

Article



Novel Decision Tool for More Severe α -Thalassemia Genotypes Screening with Functional Loss of Two or More α -Globin Genes: A Diagnostic Test Study

Patricia F. R. Siqueira ^{1,2}, Marcos K. Fleury ³, Robéria M. Pontes ^{1,4}, Renata S. P. Silva ⁵, Elaine S. Costa ^{1,2} and Marcelo G. P. Land ^{1,2,*}

- ¹ Multidisciplinary Laboratory, Instituto de Puericultura e Pediatria Martagão Gesteira (IPPMG),
- Faculty of Medicine, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro 21941-912, Brazil
 ² Internal Medicine Postgraduate Program, Faculty of Medicine, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro 21941-913, Brazil
- ³ Department of Clinical and Toxicological Analysis, Faculty of Pharmacy, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro 21941-599, Brazil
- ⁴ Translational Research Laboratory, Children's Hospital of Brasilia José Alencar, Brasilia 70684-831, Brazil
- ⁵ Hematology Department, Instituto de Puericultura e Pediatria Martagão Gesteira (IPPMG), Faculty of Medicine, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro 21941-912, Brazil
- Correspondence: land.marcelo@ippmg.ufrj.br; Tel.: +55-(21)-99649-9155

Abstract: After the exclusion of iron deficiency and β -thalassemia, molecular research for α -thalassemia is recommended to investigate microcytic anemia. Aiming to suggest more efficiently the molecular analysis for individuals with a greater chance of having a symptomatic form of the disease, we have developed and validated a new decision tool to predict the presence of two or more deletions of α -thalassemia, increasing considerably the pre-test probability. The model was created using the variables: the percentage of HbA₂, serum ferritin and mean corpuscular volume standardized by age. The model was trained in 134 patients and validated in 160 randomly selected patients from the total sample. We used Youden's index applied to the ROC curve methodology to establish the optimal odds ratio (OR) cut-off for the presence of two or more α -globin gene deletions. Using the OR cut-off of 0.4, the model's negative predictive value (NPV) was 96.8%; the cut-off point accuracy was 85.4%; and the molecular analysis pre-test probability increased from 25.9% to 65.4% after the use of the proposed model. This tool aims to assist the physician in deciding when to perform molecular studies for the diagnosis of α -thalassemia. The model is useful in places with few financial health resources.

Keywords: *α*-thalassemia; microcytic anemia; molecular diagnosis; binomial logistic regression; prediction; ROC curve; sensitivity; specificity

1. Introduction

The molecular basis of α -thalassemia is related to a reduction (α^+ -thalassemia) or absence (α^0 -thalassemia) from the synthesis of α -globin chains. The phenotypic differences in this disease are due to the number of α -globin genes affected by some mutation. In general, these mutations are caused by deletions of one or both α -globin genes from each chromosome 16 [1,2]. Its prevalence varies considerably around the world, occurring in high frequency throughout tropical and subtropical regions [2,3]. The estimated prevalence of α^+ -thalassemia is at least 20% of the world population [4]. The prevalence of α -thalassemia carriers in Brazil is not fully known. We can attribute this to the great territorial dimensions of the country and the variability of the genetic constitution of its population, in addition to the lack of availability of molecular testing for population testing in many places. It is known, however, that the - $\alpha^{3.7}$ deletion in heterozygosis, or homozygosis, is the most prevalent, due to the great miscegenation of the population with important African genetic contributions [5–7]. Studies in several Brazilian regions report variable frequencies of α^+ -thalassemia resulting from the - $\alpha^{3.7}$ deletion, ranging



Citation: Siqueira, P.F.R.; Fleury, M.K.; Pontes, R.M.; Silva, R.S.P.; Costa, E.S.; Land, M.G.P. Novel Decision Tool for More Severe α-Thalassemia Genotypes Screening with Functional Loss of Two or More α-Globin Genes: A Diagnostic Test Study. *Diagnostics* **2022**, *12*, 3008. https://doi.org/10.3390/ diagnostics12123008

Academic Editors: Rosline Hassan and Muhammad Farid Johan

Received: 11 October 2022 Accepted: 22 November 2022 Published: 1 December 2022

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



Copyright: © 2022 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/). from 0.7 to 20% [5,6,8–11]. After screening for the presence of microcytosis and/or hypochromia, the $-\alpha^{3.7}$ deletion is detected in around 50% of individuals [12,13]. The detection of these deletions is usually performed by multiplex polymerase chain reaction (multiplex-PCR) [3,14]. Other molecular techniques, such as multiplex ligation-dependent probe amplification (MLPA), gene sequencing, or amplification refractory mutations (ARMS), may be required for the investigation of unknown deletions or point mutations in α -globin genes [15].

In routine medical practice, laboratory tests exclude other causes of microcytic and/or hypochromic anemia before investigating major α -thalassemia deletions. For the initial investigation, it is recommended to study the iron kinetics, hematimetric indices, peripheral blood smears analysis, detection of H hemoglobin (HbH) inclusions, reticulocyte count, and analysis of the hemoglobin profile [2,16,17]. In the presence of microcytosis and/or hypochromia, after excluding iron deficiency and β -thalassemia minor, molecular tests for α -thalassemia by PCR are recommended [18,19]. However, the availability of molecular analysis of patients in primary care hospitals is not a reality in developing or underdeveloped countries. For this reason, using inexpensive and efficient screening methods to select individuals who require molecular testing for α -thalassemia can reduce health system costs.

In the present study, we developed and validated a binomial logistic regression (BLR) model using the results of laboratory tests, usually applied as screening, to decide when to perform the molecular technique. The use of this model aims to increase the molecular analysis pre-test probability of selecting patients before performing multiplex-PCR for the main deletions in the α -globin gene. Thus, the number of patients tested by PCR can considerably decrease, generating significant cost savings and reducing the need for blood sampling, which is especially important in infancy.

For this purpose, we performed the detection of the deletions $-\alpha^{3.7}$ and $-\alpha^{4.2}$ in 294 patients who underwent analysis for investigation of α -thalassemia at the Clinical Analysis Laboratory of Faculty of Pharmacy (LACFAR) at the Federal University of Rio de Janeiro (UFRJ). Four of these patients presented a characteristic peak of HbH in HPLC, in addition to HbH in erythrocytes staining with brilliant cresyl blue. The deletions $-(\alpha)^{20.5}$, $-^{SEA}$, and $-^{MED}$ were also tested in these patients.

2. Materials and Methods

This study was previously approved by the Ethics Committee of the IPPMG hospital (Comitê de Ética em Pesquisa do Instituto de Puericultura e Pediatria Martagão Gesteira— Federal University of Rio de Janeiro). The first acceptance (Acceptance Number 38/11, Approval Date: 14 June 2011) was granted to evaluate the prevalence of α -thalassemia in children with microcytosis in a pediatric hospital in Rio de Janeiro. For the construction and validation of the mathematical model, we requested approval from the same Ethics Committee to use the data of other patients evaluated for α -thalassemia in the Clinical Analysis Laboratory of the Faculty of Pharmacy at the Federal University of Rio de Janeiro (UFRJ) (Acceptance Number 847.419—Approval Date: 28 October 2014).

2.1. Analytical Methods

This was a retrospective observational study of a cohort of patients (adults, adolescents, and children) from November 2010 to July 2015. Peripheral blood samples of 294 patients presenting microcytosis, or suspected of being α -thalassemia carriers, were submitted to PCR for detection of the deletions $-\alpha^{3.7}$ and $-\alpha^{4.2}$ for the investigation of α^+ -thalassemia, as well as for a complete blood count (CBC), high-performance liquid chromatography (HPLC), reticulocyte count, research of HbH inclusions in erythrocytes, measurements of iron status: serum iron, total iron-binding capacity (TIBC), serum ferritin, transferrin saturation (TS), and biochemical measurements such as lactate dehydrogenase (LDH), direct bilirubin (DBil), indirect bilirubin (IBil), and total bilirubin (TBil). Four patients with HbH inclusions in their erythrocytes and the HbH characteristic peak in the HPLC were also tested for other deletions, $-^{MED}$, $-^{SEA}$, and $-(\alpha)^{20.5}$, related to α^0 -thalassemia.

These results were recorded in a database associated with the following variables: age (years), gender, number of α -globin genes deleted, CBC including red cell count (RCC), hemoglobin (Hb), hematocrit (Ht), mean corpuscular volume (MCV), MCV standardized by age, mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), red blood cell distribution width (RDW), percentage of Hb A₁, Hb A₂, and fetal Hb, serum iron, serum ferritin, TIBC, TS, LDH, DBil, IBil, and TBil.

The CBC was performed on an automated cell counter, the ABX Pentra 60C+ (Horiba, Montpellier, France). The serum ferritin levels were determined by an enzyme immunoassay with the Unicel DxI 600 (Beckman Coulter, Brea, CA, USA). The blood of all the patients was analyzed by cation-exchange high-performance liquid chromatography (HPLC Variant, Bio-Rad Laboratories, Hercules, CA, USA) for the quantification of hemoglobin (Hb) A₁, A₂, and fetal and variant hemoglobin, when identified by the methodology. All the samples were analyzed for the presence of HbH inclusion bodies and the reticulocyte count after staining the peripheral blood cells with 1% brilliant cresyl blue for 1 h at 37 °C. The other biochemical measurements and iron status were performed but were not used to build the BLR model.

The genomic DNA was isolated from the peripheral blood leucocytes using a Biopur Mini Spin Plus kit (Biometrix Diagnóstica, Curitiba, Brazil). In all the peripheral blood samples, the molecular characterization of α -thalassemia was determined by multiplex polymerase chain reaction (multiplex-PCR) using previously designed primers to detect the deletions $-\alpha^{3.7}$, $-\alpha^{4.2}$ and the not deleted α^2 gene [3]. Each set of reactions included positive and negative controls. Furthermore, the deletions $-(\alpha)^{20.5}$, $-^{SEA}$, and $-^{MED}$ were also investigated in four patients who presented with HbH in the erythrocytes staining with brilliant cresyl blue, associated with the characteristic peak of HbH in the HPLC.

Each 25 µL reaction contained 20 mmol/L Tris-HCl (pH 8.4), 50 mmol/L KCl, 2.0 mmol/L MgCl₂, 1,0 mol/L betaine (Sigma-Aldrich, St. Louis, MO, USA), 0.5 mmol/L of each dNTP, 15 pmoles of each specific primer for the deletions $-\alpha^{3.7}$, $-\alpha^{4.2}$ and the not deleted α^2 gene, 2.5 units of Taq DNA Polymerase (Invitrogen, Life Technologies, São Paulo, Brazil), and 100 ng of genomic DNA. When the multiplex-PCR was performed for the investigation of other deletions $(-^{MED}, -^{SEA}, \text{ and } -(\alpha)^{20.5})$, 5 pmoles of each specific primer were added in this reaction, maintaining the final reaction volume of 25 µL. The reactions were carried out on a thermal cycler Veriti (Applied Biosystems, Foster City, CA, USA) with an initial denaturation at 95 °C for 5 min, 30 cycles of 98 °C for 45 s, 63 °C for 1 min 30 s, 72 °C for 2 min 15 s, and a final extension at 72 °C for 5 min. After amplification, 15 µL of the product was electrophoresed through a 1% agarose gel in 1 X TBE for 1 h. The agarose gel was stained with ethidium bromide and visualized on an ultraviolet transilluminator.

Due to the large variation of the reference values of the MCV related to patient age, we created a standardization of MCV for all patients relative to the minimum reference value for adults (80 fL) [13]. This strategy permitted the MCV to be homogeneously analyzed in the different age ranges studied. The MCV minimum expected values for the age were previously described [16]. The normal values of the MCV standardized by age were the same as those used for the adult population (80–100 fL).

The following formula was used to calculate the MCV standardized by age:

MCV standardized by age = $\frac{MCV \text{ observed}}{MCV \text{ minimum expected for the age}} \times 80$

2.2. Binomial Logistic Regression (BLR) Model Building

The statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) software program, version 20.0, SPSS Inc., Chicago, IL, USA.

The group of 294 patients used for the construction and validation of the BLR model was named the total sample. The training set consisted of 134 individuals who were selected randomly by the SPSS software from the data of the 294 patients.

The univariate and multivariate analyses were performed with the training set to identify the most significant variables for the development of the BLR model.

Several variables were evaluated using univariate analysis. The variables with *p*-value ≤ 0.2 were selected for the multivariate analysis. We first evaluated the collinearity between the variables, using the Pearson correlation coefficient, Spearman correlation coefficient, and Kendall's Tau test. The candidate variables were included in a full model. The parsimonious model was built through a process of backward elimination using the likelihood-ratio test. This final model included the serum ferritin, MCV standardized by age and percentage of Hb A₂.

A receiver operating characteristic (ROC) analysis was used to determine the optimal odds ratio (OR) cut-off point to predict the presence of two or more deletions related to α -thalassemia. Based on Youden's index applied to the ROC curve, the cut-off (OR > 0.4 or probability > 28.6%) achieved the highest possible sensitivity and specificity. The same cut-off was used for the analysis of the validation group.

The validation set (n = 160) was composed of other randomly selected individuals from the total sample. The performance of the BLR model in both sets was also evaluated in terms of the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio and negative likelihood ratio, the area under the curve (AUC), and cut-off point accuracy.

3. Results

3.1. Univariate Analysis

The variables evaluated by the univariate analysis are reported in Table 1. The variables with a p-value < 0.2 appear in bold.

Variables	OP	95% Confidence	" Value	
	OK	Lower Bound	Upper Bound	<i>p</i> -value
RCC	8.813	3.365	23.087	<0.001
Hemoglobin	0.807	0.598	1.089	0.161
Hematocrit	0.986	0.889	1.095	0.797
MCV	0.860	0.804	0.919	<0.001
MCV standardized by age	0.814	0.751	0.883	<0.001
MCH	0.618	0.508	0.752	<0.001
MCHC	0.596	0.444	0.800	0.001
RDW	1.040	0.909	1.189	0.567
Reticulocyte Count	1.005	0.995	1.015	0.343
% Hb A ₁	1.041	0.989	1.097	0.124
% Hb A ₂	0.393	0.176	0.878	0.023
% Hb F	0.754	0.471	1.207	0.240
Serum Iron	1.019	1.006	1.032	0.003
TIBC	0.996	0.990	1.003	0.264
Serum Ferritin	1.007	0.999	1.014	0.074
Transferrin Saturation	1.002	0.992	1.012	0.688
LDH	0.997	0.993	1.000	0.085
Indirect Bilirubin	1.737	0.300	10.043	0.538

Table 1. Univariate analysis of the variables tested for the BLR model building (n = 134).

3.2. Development of BLR Model

The variables included in the final parsimonious model were as follows: the MCV standardized by age, serum ferritin, and percentage of Hb A₂. The parameters of the multivariate analysis of these variables are shown in Table 2.

Table 2. Parsimonious BLR model for prediction of genotypes with 2 or more α -globin genes deletions.

Two or More α-Globin Gene Deletions	В	<i>p</i> -Value	OR (Exp(B))	IC95
Intercept	21.905	< 0.001		
MCV standardized by age	-0.284	< 0.001	0.753	0.676-0.839
Serum Ferritin	0.024	< 0.001	1.024	1.011-1.037
% Hb A ₂	-1.142	0.006	0.319	0.142-0.715

The resulting BLR model after fitting to the training data can be expressed as the OR and as a probability. Table 3 shows how the value of OR and probability are calculated using the proposed model.

Table 3. Summary of the BLR Model for prediction of genotypes with 2 or more α -globin gene deletions.

	Prediction Equation		
OR	=exp [21.905 $-$ 0.284 (MCV standardized by age(fL)) + 0.024 (serum ferritin(ng/mL)) $-$ 1.142 (Hb A ₂ (%))]		
Probability (<i>P</i> α -thal with 2 or more α -globin genes deletions)	=OR/(1 + OR)		

3.3. Determination of the Best Cut-Off Point and Its Performance

To determine the best cut-off point to the OR to predict the outcome, we used the ROC curve. The optimal cut-off point to predict two or more α -globin gene deletions was OR > 0.4.

We used the OR > 0.4 cut-off point to create the predicted group variable. We compared the predicted group with the gold standard (two or more α -globin gene deletions detected by multiplex-PCR). Figure 1 shows that the BLR model gave an AUC of 0.927 with an accuracy of the cut-off point of 85.1% when setting the cut-off point at OR > 0.4 (or probability >28.6%). We compared the predicted group with the gold standard (two or more α -globin gene deletions detected by multiplex-PCR). This model performed a sensitivity of 88.6%, specificity of 83.8%, PPV of 66.0%, and NPV of 95.4% (Supplementary Table S1).



Figure 1. Determination of the optimal cut-off point using the ROC curve.

3.4. Validation Process

Table 4 shows the genotypic characteristics, the age distribution, and the variables associated with the BLR model in both groups (training and validation data sets). There was no significant difference between the groups. The variables of the BLR model and age were analyzed using the Mann–Whitney test or Student's *t*-test. The comparison between the genotypes of the α -globin gene of the two groups and sex was performed using the chi-square test or Fisher's exact test.

	Training Set Frequency	Validation Set Frequency			
	(%)/Median (IQR)/Mean (SE)	(%)/Median (IQR)/Mean (SE)	<i>p</i> -Value		
	n = 134	n = 160			
Age (years)	9.5 (10.3)	9.5 (10.3) 8.5 (10.1)			
Gender					
Male	67 (50.0%)	86 (53.8%)	0.559 ***		
Female	67 (50.0%)	74 (46.2%)			
The Genotype of the α -Globin Gene					
αα/αα	53 (39.6%)	69 (43.1%)			
$-\alpha^{3.7}/\alpha\alpha$	46 (34.3%)	50 (31.3%)	0.801 ***		
$-\alpha^{3.7}/-\alpha^{3.7}$	33 (24.6%)	41 (25.6%)	0.001		
$-\alpha^{3.7}/-$ SEA	1 (0.7%)	1 (0.7%) —			
$\alpha \alpha / - SEA$	1 (0.7%)	_			
Variables of the BLR model					
MCV standardized by age (fL)	76.3 (± 0.71)	75.6 (± 0.63)	0.470 **		
Serum Ferritin (ng/mL)	32.1 (35.7)	31.2 (36.2)	0.726 *		
Hb A ₂ (%)	2.8 (0.4)	2.8 (0.4)	0.514 *		
4 3 4 TATI		T :			

Table 4. Basic population characteristics and key BLR model variables.

* Mann-Whitney Test; ** Student's *t*-test; *** Fisher's Exact Test.

The model validation was performed on another 160 patients randomly selected from the total sample, using the OR > 0.4 cut-off point to create the predicted group variable. We compare the predicted group with the gold standard (two or more α -globin gene deletions detected by multiplex-PCR). This model gave an AUC of 0.902 with an accuracy of the cut-off point of 85.6% using the same probability cut-off of 28.6% for the validation of the training set. This model performed a sensitivity of 95.1%, specificity of 82.4%, PPV of 65.0%, and NPV of 98.0% (Supplementary Table S2).

Applying the same cut-off point in the total sample (n = 294), results were obtained that were very similar to those of the training and validation sets, which demonstrates the good randomization of the cohort, as can be seen in Table 5 (Supplementary Table S3).

Table 5. Performance evaluation of the proposed model.

	Training Set		Validation Set		Total Sample	
	(n = 134)		(n = 160)		(n = 294)	
		C.I. (95%)		C.I. (95%)		C.I. (95%)
Sensitivity	88.6%	73.3–96.8%	95.1%	83.5–99.4%	92.1%	83.6-97.1%
Specificity	83.8%	75.1-90.5%	82.4%	74.3-88.7%	83.0%	77.4-87.8%
Positive Likelihood Ratio	5.48	3.45 -8.72	5.39	3.63-8.00	5.43	4.02-7.33
Negative Likelihood Ratio	0.14	0.05-0.34	0.06	0.02-0.23	0.10	0.04-0.21
Disease Prevalence	26.1%	18.9-34.4%	25.6%	19.1–33.1%	25.9%	20.9-31.3%
Positive Predictive Value	66.0%	54.9-75.5%	65.0%	55.6-73.4%	65.4%	58.3-71.9%
Negative Predictive Value	95.4%	89.2-98.1%	98.0%	92.7–99.5%	96.8%	93.3–98.5%
Cut-off Point Accuracy	85.1%	77.9–90.6%	85.6%	79.2–90.7%	85.4%	80.8-89.2%
Global Accuracy (AUC)	0.927	0.883-0.970	0.902	0.853-0.950	0.914	0.881-0.947

3.5. Improving the Decision-Making Process

The pre-test probability, without using the model, calculated by the percentage of positive patients for two or more deletions tested by PCR, was 26.1% (35/134) in the training, 25.6% (41/160) in the validation set, and 25.9% (76/294) in the total sample. After the application of the BLR model (selecting molecular analysis only for the patients considered positive), the probability pre-test increased to 66.0% (31/47), 65.0% (39/60), and 65.4% (70/107), respectively. In the model established, the percentage of false negatives was 1.3–3.0%.

Using our model, about two-thirds of patients with suspected α -thalassemia would not need multiplex-PCR. We propose an algorithm for the evaluation of all patients suspected of being α -thalassemia carriers, to be applied in clinical practice (Supplementary Excel worksheet S4). If the result of the model was negative, the patient can be reassessed at another time, maybe a year later, if the suspicion persists, as shown in Figure 2.



Figure 2. Algorithm to use the BRL model for investigation of α -thalassemia.

4. Discussion

Several algorithms based on RBC indices have been proposed to aid in the differential diagnosis of iron deficiency anemia (IDA) and thalassemia traits. [20–24]. Most of them were constructed to distinguish IDA from β -thalassemia [24–27]. An exception is the Huber-Herklotz index, developed to differentiate cases of IDA from α -thalassemia using the parameters of complete blood cell counts [28–30]. Recently, some groups have developed models that allow the distinction of α -thalassemia from β -thalassemia and iron deficiency. In China, a model was described associating some indices of red blood cells with Hb A₂ levels to differentiate β -thalassemia, α -thalassemia, and iron deficiency [31]. An algorithm has been developed to differentiate between α -thalassemia and β -thalassemia traits ($\alpha\beta$ -algorithm), especially in geographic regions with large numbers of individuals with these inherited blood disorders [32].

The diagnosis of thalassemia traits requires time and resources, with the performance of several screenings and specific exams. Beta thalassemia is usually identified in the laboratory by the observation of Hb A₂ greater than 3.5%, although, sometimes, when there is an association with iron deficiency or coinheritance with α - or δ -thalassemia or depending on the type of mutation in the β -globin gene, the Hb A₂ can be detected in the normal concentration [15,31]. On the other hand, the definitive diagnosis of α -thalassemia

is only possible using molecular techniques to identify α -globin gene mutations, in addition to other screening tests.

The most frequent α -thalassemia-related mutations detected in the Brazilian population are the $-\alpha^{3.7}$ and $-\alpha^{4.2}$ deletions, due to the African heritage, which is very present in this population [5,7,9,10,33–35]. These mutations, when in heterozygosity, affect one α -gene, making this individual a silent carrier who is asymptomatic and does not need treatment. To optimize molecular testing by identifying the patients with a higher chance of having two or more α -globin genes mutated, we created this model. This is the first study to develop and validate a BLR model to predict the OR that a patient has two or more deletions related to α -thalassemia.

We know that, under ideal conditions, molecular testing for α -thalassemia should be performed in all suspected patients because it is the gold standard for this diagnosis. However, the financial limitations of the health systems of poorer populations make us look for alternatives to optimize this diagnosis, to prioritize the investigation of individuals who may have impairment of a larger number of α -globin genes, and who have greater clinical impacts. Besides the excessive cost of the molecular technique, it is important to emphasize the need for technical conditions and qualified personnel for conducting this analysis, and all other tests used for the preliminary identification of thalassemia carriers. If we can distinguish individuals who have two or more α -genes deleted using the proposed method, this can generate savings for the health systems in lower and middle-income countries.

The model developed with the training set showed excellent accuracy (AUC = 0.927, accuracy of 85.1%). It also demonstrated good discriminatory power in the validation set (AUC = 0.902, accuracy of 85.6%) and in the total sample (n = 294). This model also presented high sensitivity and NPV, which allows us to ensure, with a higher level of confidence, that the individual who has a negative result has a minor OR of having two or more deletions related to α -thalassemia. In our total sample, 63.6% (187/294) of the patients would not need an immediate PCR for the most frequent α -thalassemia deletions. This contributes to the decision-making process in the diagnostic investigation, benefiting both the physician and the patient. The physician will have a simple tool that will increase the molecular analysis pre-test probability, based on the most easily available screening tests. Hence, using the BLR model, almost two-thirds of the patients will be spared the exam, which will directly impact the health costs and, additionally, diminish the psychological stress related to the test.

Those patients showing laboratory characteristics strongly related to α -thalassemia will be forwarded for PCR analysis. The model has a high NPV. It indicates that patients with OR \leq 0.4 have a high chance of preserving three or four functioning α -genes, excluding a clinically significant disease. It is important to note that the results could be different, and, in reality, better when applied to the general population since the model originated from a specialized laboratory in which patients were sent for analysis due to the suspicion of α -thalassemia or the presence of microcytosis in a previously performed blood count.

We think that another advantage of the model is its strong biological justification. The theoretical basis of the proposed model involves the knowledge that the MCV value and the percentage of Hb A₂ gradually decrease with the increasing number of α -genes mutated, which results in a consequent reduction in the synthesis of α -globin chains [2,36,37]. Due to the unbalanced synthesis of the α -globin chains, there is a deficiency in the production of hemoglobin per cell and, consequently, a decrease in MCV [38]. Simultaneously, the production of Hb A₂ decreases due to the decline of $\alpha_2\delta_2$ tetramers formation [39]. Since the δ -globin chains have a higher positive charge than the β -globin chains, the Hb A₂ production is usually hampered relative to the Hb A₁ [19,39]. In the milder types of α -thalassemia (α^+ heterozygous, α^+ homozygous, and α^0 heterozygous), there is an overlapping percentage of Hb A₂ among these genotypes, which does not always allow a clear distinction between the distinct types of α -thalassemia. The reduced level of Hb A₂ is usually more evident in patients with HbH disease [19,38,40]. Serum ferritin tends to increase with the number of the α -globin genes mutated. Some older patients with HbH disease can have iron overload, due

to increased iron absorption in the intestine, independent of the transfusion received [2,41,42]. The model shows that there is a greater probability that the patients with two or more deletions have the smaller MCV standardized by age, the smaller percentage of Hb A₂, and the greater serum ferritin.

Among the possible limitations of our study, we can mention that the BRL model needs to be evaluated in another cohort, such as an interlaboratory cohort. It would be interesting to validate the model in a larger number of patients to prove its efficacy and reproducibility in distinct regions, populations, and in different age and ethnic groups. Furthermore, our model has not been evaluated for non-deletional α -thalassemia and other less frequent α -globin gene deletions. As a result, we cannot affirm that it would be informative under these conditions.

Our model was built to increase the molecular analysis pre-test probability for two or more deletions. We also tried to apply this model to screen patients with one α -globin gene deletion. However, the test performance did not distinguish the individuals without α -thalassemia from those with a deletion in an α -globin gene. Thus, the proposed BLR model cannot identify patients with only one α -globin gene deleted (silent carrier). It is known that patients suffering from a single gene deletion have minimal or absent laboratory changes, due to a decreased synthesis of α -chains. Hence, they have very similar laboratory tests to individuals who have no alteration in α -globin genes.

Although the patients with α^+ -thalassemia are asymptomatic, they can generate children with HbH disease in countries where there is a high prevalence of α^0 -thalassemia [2,43–45], and in countries where the immigration of the Asian population has increased in recent years. It is important to emphasize that couples with microcytosis and/or anemia who claim to be a parent, as well as pregnant women, should be assessed for mutations related to α -thalassemia, regardless of the outcome of our model. That is because there is a risk of producing a child with HbH disease, and risks of complications in pregnancy due to α -thalassemia, which can be associated with a common iron deficiency in pregnancy (fetal growth restriction, preterm birth, and low birth weight). In such cases, the cause of anemia or microcytosis should be identified correctly and early for appropriate intervention and genetic counseling [46,47].

5. Conclusions

This BLR model will help in the decision-making process to indicate molecular testing for α -thalassemia for patients who have a higher OR of presenting α -thalassemia traits (with two α -genes functioning) or intermedia (HbH disease, with one α -gene functioning). The model is easy to apply, requiring data commonly obtained during the investigation of microcytic and hypochromic anemias. As with other mathematical formulas, it is recommended to validate it in larger cohorts and different populations. The model is useful as a decision tool, especially in regions with limited healthcare resources, and particularly in developing countries.

Supplementary Materials: The following supporting information can be downloaded at: https: //www.mdpi.com/article/10.3390/diagnostics12123008/s1, Table S1: 2 × 2 Table Comparing BRL Model Prediction Group to Multiplex PCR Results (Training Set); Table S2: 2 × 2 Table Comparing BRL Model Prediction Group to Multiplex PCR Results (Validation Set); Table S3: 2 × 2 Table Comparing BRL Model Prediction Group to Multiplex PCR Results (Total Sample) and Excel worksheet S4: Table for Calculating the Standardized MCV (fL) and Table for Calculating the Prediction of Logistic Regression Model.

Author Contributions: Conceptualization, M.K.F., P.F.R.S., E.S.C. and M.G.P.L.; designed research, M.K.F., P.F.R.S., E.S.C. and M.G.P.L.; data curation, P.F.R.S. and M.G.P.L.; methodology, M.G.P.L.; resources, M.K.F.; investigation, P.F.R.S., R.M.P., E.S.C. and R.S.P.S.; formal analysis, M.G.P.L.; validation, M.G.P.L.; writing—original draft preparation P.F.R.S. and M.G.P.L.; writing—review and editing, P.F.R.S. and M.G.P.L.; visualization, P.F.R.S. and M.G.P.L. 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 following the Declaration of Helsinki and approved by the Ethics Committee of the IPPMG hospital (Comitê de Ética em Pesquisa do Instituto de Puericultura e Pediatria Martagão Gesteira—Federal University of Rio de Janeiro). The first acceptance (Acceptance Number 38/11, Approval Date: 14 June 2011) was granted to evaluate the prevalence of α -thalassemia in children with microcytosis in a pediatric hospital in Rio de Janeiro. In this first study, consent terms were collected from all patients. For the construction and validation of the mathematical model, we requested approval from the same Ethics Committee to use the data of other patients evaluated for α -thalassemia in the Clinical Analysis Laboratory of the Faculty of Pharmacy at the Federal University of Rio de Janeiro (UFRJ). Patient consent was waived because the data would be used only for the construction and validation of the mathematical model (Acceptance Number 847.419—Approval Date: 28 October 2014).

Informed Consent Statement: Informed consent was obtained from the subjects involved in the prevalence study (Acceptance Number 38/11, Approval Date: 14 June 2011). For the construction and validation of the mathematical model, the consent of the other patients was not required because only laboratory data would be used, without the need for new blood collection, as well as ensuring that the individuals' data would not be disclosed. (Acceptance Number 847.419—Approval Date: 28 October 2014).

Data Availability Statement: The data used in this study are available as Supplementary Materials (Tables S1–S3 and Excel worksheet S4).

Acknowledgments: The authors would like to acknowledge Isaac Lima da Silva Filho for technical support, assisting us in the establishment of molecular testing for α -thalassemia, and for providing DNA control samples of some of the α -globin gene deletions. We also thank Carlos Eduardo Pedreira for the analytical discussions of the data.

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

References

- 1. Muncie, H.L.; Campbell, J. Alpha and Beta Thalassemia. Am. Fam. Physician 2009, 80, 339–344. [PubMed]
- 2. Harteveld, C.L.; Higgs, D.R. Alpha-Thalassaemia. Orphanet J. Rare Dis. 2010, 5, 13. [CrossRef] [PubMed]
- Chong, S.S.; Boehm, C.D.; Higgs, D.R.; Cutting, G.R. Single-Tube Multiplex-Pcr Screen for Common Deletional Determinants of Alpha-Thalassemia. *Blood* 2000, 95, 360–362. [CrossRef] [PubMed]
- Modell, B.; Darlison, M. Global Epidemiology of Haemoglobin Disorders And Derived Service Indicators. Bull. World Health Organ. 2008, 86, 480–487. [CrossRef]
- 5. Sonati, M.F.; Farah, S.B.; Ramalho, A.S.; Costa, F.F. High Prevalence of Alpha-Thalassemia in A Black Population Of Brazil. *Hemoglobin* **1991**, *15*, 309–311. [CrossRef]
- Adorno, E.V.; Couto, F.D.; Moura Neto, J.P.; Menezes, J.F.; Rêgo, M.; Reis, M.G.; Gonçalves, M.S. Hemoglobinopathies in Newborns from Salvador, Bahia, Northeast Brazil. *Cad. Saude Publica* 2005, 21, 292–298. [CrossRef]
- Souza, A.E.; Cardoso, G.L.; Takanashi, S.Y.; Guerreiro, J.F. Alpha Thalassemia (3.7 Kb Deletion) In A Population from The Brazilian Amazon Region: Santarém, Pará State. *Genet. Mol. Res.* 2009, *8*, 477–481. [CrossRef]
- Couto, F.D.; De Albuquerque, A.B.; Adorno, E.V.; De Moura Neto, J.P.; De Freitas Abbehusen, L.; De Oliveira, J.L.; Dos Reis, M.G.; De Souza Gonçalves, M. Alpha-Thalassemia 2, 3.7 Kb Deletion and Hemoglobin Ac Heterozygosity In Pregnancy: A Molecular And Hematological Analysis. *Clin. Lab. Haematol.* 2003, 25, 29–34. [CrossRef]
- Anselmo, F.C.; Ferreira, N.S.; da Mota, A.J.; Gonçalves, M.S.; Albuquerque, S.R.L.; Fraiji, N.A.; Ferreira, A.C.D.; de Moura Neto, J.P. Deletional Alpha-Thalassemia Alleles In Amazon Blood Donors. *Adv. Hematol.* 2020, 2020, 4170259. [CrossRef]
- Anselmo, F.C.; Soumanou, A.G.; de Aguiar Ferreira, C.; Sobrinha, F.M.V.; Castro, A.C.S.; Brito, R.O.; da Mota, A.J.; de Souza Gonçalves, M.; de Moura Neto, J.P. Detection Of Common Deletional Of A-Thalassemia 3.7 Kb From Metropolitan Region Of Manaus, Amazonas, Brazil. *Mediterr. J. Hematol Infect. Dis.* 2021, 13, E2021001. [CrossRef]
- Silva Filho, I.; Gonçalves, M.S.; Adôrno, E.V.; Campos, D.P.; Fleury, M.K. Triagem De Hemoglobinopatias E Avaliação Da Degeneração Oxidativa Da Hemoglobina Em Trabalhadores Portadores Do Traço Falciforme (Hbas), Expostos A Riscos Ocupacionais. *Rev. Bras. Hematol. Hemoter.* 2005, 27, 183–187. [CrossRef]
- Borges, E.; Wenning, M.R.; Kimura, E.M.; Gervásio, S.A.; Costa, F.F.; Sonati, M.F. High Prevalence Of Alpha-Thalassemia Among Individuals With Microcytosis And Hypochromia Without Anemia. *Braz. J. Med. Biol. Res.* 2001, 34, 759–762. [CrossRef] [PubMed]
- 13. Siqueira, P. Nova Ferramenta De Suporte Para O Rastreio De A-Talassemia Em Pacientes Com Microcitose. Master's Thesis, Federal University of Rio De Janeiro, Rio De Janeiro, Brazil, 2015.
- Clark, B.E.; Thein, S.L. Molecular Diagnosis of Haemoglobin Disorders. *Clin. Lab. Haematol.* 2004, 26, 159–176. [CrossRef] [PubMed]

- 15. Barrett, A.N.; Saminathan, R.; Choolani, M. Thalassaemia Screening and Confirmation of Carriers in Parents. *Best Pract. Res. Clin. Obstet. Gynaecol.* **2017**, *39*, 27–40. [CrossRef]
- 16. Van Vranken, M. Evaluation of Microcytosis. *Am. Fam. Physician* **2010**, *82*, 1117–1122.
- 17. Deloughery, T.G. Microcytic Anemia. N. Engl. J. Med. 2014, 371, 2537. [CrossRef]
- Langlois, S.; Ford, J.C.; Chitayat, D. Carrier Screening For Thalassemia and Hemoglobinopathies In Canada. J. Obstet. Gynaecol. Can. 2008, 30, 950–959. [CrossRef]
- 19. Mosca, A.; Paleari, R.; Ivaldi, G.; Galanello, R.; Giordano, P.C. The Role Of Haemoglobin A(2) Testing In The Diagnosis Of Thalassaemias And Related Haemoglobinopathies. *J. Clin. Pathol.* **2009**, *62*, 13–17. [CrossRef]
- Demir, A.; Yarali, N.; Fisgin, T.; Duru, F.; Kara, A. Most Reliable Indices In Differentiation Between Thalassemia Trait And Iron Deficiency Anemia. *Pediatr. Int.* 2002, 44, 612–616. [CrossRef]
- Nalbantoğlu, B.; Güzel, S.; Büyükyalçın, V.; Donma, M.M.; Güzel, E.Ç.; Nalbantoğlu, A.; Karasu, E.; Özdilek, B. Indices Used In Differentiation Of Thalassemia Trait From Iron Deficiency Anemia In Pediatric Population: Are They Reliable? *Pediatr. Hematol.* Oncol. 2012, 29, 472–478. [CrossRef]
- 22. Mentzer, W.C. Differentiation of Iron Deficiency from Thalassaemia Trait. Lancet 1973, 1, 882. [CrossRef] [PubMed]
- 23. Srivastava, P.C.; Bevington, J.M. Iron Deficiency and or Thalassaemia Trait. Lancet 1973, 1, 832. [CrossRef] [PubMed]
- 24. Matos, J.F.; Dusse, L.M.; Borges, K.B.; de Castro, R.L.; Coura-Vital, W.; Carvalho, M. A New Index To Discriminate Between Iron Deficiency Anemia And Thalassemia Trait. *Rev. Bras. Hematol. Hemoter.* **2016**, *38*, 214–219. [CrossRef] [PubMed]
- Janel, A.; Roszyk, L.; Rapatel, C.; Mareynat, G.; Berger, M.G.; Serre-Sapin, A.F. Proposal Of A Score Combining Red Blood Cell Indices For Early Differentiation Of Beta-Thalassemia Minor From Iron Deficiency Anemia. *Hematology* 2011, 16, 123–127. [CrossRef]
- 26. 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]
- Urrechaga, E.; Hoffmann, J.J.M.L. Critical Appraisal of Discriminant Formulas for Distinguishing Thalassemia From Iron Deficiency In Patients With Microcytic Anemia. *Clin. Chem. Lab. Med.* 2017, 55, 1582–1591. [CrossRef]
- Huang, T.C.; Wu, Y.Y.; Chen, Y.G.; Lai, S.W.; Wu, S.C.; Ye, R.H.; Lu, C.S.; Chen, J.H. Discrimination Index Of Microcytic. Anemia In Young Soldiers: A Single Institutional Analysis. *PLoS ONE* 2015, *10*, E0114061. [CrossRef]
- Fu, Y.K.; Liu, H.M.; Lee, L.H.; Chen, Y.J.; Chien, S.H.; Lin, J.S.; Chen, W.C.; Cheng, M.H.; Lin, P.H.; Lai, J.Y.; et al. The Tvgh-Nycu Thal-Classifier: Development of A Machine-Learning Classifier For Differentiating Thalassemia And Non-Thalassemia Patients. *Diagnostics* 2021, 11, 1725. [CrossRef]
- 30. Huber, A.R.; Ottiger, C.; Risch, L.; Regenass, S.; Hergersberg, M.; Herklotz, R. Syndromes Thalassémiques: Clinique Et Diagnostic. *Swiss. Med. Forum* **2004**, *4*, 947–952.
- 31. Zhou, Y.; Zhang, J.; Wang, C.; Zhou, L.; Zhou, L.; Ou, D.; Peng, D. Application Of Hba. *Int. J. Lab. Hematol.* **2020**, *42*, 526–532. [CrossRef] [PubMed]
- 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. [CrossRef] [PubMed]
- de Medeiros Alcoforado, G.H.; Bezerra, C.M.; Araújo Moura Lemos, T.M.; de Oliveira, D.M.; Kimura, E.M.; Ferreira Costa, F.; de Fátima Sonati, M.; de Medeiros, T.M. Prevalence Of A-Thalassemia 3.7 Kb Deletion In The Adult Population Of Rio Grande Do Norte, Brazil. *Genet. Mol. Biol.* 2012, 35, 594–598. [CrossRef] [PubMed]
- Wagner, S.C.; de Castro, S.M.; Gonzalez, T.P.; Santin, A.P.; Filippon, L.; Zaleski, C.F.; Azevedo, L.A.; Amorin, B.; Callegari-Jacques, S.M.; Hutz, M.H. Prevalence Of Common A-Thalassemia Determinants In South Brazil: Importance For The Diagnosis Of Microcytic Anemia. *Genet. Mol. Biol.* 2010, 33, 641–645. [CrossRef] [PubMed]
- Cardoso, G.L.; Takanashi, S.Y.; Guerreiro, J.F. Inherited Hemoglobin Disorders In An Afro-Amazonian Community: Saracura. Genet. Mol. Biol. 2012, 35, 553–556. [CrossRef] [PubMed]
- Fucharoen, S.; Viprakasit, V. Hb H Disease: Clinical Course and Disease Modifiers. *Hematol. Am. Soc. Hematol. Educ Program.* 2009, 1, 26–34. [CrossRef]
- Giambona, A.; Passarello, C.; Renda, D.; Maggio, A. The Significance Of The Hemoglobin A(2) Value In Screening For Hemoglobinopathies. *Clin. Biochem.* 2009, 42, 1786–1796. [CrossRef] [PubMed]
- Higgs, D.R. The Pathophysiology and Clinical Features Of Alpha-Thalassemia. In Disorders of Hemoglobin: Genetics, Pathophysiology, Clinical Management, 2nd ed.; Steinberg, M.H., Forget, B.G., Eds.; Cambridge University Press: New York, NY, USA, 2009; pp. 266–295.
- 39. Steinberg, M.H.; Nagel, R.L. Hemoglobin of the Embryo, Fetus, and Adult. In *Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management*, 2nd ed.; Steinberg, M., Forget, B., Eds.; Cambridge University Press: New York, NY, USA, 2009; pp. 119–138.
- Van Delft, P.; Lenters, E.; Bakker-Verweij, M.; de Korte, M.; Baylan, U.; Harteveld, C.L.; Giordano, P.C. Evaluating Five Dedicated Automatic Devices For Haemoglobinopathy Diagnostics In Multi-Ethnic Populations. *Int. J. Lab. Hematol.* 2009, 31, 484–495. [CrossRef]
- Chen, F.E.; Ooi, C.; Ha, S.Y.; Cheung, B.M.; Todd, D.; Liang, R.; Chan, T.K.; Chan, V. Genetic And Clinical Features Of Hemoglobin H Disease In Chinese Patients. *N. Engl. J. Med.* 2000, 343, 544–550. [CrossRef]
- Chui, D.H.; Fucharoen, S.; Chan, V. Hemoglobin H Disease: Not Necessarily A Benign Disorder. *Blood* 2003, 101, 791–800. [CrossRef]

- 43. Singer, S.T. Variable Clinical Phenotypes of Alpha-Thalassemia Syndromes. Sci. World J. 2009, 9, 615–625. [CrossRef]
- 44. Higgs, D.R.; Weatherall, D.J. The Alpha Thalassaemias. Cell Mol. Life Sci. 2009, 66, 1154–1162. [CrossRef] [PubMed]
- 45. Vichinsky, E. Complexity of Alpha Thalassemia: Growing Health Problem with New Approaches To Screening, Diagnosis, And Therapy. *Ann. N. Y. Acad. Sci.* 2010, 1202, 180–187. [CrossRef] [PubMed]
- Xiao, H.; Wang, Y.; Ye, Y.; Yang, C.; Wu, X.; Wu, X.; Zhang, X.; Li, T.; Xiao, J.; Zhuang, L.; et al. Differential Diagnosis Of Thalassemia And Iron Deficiency Anemia In Pregnant Women Using New Formulas From Multidimensional Analysis Of Red Blood Cells. Ann. Transl. Med. 2021, 9, 141. [CrossRef] [PubMed]
- 47. Tongsong, T.; Srisupundit, K.; Luewan, S. Outcomes of Pregnancies Affected By Hemoglobin H Disease. *Int. J. Gynaecol. Obstet.* **2009**, *104*, 206–208. [CrossRef] [PubMed]