

Article

A Novel Algorithm Using Cell Population Data (VCS Parameters) as a Screening Discriminant between Alpha and Beta Thalassemia Traits

Angeli Ambayya ^{1,2,*} , Santina Sahibon ³, Thoo Wei Yang ⁴, Qian-Yun Zhang ⁵, Rosline Hassan ² 
and Jameela Sathar ¹ 

¹ Haematology Department, Hospital Ampang, Ampang 68000, Selangor, Malaysia; jsathar@hotmail.com

² Department Haematology, School of Medical Sciences, Universiti Sains Malaysia, Kota Bharu 15200, Kelantan, Malaysia; roslin@usm.my

³ Gribbles Pathology Malaysia, Petaling Jaya 46100, Selangor, Malaysia; santina@gribbles.com.my

⁴ Straits Scientific Malaysia, Ampang 68000, Selangor, Malaysia; jimmy@3s.com.my

⁵ Department of Pathology, University of New Mexico, Albuquerque, NM 87131, USA; qzhang@salud.unm.edu

* Correspondence: angeli_100182@yahoo.com

Abstract: Thalassemia is one of the major inherited haematological disorders in the Southeast Asia region. This study explored the potential utility of red blood cell (RBC) parameters and reticulocyte cell population data (CPD) parameters in the differential diagnosis of α and β -thalassaemia traits as a rapid and cost-effective tool for screening of thalassemia traits. In this study, a total of 1597 subjects (1394 apparently healthy subjects, 155 subjects with α -thalassaemia trait, and 48 subjects with β -thalassaemia trait) were accrued. The parameters studied were the RBC parameters and reticulocyte CPD parameters derived from Unicel DxH800. A novel algorithm named $\alpha\beta$ -algorithm was developed: $(MN-LMALS-RET \times RDW) - MCH$ to discriminate α from β -thalassaemia trait with a cut-off value of 1742.5 [AUC = 0.966, sensitivity = 92%, specificity = 90%, 95% CI = 0.94–0.99]. Two prospective studies were carried: an in-house cohort to assess the specificity of this algorithm in 310 samples comprising various RBC disorders and in an interlaboratory cohort of 65 α -thalassaemia trait, and 30 β -thalassaemia trait subjects to assess the reproducibility of the findings. We propose the $\alpha\beta$ -algorithm to serve as a rapid, inexpensive surrogate evaluation tool of α and β -thalassaemia in the population screening of thalassemia traits in geographic regions with a high burden of these inherited blood disorders.

Keywords: cell population data; algorithm; α -thalassemia; β -thalassaemia; VCS parameters



Citation: Ambayya, A.; Sahibon, S.; Yang, T.W.; Zhang, Q.-Y.; Hassan, R.; Sathar, J. A Novel Algorithm Using Cell Population Data (VCS Parameters) as a Screening Discriminant between Alpha and Beta Thalassemia Traits. *Diagnostics* **2021**, *11*, 2163. <https://doi.org/10.3390/diagnostics11112163>

Academic Editor: Eloisa Urrechaga

Received: 15 October 2021

Accepted: 18 November 2021

Published: 22 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Thalassaemia and haemoglobin variant disorders occur widely worldwide, with an estimation of 7% of the world population being carriers, as reported by the World Health Organisation (WHO) [1]. About 300,000–500,000 children were diagnosed with significant haemoglobin disorders, and of these, about 80% were born in developing countries [1]. In Malaysia, thalassaemia is the commonest single gene disorder characterised by defects in synthesising one or more globin chains [2]. Individuals with homozygous and double heterozygous mutations are associated with thalassaemia major or intermedia phenotype, high morbidity and mortality, whereas individuals with heterozygous mutations are carriers without exhibiting adverse morbidity [2]. Alpha (α) and beta (β) thalassaemia are the commonest types of thalassaemia in Malaysia [3].

Malaysian Thalassaemia Registry (2009) revealed that 3310 out of 4541 registered patients consist of transfusion-dependent β -thalassaemia major and HbE/ β -thalassaemia patients [4]. Clinical management of these patients includes transfusion, iron chelation therapy, and haematopoietic stem cell transplantations, which are expensive and a burden

to the country's healthcare system [5]. Besides patient management, the National Thalassaemia Prevention and Control Programme was actively carried out nationwide to provide population screening and counselling at primary healthcare facilities and health education and promotions to disseminate knowledge and awareness about the consequences of these disorders [4].

α -thalassaemia is characterised by the deficiency or absence of α -globin chain synthesis due to deletions of one or more α -globin genes located in chromosome 16 [6]. The severity of α -thalassaemia is characterised by the number of gene deletions, where single gene deletions result in α -thalassaemia silent carriers, two gene deletions ensue α -thalassaemia trait (minor), three gene deletions eventuate in haemoglobin H (HbH), and four-gene deletions result in haemoglobin Bart's, which usually results in fatal hydrops fetalis that have very high fatality rate [7]. On the other hand, β -thalassaemia is caused by the reduced or absent β -globin chain synthesis due to mutations in the β -globin gene on chromosome 11 [8]. Generally, β -thalassaemia is categorised as β -thalassaemia trait (minor), β -thalassaemia intermedia, and β -thalassaemia major based gene defects and severity of the decrease in β -globin chains production [9].

Preliminary identification of thalassaemia carriers is carried out by a screening program by full blood count (FBC) analysis as the baseline test to scrutinise the red blood cell indices, followed by morphologic examination of peripheral blood smears and subsequent confirmation by high-performance liquid chromatography (HPLC), haemoglobin (Hb) electrophoresis, and molecular genetics testing [4]. Each of the confirmation tests mentioned has its limitation, necessitating a combination of a few tests for differential diagnosis of α and β -thalassaemia. For instance, HPLC is not sufficient to discreetly detect variants as it is not sensitive and specific enough for an α -thalassaemia detection, especially in the presence of haemoglobin Constant Spring (non-deletion α -thalassaemia), indicated by the presence of a very tiny peak of Hb that is often missed [10,11]. In contrast, it is seen on the capillary electrophoresis (CE) electropherogram. Besides that, Hb electrophoresis also cannot be used as the sole technique to distinguish α - and β -thalassaemia. Haemoglobin electrophoresis with β -thalassaemia trait usually shows reduced or absent adult haemoglobin (HbA), elevated levels of haemoglobin A2 (HbA2), and increased foetal haemoglobin (HbF) [10,12]. However, a normal concentration of HbA2 does not rule out the β -thalassaemia trait, especially if there is a concurrent iron deficiency or δ -thalassaemia, which can lower HbA2 levels into the normal range. Therefore, a combination of other methods is required for confirmatory diagnosis [12]. Combinations of these multiple testing are rather costly and require expertise to make accurate and reliable diagnoses as there are substantial overlaps among these disorders [10]. Numerous combinations of red cell indices have been suggested to distinguish different types of thalassaemia as this would allow cost optimisation as only patients requiring further investigations are subjected to expensive haemoglobin electrophoresis and molecular genetics testing.

Along with the advent of new generation full blood count (FBC) analysers, advanced parameters such as cell population data (CPD), which measure the characteristics of cells based on multiple light scatters, are being explored to understand the diagnostic utilities of these parameters better. The CPD parameters provide information on the volume (V), conductivity (C), and light scatter angles (ALL, LALS, LMALS, UMALS, and MALS) for reticulocytes, which are important in the assessment of bone marrow red blood cell production efficiency and detection of haemoglobin disorders such as thalassaemia and anaemia [13]. In this study, we explored the potential utility of reticulocyte CPD and aimed to construct an algorithm that could serve as a rapid tool and an inexpensive surrogate method for the discrimination of α - and β -thalassaemia traits in population-based thalassaemia screening programs held not only in Malaysia but also in other countries with a high burden of these red cell disorders.

2. Materials and Methods

A total of 1597 subjects were included in the primary assessment and algorithm development cohort. The subjects were recruited between September 2011 and February 2019 in a study to establish reference intervals for haematological parameters in Malaysian adults reported in Ambayya et al. (2014) [14]. These cases were retrospectively analysed and categorised into 1394 apparently healthy subjects, 155 subjects with α -thalassemia trait, and 48 subjects with β -thalassemia trait.

Two ethical approvals were obtained for this study from the Medical Research Ethics Committee of the Ministry of Health Malaysia. First, for the recruitments of subjects in the reference interval establishments for haematological parameters (Research ID 10-277-5480) and second, for the exploration of haematological parameters available in FBC analysers in haematological disorders (NMRR 17-2708-38327). Written informed consent was obtained from all subjects before recruitment. Samples were processed according to the recommendation by International Council for Standardisation in Haematology [15–17].

All samples were processed as described in Ambayya et al. (2014) [14]. Initial screening test to rule out anaemia and thalassemia include full blood count (FBC) analysis of red cell indices (MCH < 27 pg, MCV < 80 fL, RDW > 14%) that was performed on Unicel DxH 800 (Beckman Coulter, Miami, FL, USA). Then, peripheral blood film (PBF) was processed using SP1000i automated slide maker (Sysmex, Kobe, Japan) and reviewed by qualified laboratory personnel. Haemoglobin analysis was carried out on Capillars 2 (Sebia, France) and followed by H inclusion detected by supravital brilliant cresyl blue (BCB) staining (Merck Millipore, Darmstadt, Germany). Iron studies assayed in this study include serum ferritin (Modular E170, Roche, Switzerland) and soluble transferrin receptor (Cobas Integra 400, Roche, Switzerland). In subjects with increased serum ferritin [males > 400 ng/mL, females > 150 ng/mL], C-reactive protein was performed using AU480 (Bekman Coulter, USA). DNA testing for confirmation of α -thalassaemia carriers was conducted in cases selected cases that required definitive diagnosis. The analyses' pipeline adopted in this study was based on Malaysia's Management of transfusion-dependent thalassaemia (Clinical Practice guideline) v2009 [4].

Comparison of red blood cell (RBC) and reticulocyte routinely reported parameters, research parameters, and cell population data (CPD) were retrieved from Unicel DxH 800. Description of the CPD parameters is in Supplementary S1. Several strategies were developed first to scrutinise these parameters between the control group (apparently healthy individuals) and subjects diagnosed with α and β -thalassaemia and then select significant indices for a mathematical algorithm development that displayed the largest area under the curve (AUC) as a single parameter.

All statistical analyses were performed using IBM SPSS Statistics 22 software (SPSS, Chicago, IL, USA). The Kolmogorov-Smirnov test of normality was carried out for distribution assessment of each parameter (RBC parameters, research parameters, and CPD). One-way variance analysis (ANOVA) with homogeneity of variance test was conducted and followed by a post-hoc test. Tukey test was done on parameters that fulfilled the homogeneity of variance, while Games-Howell test for the parameters that did not fulfil the homogeneity of variance. Then, the Welch test was done to verify parameters that did not fulfil the homogeneity of variance. For the initial assessment of differences between the α and β -thalassaemia trait subjects, an independent t-test was conducted to test the parameters at a significance level of 0.05. Receiver Operating Characteristic (ROC) curves were generated for parameters that yielded p -values of <0.05. ROC analysis with p -values of <0.05 and AUC > 0.8 was considered significant in this study, and for each shortlisted parameter, cut-off points were determined by considering the sensitivity and specificity.

A novel algorithm, termed as $\alpha\beta$ -algorithm, was formed by utilising routinely reported RBC parameters and reticulocyte CPD with AUC > 0.9 to discriminate α and β -thalassaemia traits. This algorithm was tested in comparison with a control group cohort (apparently healthy individuals), α and β -thalassaemia traits. Following that, two prospective validation cohorts were designed and carried out to ensure the reliability, repro-

ducibility, and robustness of ($\alpha\beta$ -algorithm) when heterogeneous RBC disorders are present in a diagnostic setting: an in-house cohort and an interlaboratory cohort in collaboration with Gribbles Pathology, Malaysia. For the in-house validation cohort, 310 samples were recruited from various red cell disorders subjects (119 α -thalassemia trait, 48 β -thalassemia trait, 15 haemoglobin E trait, 84 iron deficiency anaemia (IDA), and 44 iron deficiency (ID)). We compared the distribution of cases among these red cell disorders to assess the overlap of cut-offs between other cases (non- α or β -thalassemia trait). As for the interlaboratory cohort, 95 samples were included (65 α -thalassemia trait, 30 β -thalassemia trait) recruited during general health screening programs. We used VassarStats clinical calculator (<http://vassarstats.net/clin1.html>, accessed on 16 September 2021) to obtain sensitivity, specificity, positive, and negative predictive values to discriminate α and β -thalassaemia traits [15].

3. Results

3.1. Development of ($\alpha\beta$ Algorithm) Using a Primary Retrospective Cohort

A total of 1597 subjects (1394 control group/apparently healthy; 155 α -thalassemia trait; 48 β -thalassemia trait) were assessed in this study. The distribution of cases and summary of statistics for RBC parameters, sTfR, serum ferritin, and haemoglobin analysis (HBA, HBA2, and HBF) are detailed in Table 1. After performing normality testing, we then compared the groups (control, α , and β -thalassaemia traits) by performing ANOVA and post hoc tests, and the diagnostic performance was determined by ROC analyses, as summarised in Supplementary S2.

Table 1. Distribution of cases and statistics of RBC parameters, serum ferritin, haemoglobin analysis (HBA, HBA2, and HBF).

| Groups | Statistics | HB (g/dL) | RBC (10 /mCL) | HCT (%) | MCV (fL) | MCH (pg) | MCHC (g/dL) | RDW (%) | TSfR (mg/L) | Ferritin (mcg/L) | HBA (%) | HBA2 (%) | HBF (%) |
|--------------------------------|------------|-----------|---------------|---------|-------------------|-------------------|-------------------|-------------------|-------------------|------------------|-------------------|-------------------|-------------------|
| Alpha trait | N | 155 | 155 | 155 | 155 | 155 | 155 | 155 | 155 | 155 | 155 | 155 | 155 |
| | Mean | 12.56 | 5.15 | 39.03 | 76.27 | 24.57 | 32.19 | 15.43 | 4.14 | 117.03 | 97.50 | 2.43 | 0.58 |
| | SD | 1.22 | 5.08 | 3.46 | 5.66 | 2.18 | 0.86 | 1.45 | 1.39 | 138.70 | 0.36 | 0.25 | 0.53 |
| | Med | 12.45 | 0.59 | 38.60 | 77.60 | 25.10 | 32.10 | 15.30 | 3.90 | 72.00 | 97.60 | 2.40 | 0.35 |
| | Min | 10.10 | 3.57 | 32.20 | 60.10 | 19.20 | 29.90 | 12.90 | 1.50 | 3.00 | 95.50 | 1.90 | 0.20 |
| | Max | 16.30 | 6.76 | 49.60 | 98.50 | 32.80 | 35.30 | 21.40 | 9.90 | 766.00 | 98.10 | 3.50 | 2.20 |
| | Range | 6.20 | 3.19 | 17.40 | 38.40 | 13.60 | 5.40 | 8.50 | 8.40 | 763.00 | 2.60 | 1.60 | 2.00 |
| Beta trait | N | 48 | 48 | 48 | 48 | 48 | 48 | 48 | nd | nd | 48 | 48 | 48 |
| | Mean | 11.64 | 5.43 | 36.69 | 68.00 | 21.57 | 31.73 | 17.31 | - | - | 93.83 | 5.23 | 1.13 |
| | SD | 0.87 | 0.55 | 2.60 | 6.24 | 2.17 | 0.73 | 2.01 | - | - | 0.96 | 0.49 | 0.76 |
| | Med | 11.80 | 5.39 | 36.80 | 65.40 | 20.70 | 32.00 | 16.55 | - | - | 93.70 | 5.40 | 1.00 |
| | Min | 10.20 | 4.51 | 32.00 | 61.50 | 18.30 | 29.60 | 14.80 | - | - | 91.70 | 4.40 | 0.30 |
| | Max | 13.20 | 6.45 | 40.90 | 82.10 | 26.30 | 32.90 | 24.30 | - | - | 95.40 | 6.00 | 3.50 |
| | Range | 3.00 | 1.94 | 8.90 | 20.60 | 8.00 | 3.30 | 9.50 | - | - | 3.70 | 1.60 | 3.20 |
| Control group (healthy male) | N | 477 | 477 | 477 | 1394 ^a | 1394 ^a | 1394 ^a | 1394 ^a | 1394 ^a | 477 | 1394 ^a | 1394 ^a | 1394 ^a |
| | Mean | 14.91 | 4.83 | 43.84 | 89.61 | 30.12 | 33.62 | 13.51 | 3.19 | 265.56 | 97.34 | 2.61 | 0.58 |
| | SD | 1.06 | 0.41 | 2.73 | 4.00 | 1.58 | 0.87 | 0.72 | 0.90 | 230.02 | 0.32 | 0.24 | 0.42 |
| | Med | 14.90 | 4.91 | 43.60 | 89.50 | 30.10 | 33.60 | 13.40 | 3.10 | 217 | 97.40 | 2.60 | 0.40 |
| | Min | 12.50 | 3.83 | 37.40 | 80.30 | 26.9 | 30.10 | 11.70 | 0.90 | 37 | 94.90 | 1.10 | 0.20 |
| | Max | 16.90 | 5.61 | 48.20 | 107.80 | 43.50 | 48.20 | 19.00 | 8.80 | 1138 | 98.90 | 3.50 | 2.50 |
| | Range | 4.40 | 1.78 | 10.80 | 27.50 | 17.90 | 18.10 | 7.30 | 7.90 | 1101 | 4.00 | 2.40 | 2.30 |
| Control group (healthy female) | N | 917 | 917 | 917 | 1394 ^a | 1394 ^a | 1394 ^a | 1394 ^a | 1394 ^a | 917 | 1394 ^a | 1394 ^a | 1394 ^a |
| | Mean | 13.00 | 4.24 | 38.54 | 89.61 | 30.12 | 33.62 | 13.51 | 3.19 | 106.89 | 97.34 | 2.61 | 0.58 |
| | SD | 0.81 | 0.32 | 2.61 | 4.00 | 1.58 | 0.87 | 0.72 | 0.90 | 114.04 | 0.32 | 0.24 | 0.42 |
| | Med | 12.90 | 4.23 | 38.50 | 89.50 | 30.10 | 33.60 | 13.40 | 3.10 | 74 | 97.40 | 2.60 | 0.40 |
| | Min | 11.60 | 2.97 | 26.80 | 80.30 | 26.9 | 30.10 | 11.70 | 0.90 | 14 | 94.90 | 1.10 | 0.20 |
| | Max | 14.90 | 5.15 | 44.40 | 107.80 | 43.50 | 48.20 | 19.00 | 8.80 | 669 | 98.90 | 3.50 | 2.50 |
| | Range | 4.60 | 2.18 | 17.60 | 27.50 | 17.90 | 18.10 | 7.30 | 7.90 | 655 | 4.00 | 2.40 | 2.30 |

nd refers to no data. Serum ferritin and sTfR are not tested in β -thalassaemia trait subjects. ^a Based on the haematological reference range for Malaysian adults [18], specific gender-based calculations were performed for HB, RBC, HCT, and serum ferritin. Whereas for the other parameters (MCV, MCH, MCHC, RDW, TSfR, Ferritin, HBA, HBA2, and HBF), no gender-based calculation was done as no clinically gender-based associated variation was reported [18].

ROC analysis involved 203 subjects (155 α -thalassaemia trait and 48 β -thalassaemia trait), in which five parameters were notably higher (MCV, MCH, MAF, MN-V-RET, MN-V-NRET) in α than in the β -thalassaemia trait. The following parameters were lower in α than in the β -thalassaemia trait: RDW, MN-MALS-RET, MN-LMALS-RET, and MN-LMALS-NRET, as shown in Table 2. Three parameters (MN-LMALS-RET, MCH, and RDW) possessed the largest AUC distinguishing α from β -thalassaemia. To improve the sensitivity and specificity of these parameters, a mathematical algorithm ($\alpha\beta$ -algorithm) was devised ($\text{MN-LMALS-RET} \times \text{RDW} - \text{MCH}$) to robustly distinguish α from β -thalassaemia trait with a cut-off of 1742.5, the AUC, sensitivity, and specificity were 0.966 (95% CI: 0.94, 0.99), 92% and 90%, respectively. ROC curves of RBC parameters and reticulocyte CPDs are depicted in Figure 1a,b. Figure 2 displays the ROC curve for the $\alpha\beta$ -algorithm. Box and whisker plots for the for MCH, RDW, MN-LMALS-RET, and the $\alpha\beta$ -algorithm that delineate the α , β -thalassaemia trait and the control group (apparently healthy subjects) are represented in Figure 3a–d. The $\alpha\beta$ -algorithm is incorporated into the screening algorithm for voluntary and cascade screening in the Clinical Practice Guideline for Management of Transfusion Dependent Thalassaemia in Malaysia, as depicted in Figure 4 [4].

Table 2. The ROC data elaboration for each reticulocyte CPD to distinguish α from β -thalassaemia.

| Parameters | AUC | 95% CI | Cut-Off | | Sens (%) | Spec (%) |
|--------------------------|-------|--------------|------------------------|-----------------------|----------|----------|
| | | | α -Thalassaemia | β -Thalassaemia | | |
| $\alpha\beta$ -algorithm | 0.966 | 0.939, 0.993 | <1742.50 | >1742.50 | 91.49 | 90.22 |
| RDW | 0.911 | 0.864, 0.958 | <15.85 | >15.85 | 82.98 | 84.78 |
| MN-LMALS-RET | 0.868 | 0.810, 0.927 | <109.50 | >109.50 | 80.85 | 75.00 |
| MCH | 0.861 | 0.796, 0.925 | >22.30 | <22.30 | 76.09 | 74.47 |
| MN-V-RET | 0.853 | 0.783, 0.923 | >47.50 | <47.50 | 81.52 | 76.59 |
| MN-V-NRET | 0.853 | 0.780, 0.926 | >35.50 | <35.50 | 85.87 | 74.47 |
| MCV | 0.836 | 0.764, 0.909 | >70.80 | <70.80 | 75 | 74.47 |
| MAF | 0.823 | 0.753, 0.894 | >8.95 | <8.95 | 70.65 | 70.21 |
| MN-MALS-RET | 0.816 | 0.744, 0.888 | <118.50 | >118.50 | 74.47 | 73.91 |
| MN-LMALS-NRET | 0.813 | 0.743, 0.883 | <43.50 | >43.50 | 72.34 | 70.65 |

AUC, area under the curve; 95% CI, 95% confidence interval; thal, thalassaemia; MCV, mean corpuscular volume; MCH, mean corpuscular haemoglobin; MAF, microcytic anaemia factor; RDW, red blood cells distribution width; MN-LMALS-RET, lower median angle light scatter reticulocyte; Sens, sensitivity; Spec, specificity. Refer to Supplementary S1, Description of CPD Parameters for the CPD abbreviations).

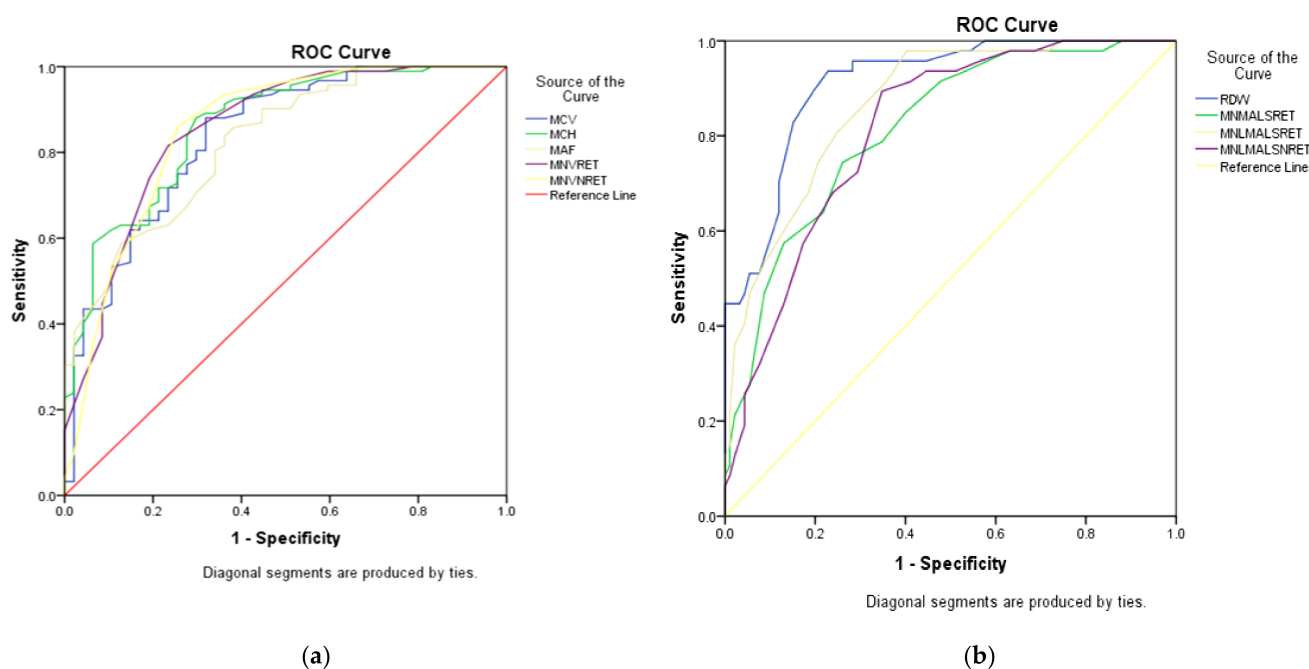


Figure 1. (a,b) ROC curves of RBC parameters and reticulocyte CPD to distinguish α from β thalassaemia trait.

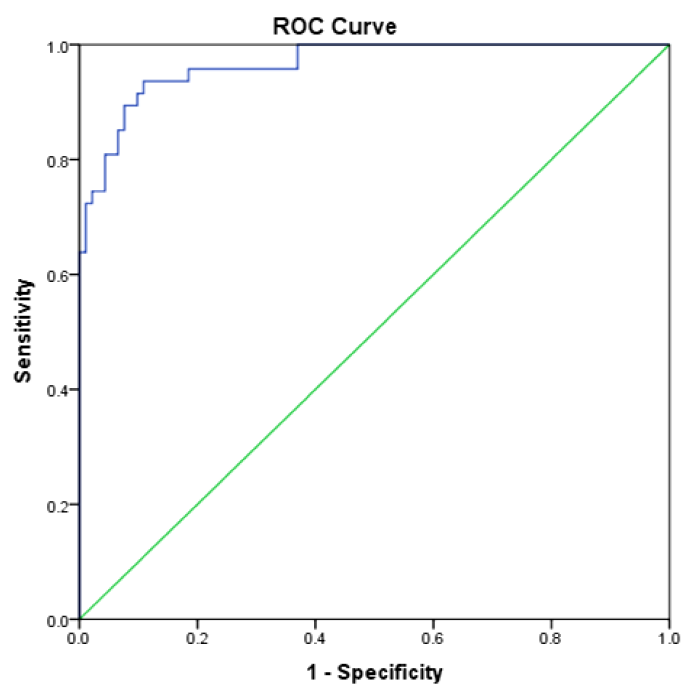


Figure 2. The ROC curve for discriminating α from β -thalassaemia traits generated from the $\alpha\beta$ -algorithm.

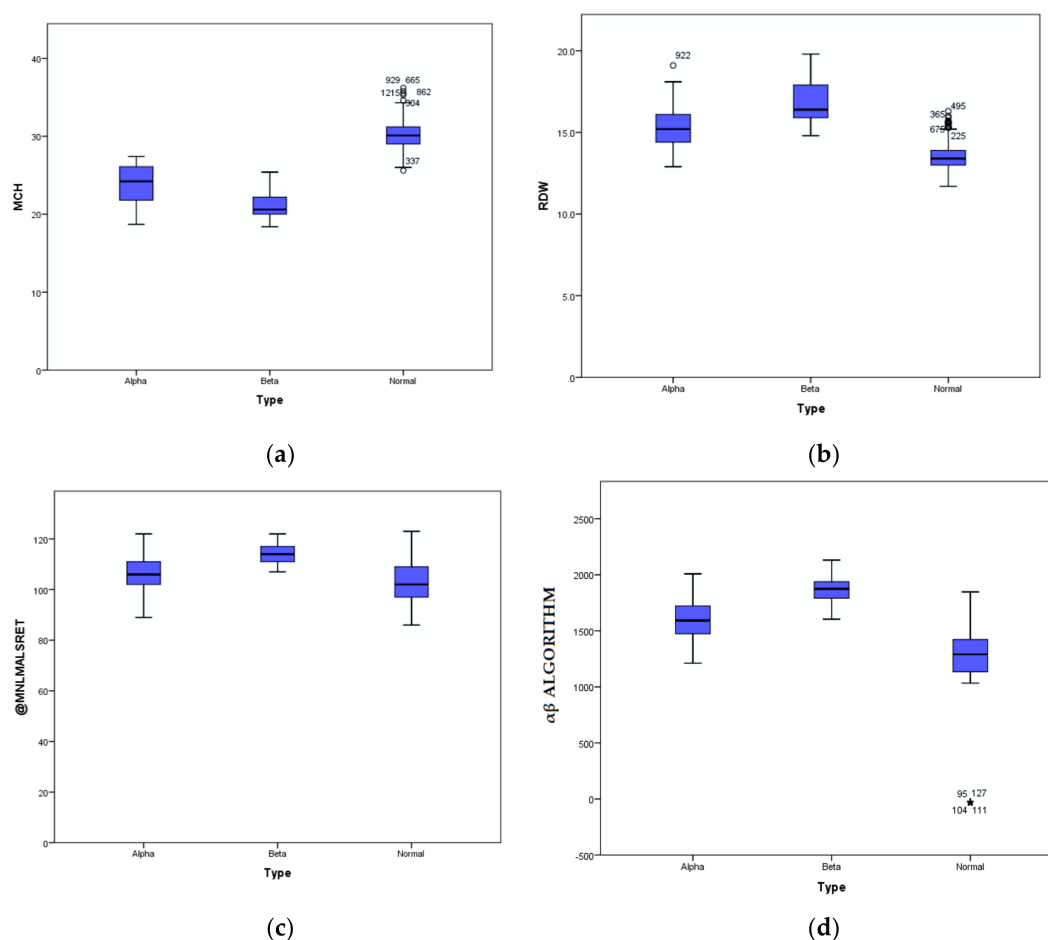


Figure 3. (a–d) Box and whisker plot for MCH, RDW, MN-LMALS-RET (labelled as @MNLMLSRET in Figure 3c and the $\alpha\beta$ -algorithm between the α and β -thalassaemia trait and the control group (apparently healthy subjects). * refers to the outliers.

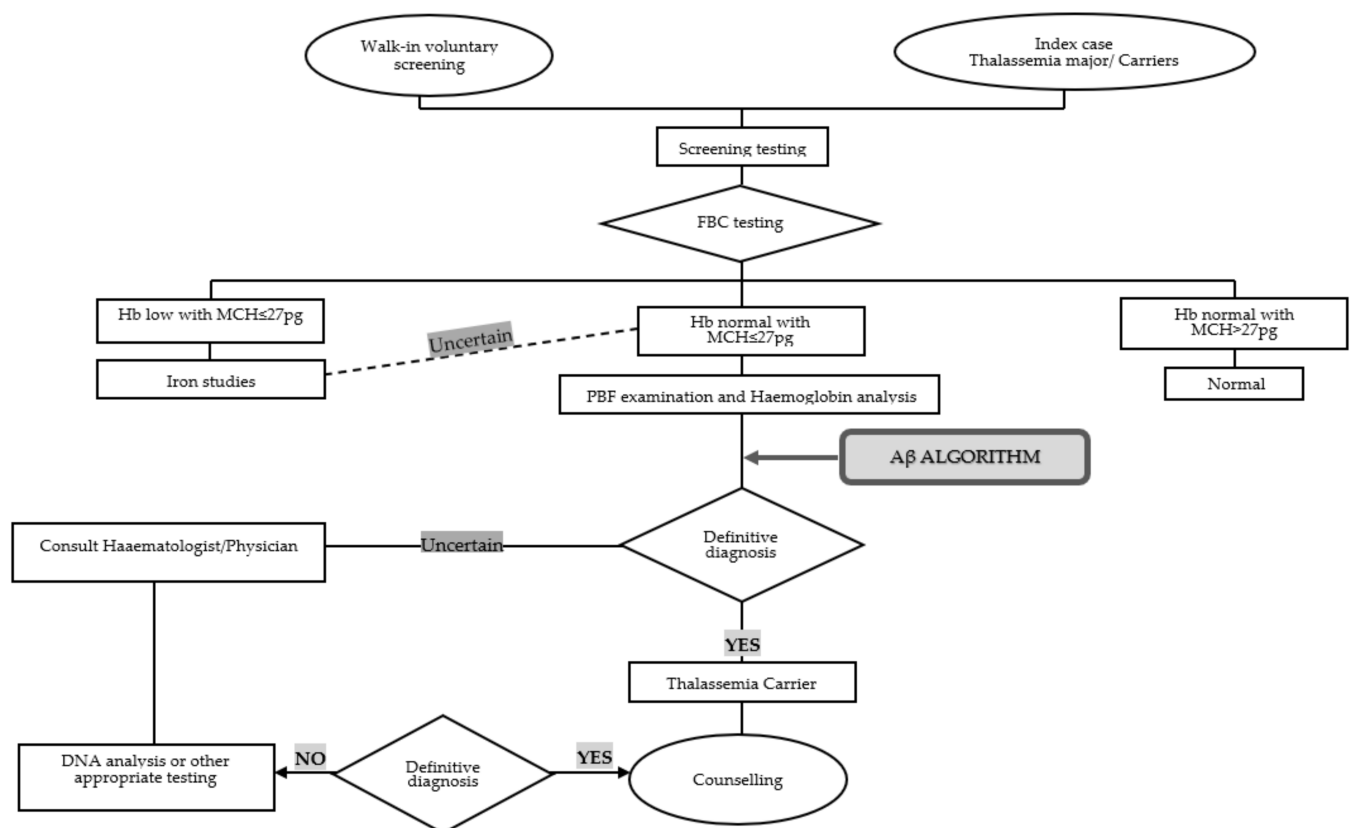


Figure 4. Flowchart representing the $\alpha\beta$ -algorithm incorporated into the screening algorithm for voluntary and cascade screening in Malaysia's Clinical Practice Guideline for Management of Transfusion Dependent Thalassemia [4].

3.2. $\alpha\beta$ -Algorithm Validation Studies

3.2.1. In-House Validation

In-house validation in a cohort of 310 subjects comprising various red cell disorders as summarised in Table 3 and Supplementary S3 and corresponding box and whisker plot is depicted in Figure 5. As shown, between the quartile 1 (Q1) to quartile 3 (Q3) of α thalassemia and β thalassemia traits, there was no significant overlap, comprising the majority of the cases included in this cohort. The overlapping cases require confirmation by HPLC and DNA testing. There is a significant overlap between the IDA and β thalassemia trait.

Table 3. In-house validation of the $\alpha\beta$ -algorithm to distinguish between various red cell disorders.

| Algorithm | Alpha Trait | Beta Trait | HbE Trait | IDA | ID |
|-----------|-------------|------------|-----------|-----------|-----------|
| N | 119 | 48 | 15 | 84 | 44 |
| Mean | 1667.82 | 1960.19 | 1586.31 | 1874.90 | 1430.13 |
| Med | 1660.10 | 1890.40 | 1587.60 | 1857.90 | 1403.65 |
| SD | 148.21 | 248.65 | 80.58 | 245.80 | 153.62 |
| Max | 2058.20 | 2819.50 | 1771.80 | 2526.20 | 1745.80 |
| Min | 1422.70 | 1604.30 | 1441.20 | 1422.80 | 1132.90 |
| Variance | 21,966.37 | 61,829.15 | 6493.59 | 60,417.72 | 23,598.58 |

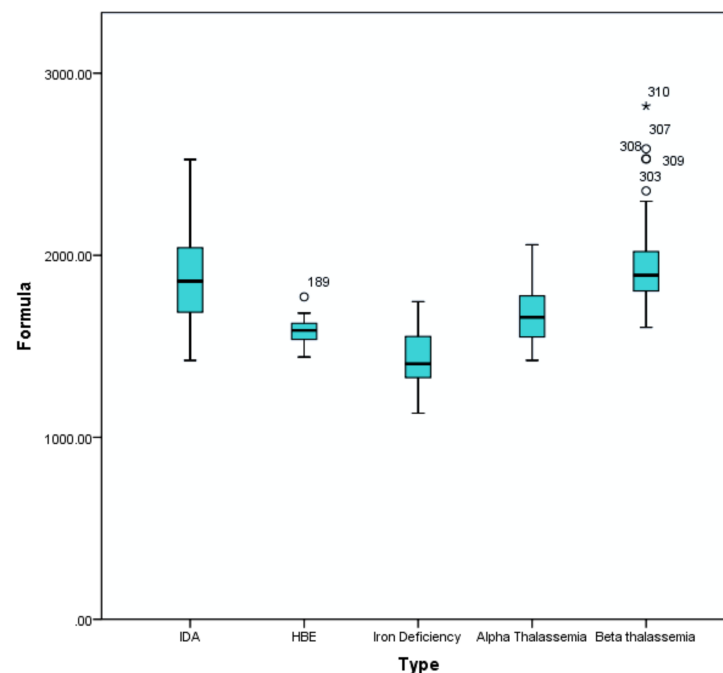


Figure 5. Box and whisker plot for the $\alpha\beta$ -algorithm in various red cell disorders (IDA, HbE trait (labelled as HbE in this figure), iron deficiency, α thalassaemia trait (labelled as alpha thalassemia), and β -thalassaemia trait (labelled as beta-thalassemia). Outliers are marked as “*” and “0” in the HbE trait and β -thalassaemia trait groups.

3.2.2. Interlaboratory Validation Cohort

An interlaboratory validation cohort that included 95 samples (65 α -thalassaemia trait, 30 β -thalassaemia trait) were subjected to the developed mathematical algorithm. β -thalassaemia trait was assigned as the positive group, and α -thalassaemia trait was assigned as the negative group, as shown in Table 4 and Supplementary S4. This $\alpha\beta$ -algorithm yielded a sensitivity of 92.85% (CI:0.75, 0.99) and specificity of 94.03% (CI:0.84, 0.98) in discriminating β -thalassaemia trait from the α -thalassaemia trait with a positive predictive accuracy of 86.67% (CI:0.68, 0.95) and negative predictive accuracy of 96.9% (CI:0.88, 0.99).

Table 4. Interlaboratory comparison of the $\alpha\beta$ -algorithm in differentiating α -thalassaemia trait and 30 β -thalassaemia trait.

| Algorithm | Prediction | Condition | | |
|----------------------------------|------------|-----------|---------|-------|
| | | Absent | Present | Total |
| (MNLMALESRET \times RDW) – MCH | Positive | 4 | 26 | 30 |
| | Negative | 63 | 2 | 65 |
| | Total | 67 | 28 | 95 |

4. Discussion

Distinguishing between various thalassaemia traits is essential in clinical decision-making as it will influence the treatment options and outcome of the patients [16]. However, the diagnosis of thalassaemia traits requires time and resources as various screening and confirmatory testing, including FBC, morphological review of blood smears, haemoglobin electrophoresis, and molecular genetic analysis to establish a reliable and robust diagnosis [17]. Several algorithms based on RBC indices have been proposed to differentiate iron deficiency anaemia and thalassaemia traits [17], including Mentzer Index [19], England & Fraser Index [20], Shine and Lal formula [21], Ehsani formula [22], Srivastava formula [23], Palestinian population [24], Green & King Index [25], and RDW Index [26].

In more recent studies, researchers have been exploring machine learning algorithm-based studies with advancements in data science [27,28]. In a study in Thailand, a web-based prediction tool for discrimination of thalassemia trait and IDA was developed using a machine learning algorithm. However, in this study, the authors did not delineate the subtypes of thalassemia but created a support vector machine (SVM) model to distinguish IDA from thal trait, named ThalPred [27]. One of the largest studies performed in Israel validated Shine's formula and in-house developed SVM formula. This study included a total of 64,586 subjects. Their SVM formula displayed high sensitivity (>98%) and >99.77% negative predictive value that is robust in distinguishing the β -thalassemia carrier from normal count subjects and iron-deficient women [28].

Previously, using the CPD parameters generated by Beckman Coulter DxH 800, Ng and the team proposed an algorithm to differentiate IDA from thalassaemia traits among its subjects. With a cut-off value of 23, the area under the curve (AUC) of 0.995 (95% CI of 0.99–1.00), the algorithm achieved a sensitivity of 97% and specificity of 99.1%. They suggested that no biochemical marker of iron status such as serum ferritin testing is required with this formula; hence, simplified diagnostic workup of IDA and thalassemia was proposed [16]. However, these formulas were devised from various populations with varying sensitivity and specificity, with none specific for differentiation of α and β -thalassaemia, without incorporating reticulocyte CPD parameters, as done in this present study [17,29,30].

The differentiation of α and β -thalassaemia based on the red blood cells parameter (Hb level, MCH, MCV, and RBC level) was performed in 2018 by R. Azma et al. in an LH 750 FBC analyser by Beckman Coulter (USA). A total of 299 subjects (α -thalassaemia traits, $n = 160$ and β -thalassaemia traits, $n = 139$) were included for this study. It showed that α -thalassaemia carriers have higher Hb level (α -thalassaemia carrier: 12.1 ± 1.31 ; β -thalassaemia carrier: 10.6 ± 1.8 , p -value: 0.00 *), MCV (α -thalassaemia carrier: 71.0 ± 6.1 ; β -thalassaemia carrier: 68.7 ± 7.2 , p -value: 0.044 *), MCH (α -thalassaemia carrier: 22.8 ± 2.3 ; β -thalassaemia carrier: 21.8 ± 2.7 , p -value: 0.056) RBC level (α -thalassaemia carrier: 5.3 ± 0.7 ; β -thalassaemia carrier: 4.9 ± 0.9 , p -value: 0.001 *), compared to β -thalassaemia carriers [31]. Although this study was not done based on CPD parameters, similar results were observed in our present study. The MCV and MCH parameters are shown to have notably higher values among subjects with α -thalassaemia trait compared to β -thalassaemia trait, in accordance with the findings from this study, as depicted in Table 1.

In this present study, we successfully developed an algorithm ($\alpha\beta$ -algorithm) based on reticulocyte CPD parameters that will serve as the downstream tool after FBC is performed to distinguish α versus β -thalassaemia trait, using Beckman Coulter Unicel DxH 800. The ROC-curve analysis revealed three parameters with the highest AUC, MN-LMALS-RET, MCH, and RDW, with good sensitivity and specificity (Figures 1a,b and 2). The $\alpha\beta$ -algorithm was devised by combining the aforementioned parameters ((MN-LMALS-RET \times RDW) – MCH) that resulted in higher AUC with better sensitivity and specificity to discriminate α from β -thalassaemia traits. The rationales for the $\alpha\beta$ -algorithm were MN-LMALS-RET, which measures the red blood cell volume in the reticulocyte channel through lower median angle light scattering [32], MCH indicates the amount of haemoglobin in a red blood cell [33], and RDW helps to measure variation in red blood cell size [12].

Our prospective study revealed the usefulness of the $\alpha\beta$ -algorithm as a downstream pipeline following FBC analysis, as depicted in Figure 4. Based on Figure 5, when this algorithm was evaluated in a cohort of various red cell disorders, there was an overlap between the subgroups, especially between the β -thalassemia trait and IDA groups. Nevertheless, the diagnostic pipeline for cases suspected as IDA differs in the Malaysian clinical practice guideline, so this overlap will not impact the role of this algorithm in the differential diagnosis of α and β -thalassemia. Interlaboratory validation showed that the findings are reproducible in another laboratory setting with similar Unicel DxH 800, hence reiterating the usefulness of the $\alpha\beta$ -algorithm in the screening of thalassemia cases, in particular, α

and β -thalassemia in geographical locations with a high prevalence of thalassaemias such as Malaysia and other Asian countries.

There are several limitations of this study: First, it included a relatively low number of α and β -thalassemia cases. Secondly, β -thalassemia group, the maximum value of MCV was 82 that overlaps with the minimum MCV of 80 in the healthy control group. In such cases, other RBC indices (RBC, MCH) and HPLC would aid in the differential diagnosis of β -thalassemia. Ideally, molecular analysis for both alpha and beta mutations needs to be performed on all samples so as not to miss out on the silent α or β thalassemia during screening. However, we did not perform molecular analysis on all cases because of the high testing cost and rarity of silent β -thalassemia [34,35].

We propose the vigorousness of this $\alpha\beta$ -algorithm be validated in a larger cohort in other geographical locations that exhibit high prevalence and incidence of α and β -thalassaemia traits. Based on this study, we recommend the FBC analysis is performed within 6 h of sampling following International Council for Standardisation in Haematology (ICSH) guidelines [36]. Apart from that, the transportation and storage of samples condition must meet the guidelines recommended by The Clinical and Laboratory Standard Institute (CLSI) [37]. Any deviation from adhering strictly to these guidelines may affect the sensitivity and specificity of this algorithm as CPD parameters are highly affected by the structural changes of the cells.

5. Conclusions

Devising an algorithm that accurately distinguishes α and β -thalassaemia traits using CPD parameters derived from FBC analyser is essential in the pipeline of large population-based screening carried out in Malaysia as high prevalence of these inherited disorders are reported. Implementing the $\alpha\beta$ -algorithm in the screening laboratory will promote cost optimisation as only suspected subjects will be included in the downstream pipeline before performing haemoglobin electrophoresis and/or further specific genetic testing, as depicted in Figure 4. The applicability of the $\alpha\beta$ -algorithm is relatively straightforward without involving any additional cost to the diagnosis and requires no sophisticated analysis and expertise. Hence, we strongly recommend adopting this $\alpha\beta$ -algorithm in all screening laboratories with access to similar CPD parameters nationwide and other geographical regions with a high prevalence of thalassaemia-related disorders [38,39].

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/diagnostics11112163/s1>. Supplementary Materials are available as the following files: Supplementary S1 Description of CPD Parameters.docx, Supplementary S2 ROC analysis.xlsx, Supplementary S3 in-house cohort.xlsx, and Supplementary S4 Interlaboratory Cohort.xls.

Author Contributions: A.A. developed the study design and prepared the manuscript. A.A. collected retrospective patient data. S.S. performed the interlaboratory prospective study and analysed the samples, and T.W.Y. collected the prospective study cohort data and handled the logistics for the interlaboratory comparison cohort. A.A. performed the statistical analysis and prepared all figures and tables. J.S., Q.-Y.Z. and R.H. oversaw the study and supervised. R.H. critically reviewed the technical aspect of the study. All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Medical Research Ethics Committee of the Ministry of Health of Malaysia [Research ID 10-277-5480 (date of approval: 26 October 2010 and NMRR 17-2708-38327 (date of approval: 19 December 2017))].

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data used in this study are available as Supplementary Materials (Supplementary S1–S4).

Acknowledgments: We gratefully acknowledge the following: Ministry of Health Malaysia, Gribbles Pathology, Malaysia and School of Medical and Health Sciences, Universiti Sains Malaysia, for data access and publication permission. Special thanks to Straits Scientific Malaysia for their scientific support and Wan Siti Nurbaya for technical assistance. By the graciousness of Almighty God, this project was successfully completed.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Joint WHO-TIF Committee. Management of haemoglobin disorders. In Proceedings of the Joint WHO-TIF Meet, Nicosia, Cyprus, 16–18 November 2008.
2. Williams, T.N.; Weatherall, D.J. World Distribution, Population Genetics, and Health Burden of the Hemoglobinopathies. *Cold Spring Harb. Perspect. Med.* **2012**, *2*, a011692. [CrossRef] [PubMed]
3. Azma, R.Z.; Ainoon, O.; Hafiza, A.; Azlin, I.; Noor Farisah, A.R.; Nor Hidayati, S.; Noor Hamidah, H. Molecular characteristic of alpha thalassaemia among patients diagnosed in UKM Medical Centre. *Malays. J. Pathol.* **2014**, *36*, 27–32. [PubMed]
4. Ibrahim, H.; Hassan, A.; George, E.; Sim, G.A.; Othman, I.S.; Mudla, I.; Hassan, H. Clinical Practice Guidelines: Management of Transfusion Dependent Thalassaemia. Available online: <http://www.moh.gov.my/penerbitan/CPG2017/8318.pdf> (accessed on 10 September 2021).
5. Wahab, I.A.; Naznin, M.; Nora, M.Z.; Suzanah, A.R.; Zulaiho, M.; Faszrul, A.R.A.; Kamaruzaman, W.S. Thalassaemia: A study on the perception of patients and family members. *Med. J. Malays.* **2011**, *66*, 326–334.
6. Higgs, D.R. The molecular basis of α -thalassemia. *Cold Spring Harb. Perspect. Med.* **2013**, *3*, a011718. [CrossRef] [PubMed]
7. Rosnah, B.; Rosline, H.; Zaidah, A.W.; Noor Haslina, M.N.; Marini, R.; Shafini, M.Y.; Nurul Ain, F.A. Detection of Common Deletional Alpha-Thalassemia Spectrum by Molecular Technique in Kelantan, Northeastern Malaysia. *ISRN Hematol.* **2012**, *2012*, 462969. [CrossRef] [PubMed]
8. Galanello, R.; Origa, R. Beta-thalassemia. *Orphanet J. Rare Dis.* **2010**, *5*, 11. [CrossRef] [PubMed]
9. George, E.; Ann, T.J.A.M. Genotype-phenotype diversity of beta-thalassemia in Malaysia: Treatment options and emerging therapies. *Med. J. Malaysia* **2010**, *65*, 256–260. [PubMed]
10. Sabath, D.E. Molecular Diagnosis of Thalassemias and Hemoglobinopathies. *Am. J. Clin. Pathol.* **2017**, *148*, 6–15. [CrossRef]
11. Pornprasert, S.; Saoboontan, S.; Wiengkum, T. Hemoglobin Constant Spring (Hb CS) Missed by HPLC in an Hb E Trait Pregnancy Resulting in Hb H-CS Disease in a Thai Girl: Utility of Capillary Electrophoresis. *Indian J. Hematol. Blood Transfus.* **2016**, *32*, 254–257. [CrossRef] [PubMed]
12. Muncie, H.L.; Campbell, J. Alpha and beta thalassemia. *Am. Fam. Physician* **2009**, *80*, 339–344.
13. Lopez, R.S.; Solenthaler, M.J.C. *Unicel DxH Series Case Book*; Beckman Coulter, Inc.: Brea, CA, USA, 2016.
14. Ambayya, A.; Sasmita, A.O.; Zhang, Q.-Y.; Ting, A.S.; Meng, C.K.; Sathar, J.; Yegappan, S. Reference interval establishment of full blood count extended research parameters in the multi-ethnic population of Malaysia. *Med. J. Malays.* **2019**, *74*, 395.
15. Lowry, L. VassarStats: Website for Statistical Computation. Available online: <http://vassarstats.net/clin1.html> (accessed on 16 September 2021).
16. Ng, E.H.Y.; Leung, J.H.W.; Lau, Y.S.; Ma, E.S.K. Evaluation of the new red cell parameters on Beckman Coulter DxH800 in distinguishing iron deficiency anaemia from thalassaemia trait. *Int. J. Lab. Hematol.* **2015**, *37*, 199–207. [CrossRef]
17. Hafeez Kandhro, A.; Shoombuatong, W.; Prachayasittikul, V.; Nuchnoi, P. New Bioinformatics-Based Discrimination Formulas for Differentiation of Thalassemia Traits From Iron Deficiency Anemia. *Lab. Med.* **2017**, *48*, 230–237. [CrossRef] [PubMed]
18. Ambayya, A.; Su, A.T.; Osman, N.H.; Nik-Samsudin, N.R.; Khalid, K.; Chang, K.M.; Sathar, J.; Rajasuriar, J.S.; Yegappan, S. Haematological Reference Intervals in a Multiethnic Population. *PLoS ONE* **2014**, *9*, e91968. [CrossRef]
19. Mentzer, W. Differentiation of iron deficiency from thalassaemia trait. *Lancet* **1973**, *301*, 882. [CrossRef]
20. England, J.M.; Fraser, P. Differentiation of iron deficiency from thalassaemia trait by routine blood-count. *Lancet* **1973**, *301*, 449–452. [CrossRef]
21. Shine, I.; Lal, S. A strategy to detect β -thalassaemia minor. *Lancet* **1977**, *309*, 692–694. [CrossRef]
22. Ehsani, M.A.; Shahgholi, E.; Rahiminejad, M.S.; Seighali, F.; Rashidi, A. A new index for discrimination between iron deficiency anemia and beta-thalassemia minor: Results in 284 patients. *Pak. J. Biol. Sci.* **2009**, *12*, 473–475. [CrossRef] [PubMed]
23. Srivastava, P.C.; Bevington, J.M. Iron deficiency and-or thalassaemia trait. *Lancet* **1973**, *1*, 832. [CrossRef]
24. Sirdah, M.; Tarazi, I.; Al Najjar, E.; Al Haddad, R. Evaluation of the diagnostic reliability of different RBC indices and formulas in the differentiation of the beta-thalassaemia minor from iron deficiency in Palestinian population. *Int. J. Lab. Hematol.* **2008**, *30*, 324–330. [CrossRef]
25. Green, R.; King, R. A new red cell discriminant incorporating volume dispersion for differentiating iron deficiency anemia from thalassemia minor. *Blood Cells* **1989**, *15*, 481–491.
26. Ricerca, B.M.; Storti, S.; D'Onofrio, G.; Mancini, S.; Vittori, M.; Campisi, S.; Mango, G.; Bizzi, B. Differentiation of iron deficiency from thalassaemia trait: A new approach. *Haematologica* **1987**, *72*, 409–413. [PubMed]
27. Laengsri, V.; Shoombuatong, W.; Adirojananon, W.; Nantasenamat, C.; Prachayasittikul, V.; Nuchnoi, P. ThalPred: A web-based prediction tool for discriminating thalassemia trait and iron deficiency anemia. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 1–14. [CrossRef]

28. Roth, I.L.; Lachover, B.; Koren, G.; Levin, C.; Zalman, L.; Koren, A. Detection of β -thalassemia carriers by red cell parameters obtained from automatic counters using mathematical formulas. *Mediterr. J. Hematol. Infect. Dis.* **2018**, *10*, e2018008. [[CrossRef](#)] [[PubMed](#)]
29. Mehdi, S.; Al Dahmash, B. A comparative study of hematological parameters of α and β thalassemias in a high prevalence zone: Saudi Arabia. *Indian J. Hum. Genet.* **2011**, *17*, 207. [[CrossRef](#)] [[PubMed](#)]
30. Koh, D.X.R.; Raja Sabudin, R.Z.A.; Mohd Yusoff, M.; Hussin, N.H.; Ahmad, R.; Othman, A.; Ismail, E. Molecular Characterisation of α - and β -thalassaemia among Indigenous Senoi Orang Asli Communities in Peninsular Malaysia. *Ann. Hum. Genet.* **2017**, *81*, 205–212. [[CrossRef](#)]
31. Azma, R.; Jalil, N.; Ma, M.-H.; Skk, J.; Sa, F.-H.; Nm, N.-N.; Am, F.-A.; Alauddin, H.; Ithnin, A.; Othman, A. A Comparative Study of Red Blood Cell Parameters of Alpha and Beta Thalassaemia Patients Diagnosed in a University Hospital in Cheras, Malaysia. *ARC J. Hematol.* **2018**, *3*, 23–27.
32. Jung, Y.-J.; Kim, J.-H.; Park, Y.-J.; Kahng, J.; Lee, H.; Lee, K.-Y.; Kim, M.Y.; Han, K.; Lee, W. Evaluation of cell population data on the UniCel DxH 800 Coulter Cellular Analysis system as a screening for viral infection in children. *Int. J. Lab. Hematol.* **2012**, *34*, 283–289. [[CrossRef](#)] [[PubMed](#)]
33. Jolobe, O.M.P. More on identification of iron deficiency when it coexists with chronic inflammation. *J. Clin. Pathol.* **2011**, *64*, 833. [[CrossRef](#)]
34. Ristaldi, M.S.; Murru, S.; Loudianos, G.; Casula, L.; Porcu, S.; Pigheddu, D.; Fanni, B.; Sciarratta, G.V.; Agosti, S.; Parodi, M.I.; et al. The C-T substitution in the distal CACCC box of the β -globin gene promoter is a common cause of silent β thalassaemia in the Italian population. *Br. J. Haematol.* **1990**, *74*, 480–486. [[CrossRef](#)] [[PubMed](#)]
35. Cao, A.; Galanello, R. Beta-thalassemia. *Genet. Med.* **2010**, *12*, 61–76. [[CrossRef](#)] [[PubMed](#)]
36. Briggs, C.; Culp, N.; Davis, B.; D'Onofrio, G.; Zini, G.; Machin, S.J. ICSH guidelines for the evaluation of blood cell analysers including those used for differential leucocyte and reticulocyte counting. *Int. J. Lab. Hematol.* **2014**, *36*, 613–627. [[CrossRef](#)]
37. Calam, R.R.; Bessman, J.D.; Smith, S.S.; Szamosi, D.I.; Warunek, D.J.; Wiseman, J.D. Procedures for the Handling and Processing of Blood Specimens. In *Approved Guideline*, 3rd ed.; NCCLS: Wayne, PA, USA, 2004; Volume 24, ISBN 1562385550.
38. Chiruka, S.; Darbyshire, P. Management of thalassaemia. *Paediatr. Child Health* **2011**, *21*, 353–356. [[CrossRef](#)]
39. Mohd Pauzy, L.H.; Esa, E.; Mahani Mokhri, N.; Yusoff, Y.M.; Jamaludin, N.A.; Zakaria, Z. Thalassemia Distribution Based on Screening Programs in the Population of the East Malaysian State of Sabah. *J. Blood Disord. Transfus.* **2018**, *9*, 1. [[CrossRef](#)]