version control over time are necessary to keep models valid.
**Ethical, Privacy, and Regulatory Issues**
Use of genetic data, imaging, and identifying health data demands strict data security, de-identification, and informed consent frameworks. Regulatory oversight (e.g., as medical devices or decision

support tools) may require validation trials, audits, and approvals which differ across jurisdictions.
Clinicians and patients may demand transparent reasoning rather than black box predictions. Liability:Legal and ethical frameworks must be established for AI mistakes

**Equity and Accessibility**

Many high-prevalence regions are resource-limited. AI tools requiring advanced computation, stable internet, or high-resolution imaging may not be feasible. Implementation costs, training of personnel, maintenance, and infrastructure may not be affordable in low-income settings. If AI models are developed using data from high-income settings and directly imported to underprivilaged populations, disparities may worsen.

**Model Generalisability and Performance Inflation**

Although several studies report diagnostic accuracies exceeding 95%, these results are often obtained under controlled conditions using relatively homogeneous or single-centre datasets. The lack of large, multi-centre external validation raises concerns regarding overfitting and real-world generalisability. Therefore, high reported accuracy should be interpreted cautiously until prospective clinical validation and regulatory approval pathways are established.

## 8. Future Perspectives

### 8.1 Use of AI in Large-Scale Screening Programs

Putting artificial intelligence into big testing programs has greatly changed diseases detection, making things reiable, fast, and versatile. Computer programs that based on deep learning model have been better at predicting diabetes and prediabetes, which means fewer pointless tests and overlooked sicknesses. Using artificial intelligence in large breast cancer tests has kept the right diagnosis rate while greatly lowering the load of work, cutting the need for another reader. In the same way, AI based lung cancer tests help radiologists feel confident by giving other readings of lung lumps and finding health problems related to smoking. Computer learning systems is also being used for big virtual drug tests, grouping chemicals well and sorting out correct and wrong results[40]. Also, deep learning methods in diabetic eye disease tests have been accurate, making things easier to grow and better using resources in public health areas. Artificial intelligence is promising for making work better through sorting tools, improving how things work, lowering false negatives, and keeping the same detection rate in tests for everyone.

### 8.2 Combining AI with Genetic Testing for Precision Medicine

Artificial intelligence has sped up the progress of tailored medical care by making genetic tests and data analysis better. By adding AI, complicated genetic information can be turned into useful understandings, which helps with personalized diagnoses and treatments. Using machine learning to mix health and genetic details has made guesses about how people will react to medicine better, like in the production of seizure medicines. Machine learning programs have gotten total model accuracy up to 88%, making choices about treatment and predictions better. AI-based drug response analyses also make custom-made medicine plans easier, which makes treatments work better and lowers bad side effects [41]. Also, using AI to mix different kinds of biological information helps find diseases early, change treatments for each person, and make drug creation processes work better. Overall, these uses highlight how AI is changing tailored medicine, helping with the ongoing big change in healthcare.

### 8.3 Mobile/Point-of-Care AI Tools for Low-Resource Settings

New improvements highlight how helpful artificial intelligence could be in portable and on-site medical tools, mainly for places with few resources and urgent situations. AI make processes work better and improve accuracy of diagnoses in poorer countries. Computer programs that learn deeply can help people who respond to emergencies by making it easier to find diseases right away using tools that can be carried around. Even with these improvements, there are still problems with how well they work in different situations, how easy it is to get data, and making sure POCUS tools and programs are all the same [42]. Thinking about what is right and wrong, how well patients follow instructions, and making sure everyone has fair access are also very important things that affect whether these tools are used successful. Even so, portable medical tests that use AI are still a reliable way to make healthcare more available everywhere.

### 8.4 Potential Expansion to Other haemoglobinopathies

Conditions such as β-thalassaemia and sickle cell disease are widespread inherited problems globally and lead to serious health concerns. Although bone marrow transplants represent the sole method for completely fixing them, there aren't sufficient individuals available for donation. Another potential method involves gene therapy, which might introduce

genes, stop them, or accurately adjust areas influencing globin [43]. However, it remains challenging to alter cells sufficiently and guarantee that blood stem cells continue multiplying over an extended period. Looking ahead, we must discover improved methods to render gene therapy more affordable and simpler to manufacture, while also addressing ethical and realistic concerns. These innovative concepts may assist in decreasing the quantity of individuals becoming ill and dying due to inherited blood conditions across the globe [44].

**Conclusion:**

The field of Artificial Intelligence has demonstrated enormous potential in transforming the diagnostic environment of thalassaemia. Although traditional indices and methods of molecular analysis still hold a core role, AI-based methods introduce a whole new accuracy, efficiency, and availability of the process. Clinicians are now able to diagnose thalassaemia carriers more efficiently, through machine learning and deep learning models, and avoid cases of misdiagnosis and unwarranted treatment. The application of AI in clinical practice also creates new opportunities in terms of round-the-clock monitoring, genetic counseling, and screening of a large population. Nevertheless, to be able to put these innovations into practical use, the problem of standardization of data, transparency of algorithms, and their ethical application should be resolved. AI must not be considered the alternative to molecular testing; it can be regarded as the supplementary tool that improves diagnostic accuracy and expands the access to early diagnosis, particularly in the resource-limited setting. As the world continues to evolve and AI is implemented responsibly, it is bound to be a pillar in the battle against thalassaemia in the world.

**References**:

1. Tuo Y, Li Y, Li Y, Ma J, Yang X, Wu S, Jin J, He Z. Global, regional, and national burden of thalassemia, 1990–2021: a systematic analysis for the global burden of disease study 2021. EClinicalMedicine. 2024 Jun 1;72.
2. Gao J, Liu W. Advances in screening of thalassaemia. Clinica Chimica Acta. 2022 Sep 1;534:176-84.
3. Damkliang K, Wongsirichot T, Suwanno K, Cheranakhorn C, Sangkeaw S, Sottisataworn S. An AI-Based Decision Support Framework for Clinical Screening of Iron Deficiency Anemia and Thalassemia. IEEE Access. 2025 Jul 25.
4. Ferih K, Elsayed B, Elshoeibi AM, Elsabagh AA, Elhadary M, Soliman A, Abdalgayoom M, Yassin M. Applications of artificial intelligence in thalassemia: A comprehensive review. Diagnostics. 2023 Apr 26;13(9):1551.
5. Uçucu S, Azik F. Artificial intelligence-driven diagnosis of β-thalassemia minor & iron deficiency anemia using machine learning models. Journal of Medical Biochemistry. 2024 Jan 25;43(1):11.
6. Peters M, Heijboer H, Smiers F, Giordano PC. Diagnosis and management of thalassaemia. BMJ. 2012 Jan 25;344.
7. Çil B, Ayyıldız H, Tuncer T. Discrimination of β-thalassemia and iron deficiency anemia through extreme learning machine and regularized extreme learning machine based decision support system. Medical hypotheses. 2020 May 1;138:109611.
8. Liang ST, Wong VC, So WW, Ma HK, Chan V, Todd D. Homozygous α-thalassaemia: clinical presentation, diagnosis and management. A review of 46 cases. BJOG: An International Journal of Obstetrics & Gynaecology. 1985 Jul;92(7):680-4.
9. Tepakhan W, Srisintorn W, Penglong T, Saelue P. Machine learning approach for differentiating iron deficiency anemia and thalassemia using random forest and gradient boosting algorithms. Scientific Reports. 2025 May 15;15(1):16917.
10. Munkongdee T, Chen P, Winichagoon P, Fucharoen S, Paiboonsukwong K. Update in laboratory diagnosis of thalassemia. Frontiers in molecular biosciences. 2020 May 27;7:74.
11. Jahan A, Singh G, Gupta R, Sarin N, Singh S. Role of red cell indices in screening for beta thalassemia trait: an assessment of the individual indices and application of machine learning algorithm. Indian Journal of Hematology and Blood Transfusion. 2021 Jul;37(3):453-7.
12. Diaz-del-Pino S, Trelles-Martinez R, González-Fernández FA, Guil N. Artificial intelligence to assist specialists in the detection of haematological diseases. Heliyon. 2023 May 1;9(5).
13. Fu YK, Liu HM, Lee LH, Chen YJ, Chien SH, Lin JS, Chen WC, Cheng MH, Lin PH, Lai JY, Chen CM. The TVGH-NYCU thal-classifier: development of a machine-learning classifier for differentiating thalassemia and non-thalassemia patients. Diagnostics. 2021 Sep 20;11(9):1725.
14. Tran DC, Dang AL, Hoang TN, Nguyen CT, Le TM, Dinh TN, Tran VA, Doan TK, Nguyen TT. Prevalence of thalassemia in the Vietnamese

population and building a clinical decision support system for prenatal screening for thalassemia. Mediterranean Journal of Hematology and Infectious Diseases. 2023 May 1;15(1):e2023026.

15. Mahmood S, Khalaf AA, Hamadi SS. Basrah Score: a novel machine learning-based score for differentiating iron deficiency anemia and beta thalassemia trait using RBC indices. Frontiers in Big Data. 2025 Aug 4;8:1634133.
16. Zhang F, Zhan J, Wang Y, Cheng J, Wang M, Chen P, Ouyang J, Li J. Enhancing thalassemia gene carrier identification in non-anemic populations using artificial intelligence erythrocyte morphology analysis and machine learning. European Journal of Haematology. 2024 May;112(5):692-700.
17. Origa R. β-Thalassemia. Genetics in medicine. 2017 Jun;19(6):609-19.
18. Nasir MU, Naseem MT, Ghazal TM, Zubair M, Ali O, Abbas S, Ahmad M, Adnan KM. A comprehensive case study of deep learning on the detection of alpha thalassemia and beta thalassemia using public and private datasets. Scientific Reports. 2025 Apr 17;15(1):13359.
19. Zaylaa AJ, Makki M, Kassem R. Thalassemia Diagnosis Through Medical Imaging: A New Artificial Intelligence-Based Framework. In2022 International Conference on Smart Systems and Power Management (IC2SPM) 2022 Nov 10 (pp. 41-46). IEEE.
20. Lai JX, Tang JW, Gong SS, Qin MX, Zhang YL, Liang QF, Li LY, Cai Z, Wang L. Development and validation of an interpretable risk prediction model for the early classification of thalassemia. npj Digital Medicine. 2025 Jun 10;8(1):346.
21. Farooq MS, Younas HA. Beta Thalassemia Carriers detection empowered federated Learning. arXiv preprint arXiv:2306.01818. 2023 Jun 2.
22. Wang J. Deep learning in hematology: From molecules to patients. Clinical Hematology International. 2024 Oct 8;6(4):19.
23. Sadiq S, Khalid MU, Ullah S, Aslam W, Mehmood A, Choi GS, On BW. Classification of β-thalassemia carriers from red blood cell indices using ensemble classifier. IEEE access. 2021 Mar 17;9:45528-38.
24. Purwar S, Tripathi RK, Ranjan R, Saxena R. Detection of microcytic hypochromia using cbc and blood film features extracted from convolution neural network by different classifiers. Multimedia Tools and Applications. 2020 Feb;79(7):4573-95.
25. Phirom K, Charoenkwan P, Shoombuatong W, Charoenkwan P, Sirichotiyakul S, Tongsong T. DeepThal: a deep learning-based framework for the large-scale prediction of the α+-thalassemia trait using red blood cell parameters. Journal of Clinical Medicine. 2022 Oct 26;11(21):6305.
26. Christensen F, Kılıç DK, Nielsen IE, El-Galaly TC, Glenthøj A, Helby J, Frederiksen H, Möller S, Fuglkjær AD. Classification of α-thalassemia data using machine learning models. Computer Methods and Programs in Biomedicine. 2025 Mar 1;260:108581.
27. Kilicarslan S, Celik M, Sahin Ş. Hybrid models based on genetic algorithm and deep learning algorithms for nutritional Anemia disease classification. Biomedical Signal Processing and Control. 2021 Jan 1;63:102231.
28. Long Y, Bai W. Constructing a novel clinical indicator model to predict the occurrence of thalassemia in pregnancy through machine learning algorithm. Frontiers in Hematology. 2024 Apr 4;3:1341225.
29. Wongseree W, Chaiyaratana N, Vichittumaros K, Winichagoon P, Fucharoen S. Thalassaemia classification by neural networks and genetic programming. Information Sciences. 2007 Feb 1;177(3):771-86.
30. Mo D, Zheng Q, Xiao B, Li L. Predicting thalassemia using deep neural network based on red blood cell indices. Clinica Chimica Acta. 2023 Mar 15;543:117329.
31. Paydaş Hataysal E, Körez MK. Machine Learning-Based Prediction of β-Thalassemia Trait Using Red Blood Cell Indices. International Journal of Laboratory Hematology. 2025 Oct 15.
32. Saleem M, Aslam W, Lali MI, Rauf HT, Nasr EA. Predicting thalassemia using feature selection techniques: A comparative analysis. Diagnostics. 2023 Nov 14;13(22):3441.
33. Wang W, Ye R, Tang B, Qi Y. MultiThal-classifier, a machine learning-based multi-class model for thalassemia diagnosis and classification. Clinica Chimica Acta. 2025 Feb 1;567:120025.
34. Abdulkarim D, Abdulazeez AM. Machine learning-based prediction of thalassemia: A review. The Indonesian Journal of Computer Science. 2024 Jun 15;13(3

35. Salman Khan M, Ullah A, Khan KN, Riaz H, Yousafzai YM, Rahman T, Chowdhury ME, Abul Kashem SB. Deep learning assisted automated assessment of thalassaemia from haemoglobin electrophoresis images. Diagnostics. 2022 Oct 3;12(10):2405.
36. Almbaidin H, Abeden Z, Alhawamdeh M, Nazman HN, Hamdani HY, Yusoff NM, Isa A. Application of artificial intelligence in the diagnosis of haematological disorders. Biomedical Research and Therapy. 2024 Dec 31;11(12):6989-7002.
37. Abdalla HB, Ahmed A, Li G, Mustafa N, Sangi AR. Deep Maxout Network-based Feature Fusion and Political Tangent Search Optimizer enabled Transfer Learning for Thalassemia Detection. arXiv preprint arXiv:2308.02029. 2023 Aug 3.
38. Gastounioti A, Desai S, Ahluwalia VS, Conant EF, Kontos D. Artificial intelligence in mammographic phenotyping of breast cancer risk: a narrative review. Breast Cancer Research. 2022 Feb 20;24(1):14.
39. Goodman MA, Malik P. The potential of gene therapy approaches for the treatment of hemoglobinopathies: achievements and challenges. Therapeutic advances in hematology. 2016 Oct;7(5):302-15.
40. Meenakumari B, Chandramouleeswari K, Dhanasekar S. Advanced molecular approaches to thalassemia disorder and the selection of molecular-level diagnostic testing in resource-limited settings. Hematology, Transfusion and Cell Therapy. 2025 Jul 1;47(3):103860.
41. Ng AY, Glocker B, Oberije C, Fox G, Sharma N, James JJ, Ambrózay É, Nash J, Karpati E, Kerruish S, Kecskemethy PD. Artificial intelligence as supporting reader in breast screening: a novel workflow to preserve quality and reduce workload. Journal of Breast Imaging. 2023 May 1;5(3):267-76.
42. Grenier PA, Brun AL, Mellot F. The potential role of artificial intelligence in lung cancer screening using low-dose computed tomography. Diagnostics. 2022 Oct 8;12(10):2435.
43. Larsen M, Aglen CF, Hoff SR, Lund-Hanssen H, Hofvind S. Possible strategies for use of artificial intelligence in screen-reading of mammograms, based on retrospective data from 122,969 screening examinations. European radiology. 2022 Dec;32(12):8238-46.
44. Gupta A, Zhou HX. Machine learning-enabled pipeline for large-scale virtual drug screening. Journal of chemical information and modeling. 2021 Aug 17;61(9):4236-44.