Electronic Journal of General Medicine
2023, 20(4), em501
e-ISSN: 2516-3507
https://www.ejgm.co.uk/

MODESTUM

Original Article

OPEN ACCESS

Prevalence of thalassemia in the Vietnamese population and building a clinical decision support system for prenatal screening for thalassemia

Danh Cuong Tran [1,2], Anh Linh Dang [1], Thi Ngoc Lan Hoang [3], Chi Thanh Nguyen [4], Thi Minh Phuong Le [5], Thi Ngoc Mai Dinh [6], Van Anh Tran [7], Thi Kim Phuong Doan [3], Thi Trang Nguyen [3,8]*

[1] Center for Prenatal Diagnosis, National Hospital of Obstetrics and Gynecology, Hanoi, VIETNAM
[2] Department of Obstetrics and Gynecology, Hanoi Medical University, Hanoi, VIETNAM
[3] Department of Biomedical and genetics, Hanoi Medical University, Hanoi, VIETNAM
[4] Department of Specialized Software, Academy of Military Science and Technology, Hanoi, VIETNAM
[5] Department of Basic Sciences in Medicine and Pharmacy, University of Medicine and Pharmacy-Vietnam National University, Hanoi, VIETNAM
[6] Department of Pediatrics, Hanoi Medical University, Hanoi, VIETNAM
[7] Department of Pediatrics, Hanoi Medical University Hospital, Hanoi, VIETNAM
[8] Clinical Genetics Center, Hanoi Medical University Hospital, Hanoi, VIETNAM
***Corresponding Author:** trangnguyen@hmu.edu.vn

Citation: Tran DC, Dang AL, Hoang TNL, Nguyen CT, Le TMP, Dinh TNM, Tran VA, Doan TKP, Nguyen TT. Prevalence of thalassemia in the Vietnamese population and building a clinical decision support system for prenatal screening for thalassemia. Electron J Gen Med. 2023;20(4):em501. https://doi.org/10.29333/ejgm/13206

ARTICLE INFO

Received: 02 Dec. 2022

Accepted: 07 Mar. 2023

ABSTRACT

The prevalence of thalassemia among the Vietnamese population was studied, and clinical decision support systems (CDSSs) for prenatal screening of thalassemia were created. A cross-sectional study was conducted on pregnant women and their husbands visiting from October 2020 to December 2021. A total of 10,112 medical records of first-time pregnant women and their husbands were collected. CDSS including two different types of systems for prenatal screening for thalassemia (expert system [ES] and four artificial intelligence [AI]-based CDSS) was built. 1,992 cases were used to train and test machine learning (ML) models while 1,555 cases were used for specialized ES evaluation. There were 10 key variables for AI-based CDSS for ML. The four most important features in thalassemia screening were identified. Accuracy of ES and AI-based CDSS was compared. The rate of patients with alpha thalassemia is 10.73% (1,085 patients), the rate of patients with beta-thalassemia is 2.24% (227 patients), and 0.29% (29 patients) of patients carry both alpha-thalassemia and beta-thalassemia gene mutations. ES showed an accuracy of 98.45%. Among AI-based CDSS developed, multilayer perceptron model was the most stable regardless of the training database (accuracy of 98.50% using all features and 97.00% using only the four most important features). AI-based CDSS showed satisfactory results. Further development of such systems is promising with a view to their introduction into clinical practice.

Keywords: thalassemia, Vietnamese population, clinical decision support system, expert system, AI-based system

INTRODUCTION

Thalassemia–inherited autosomal recessive disease, which is characterized by impaired synthesis of hemoglobin protein chains [1]. A normal mature hemoglobin molecule (HbA) consists of two pairs of alpha and beta chains [2]. Thalassemia is caused by a gene mutation of the gene responsible for globin chains synthesis, based on which alpha-thalassemia and beta-thalassemia, are distinguished [3]. Such a gene can be inherited from one parent or two. The child's body produces fewer or no hemoglobin chains. The production of the other chains that make up globin does not end. As a result, unstable protein components are produced that destroy the blood cells [4]. Thalassemia is thus the result of reduced synthesis of at least one globin polypeptide chain, resulting in abnormal erythrocytes, anemia and often in iron overload [5].

The severity of the disease depends on the number of mutated alleles. In humans, the alpha chain of hemoglobin is encoded by two pairs of genes, while the beta chain has only one pair. Patients with one alpha+allele are clinically normal and are called asymptomatic carriers. Heterozygotes with defects in two of the four genes (small alpha thalassemia) tend to develop microcytic anemia of mild to moderate severity, but with a subclinical course. Defects in three of the four genes significantly impair alpha-chain synthesis in which case hemolytic anemia and splenomegaly are common. A defect in all four is a fatal condition that causes intrauterine fetal death [6]. Minor beta-thalassemia occurs in asymptomatic heterozygotes with a mild to moderate clinical picture of microcytic anemia. Intermediate beta-thalassemia presents a variable clinical picture due to the inheritance of two beta-thalassemia alleles. Large beta-thalassemia (Cooley's anemia) occurs in homozygous patients or complex heterozygotes as a

result of a severe beta-globin defect. These patients develop severe anemia and bone marrow hyperactivity [7, 8]. In addition, it is rare to find simultaneous disorders in both alpha- and beta-chain at once, but in this case the disease may be milder because there is little imbalance between the two types of chains [3]. Thus, severe forms of the disease seriously affect physical development, causing patients to need continuous blood transfusions for life, causing many complications in the liver, heart, endocrine glands, and bones. The disease is not only life-threatening, affecting the quality of life of the patient, but also expensive treatment costs bring a burden to the family and the whole society [9].

Worldwide, an estimated 7.00% of the population carries the thalassemia gene, and each year between 300,000 and 500,000 babies are born with severe homozygosity for the disease [10, 11]. In Vietnam, there are more than five million people who carry the gene and suffer from thalassemia, every year there are more 100,000 children carrying the disease gene and 1,700 children with severe disease due to mutations in both genes [12]. Thalassemia is distributed in all provinces and ethnic groups throughout the country, especially ethnic minorities in mountainous provinces [12].

Thalassemia is a preventable disease by screening pregnant women and their husbands at risk of carrying the disease gene to prevent having children with the disease [13, 14].

Nowadays, modern technology has been researched and applied in the field of medicine to support doctors in patient care as well as practice specialize [15]. A clinical decision support system (CDSS) is “any electronic or non-electronic system designed to aid directly in clinical decision making in which characteristics of individual patients are used to generate patient-specific assessments or recommendations that are then presented to clinicians for consideration” [16]. CDSSs are classified as expert knowledge-based systems and artificial intelligence (AI) [17]. CDSSs have many advantages, such as reducing the rate of misdiagnosis, improving efficiency and patient care, and reducing the risk of medication errors [18]. For thalassemia, there have been studies around the world applying AI in screening carriers with high efficiency. In 2002, Amendolia and colleagues [19] studied and built a real-time classification system based on artificial neural networks (ANNs) to distinguish thalassemia gene carriers and normal people with an accuracy of 94.00%, a sensitivity of 92.00% and a specificity 95.00%. In 2013, the study [20] compared the performance of radial basis function (RBF) network, probabilistic neural network (PNN), and k-nearest neural network (KNN) algorithms in thalassemia screening with 304 data samples. The results show that RBF algorithm had a sensitivity of 93.00%, a specificity of 91.00%, similar to the results of KNN of 80% and 91%, of PNN of 89.00% and 73.00% [20].

The screening for thalassemia in Vietnam is still mainly based on the two indexes (mean corpuscular volume [MCV] and mean corpuscular hemoglobin [MCH]) and performed manually, no CDSS has been built yet. This causes many difficulties in disease prevention at primary health care facilities and in ethnic minority areas because of the limited understanding of thalassemia not only by the people but also by the grassroots medical staff. The requirement is to build AI software and an expert knowledge-based system for thalassemia screening that can be applied to even primary health care facilities.

Thus, the development and introduction into clinical practice of modern CDSSs for screening thalassemia in Vietnam is an urgent task. These systems can provide significant assistance to doctors in making optimal decisions, even in primary health care institutions. Special attention should be paid to CDSSs based on machine learning (ML) algorithms, which remain poorly studied.

Therefore, our research was carried out with the following objectives:

1. Investigating the prevalence of thalassemia in the Vietnamese population
2. Building a CDSS for prenatal screening for thalassemia.

METHODS

Study Design

This is a cross-sectional study conducted from October 2020 to December 2021. Data were collected using convenient sampling method. We collected data of pregnant women and their husband when they come to National Hospital of Obstetrics and Gynecology for annual screening of birth defects through medical records. Data were collected from the medical records of patients who came to the hospital before the study.

Study Subjects

A total of 10,112 medical records of first-time pregnant women and their husbands were collected, of which 1,992 cases were used to train and test ML models while 1,555 cases were used for specialized knowledge system evaluation. All patients underwent routine screening of thalassemia: peripheral blood smear, complete blood count (CBC), hemoglobin quantification by high performance liquid chromatography (HPLC), and capillary electrophoresis (CE), iron status tests. Detection of hemoglobin gene mutations by polymerase chain reaction (PCR) was performed for 1,364 patients and 658 newborns to assess the prevalence of different forms of thalassemia in the Vietnamese population. Multiplex ligation-dependent probe amplification (MLPA) technique has been used for molecular detection of alpha-thalassemia. Reverse dot-blot PCR technique has been used for molecular detection of beta-thalassemia.

Data Analysis

Two types of CDSS models for thalassemia pre-screening have been created: four AI-based CDSS for ML and expert system (ES). The basic difference of these two types of systems is that one is based on the knowledge base gathered from the knowledge of experts, and the other is based on computer mining knowledge from medical data. The specialized ES was built based on the guideline for prenatal screening for thalassemia of the Vietnamese Ministry of Health.

The following independent variables were used in ES CDSS: four most important indicators from CBC result (according to MID and MDA algorithms) including HGB, MCH, MCV, red blood cell distribution width (RDW), serum ferritin concentration from iron status tests result, HbA2, and HbF levels from HPLC result. In addition, one should consider the history of hydrops fetalis and having children or family members diagnosed with thalassemia.

Table 1. Frequency of different types of thalassemia in patients

Types of thalassemia		Pregnant women		Husband		Total	%
		n	%	n	%		
Patients who performed genetic testing	Alpha thalassemia	566	6.40	519	40.83	1,085	10.73
	Beta thalassemia	117	1.32	110	8.65	227	2.24
	Co-inheritance of alpha- & beta-thalassemia	7	0.08	22	1.73	29	0.29
	Others	14	0.16	9	0.71	23	0.23
Patients who did not perform genetic testing		8,137	92.04	611	48.07	8,748	86.51
Total number of cases that performed CBC &d iron status test		8,841	100.00	1,271	100.00	10,112	100.00

Table 2. Prevalence of thalassemia types in fetuses who underwent genetic testing

Types of thalassemia	n	%
Alpha thalassemia	407	61.85
Beta thalassemia	86	13.07
Co-inheritance of alpha- & beta-thalassemia	19	2.89
Others	4	0.61
Normal	142	21.58
Total	658	100.00

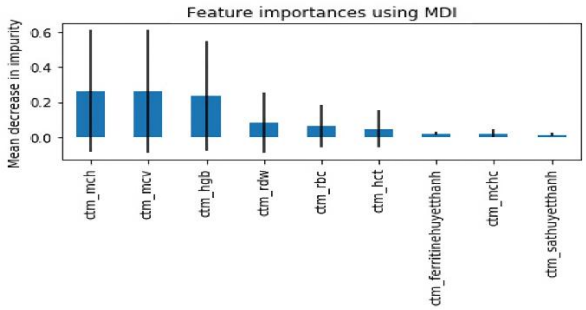

Figure 1. Feature importance using MDI algorithm-1 (Source: Authors' own elaboration)

There were 10 key variables for AI-based CDSSs for ML: the dependent variable was whether a patient had a thalassemia gene; nine independent variables including seven from the CBC result were hematocrit (HCT), MCH, mean corpuscular hemoglobin concentration (MCHC), MCV, hemoglobin (HGB), RDW, red blood cell (RBC) and the other two are iron status: serum iron and serum ferritin. In addition, AI-based CDSSs for ML were evaluated using the stated above four most important indicators (according to MID and MDA algorithms).

Data from 1,992 pregnant women and their husbands were used to train and test four ML models, which were KNN, support vector machine (SVM), random forest (RF), and multilayer perceptron (MLP). The purpose of these models is screening for thalassemia gene in pregnant women, husbands, and both pregnant women and husbands. Thus, we used data from all participants and divide it into two subsets, one with data from pregnant women only and the other from the husbands. After analyzing the dependent variable, which is whether the participant had thalassemia gene, we realized that there were more participants without thalassemia gene than those who did, which caused an unbalance in the dataset and result in the inaccuracy of all models. To solve this problem, we performed synthetic minority over-sampling technique (SMOTE). This method was introduced in 2002 [21], the idea is based on the k-nearest neighbors algorithm. We get one sample randomly from the minority layer a and one of its k-nearest neighbors in the feature space, then we choose randomly a k-nearest neighbors b and draw a line between these samples in the feature space. New samples are created on this line as the combination of a and b. These new samples

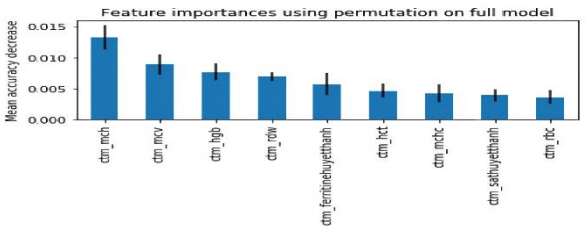

Figure 2. Feature importance using MDA algorithm-2 (Source: Authors' own elaboration)

helped balance the datasets, which mean the number of participants with thalassemia gene is now equal to those who did not. Datasets were then standardized using z-score method, in particular we used StandardScaler command from scikit-learn library. Hyperparameters were found using grid search.

All four models were tested using 10-fold cross validation using these datasets and evaluated by four indices: accuracy, precision, recall, and F1-score to find the best one.

RESULTS

Prevalence of Types of Thalassemia

10,112 pregnant women and their husbands performed the CBC, HPLC, and iron status during the study period. Based on the results of the CBC, HPLC, and iron status, 1,364 pregnant women and their husbands were prescribed a genetic test due to the suspected presence of thalassemia genes. The genetic test resulted in 1,085 (10.73% of the total number of pregnant women and their husbands) alpha-thalassemia cases and 227 (2.24%) beta-thalassemia cases. A small ratio of 0.29% inherited both α- and β-thalassemia genes and 0.23% were others, including HbE disease, alpha thalassemia/HbE, beta thalassemia/HbE and hemoglobin constant spring disease (**Table 1**). Among 658 fetuses of parents with identified thalassemia genes who performed the genetic testing, the frequency of α-thalassemia (61.85%) was also higher than β-thalassemia (13.07%) and others (including HbE disease, α-thalassemia/HbE, and β-thalassemia/HbE) (0.61%) (**Table 2**).

Determining Features of the CBC, HPLC, and Iron Status Test Needed to Screen for Thalassemia

After using MID and MDA algorithms, we found that four indices included HGB, MCV, MCH, and RDW were the most important ones in the screening of thalassemia from the database containing 10,112 cases having CBC, HPLC, and iron status results (**Figure 1** and **Figure 2**).

Table 3. Comparing results of models using four most important features with models using all features in general CBC & iron status database (pregnant women & their husbands)

Model	Accuracy (%)		Precision (%)		Recall (%)		F1-score (%)	
	Original data	SMOTE	Original data	SMOTE	Original data	SMOTE	Original data	SMOTE
All CBC & iron status indices								
KNN	98.00	95.50	93.87	90.12	100.00	99.12	98.66	95.67
SVM	95.00	94.50	94.50	90.17	97.15	97.15	95.34	94.04
RF	97.00	95.50	93.24	85.24	98.86	96.45	96.08	92.45
MLP	98.50	95.50	94.67	89.80	100.00	99.23	97.65	94.73
Only four most important indices (HGB, MCV, MCH, & RDW)								
KNN	97.00	92.50	93.75	85.90	98.68	98.68	98.15	90.91
SVM	94.50	93.50	90.12	87.18	96.05	96.05	93.59	93.00
RF	96.00	92.50	91.46	84.27	98.68	96.15	94.94	90.68
MLP	97.00	94.50	92.68	88.24	99.23	98.68	96.20	93.17

Note. Precision & recall indices were calculated for thalassemia carrier

Table 4. Comparing results of models using four most important features with models using all features in CBC & iron status database of pregnant women

Model	Accuracy (%)		Precision (%)		Recall (%)		F1-score (%)	
	Original data	SMOTE	Original data	SMOTE	Original data	SMOTE	Original data	SMOTE
All CBC & iron status indices								
KNN	**96.93**	**97.53**	**93.22**	**93.33**	**98.21**	**100.00**	95.65	**96.55**
SVM	95.68	92.60	91.52	83.33	100.00	98.21	93.91	90.16
RF	**96.91**	91.36	91.80	80.00	98.21	**100.00**	**95.72**	88.90
MLP	**96.91**	**97.53**	**93.25**	**93.33**	**100.00**	**100.00**	95.65	**96.55**
Only four most important indices (HGB, MCV, MCH, & RDW)								
KNN	**97.53**	**97.87**	**93.38**	**93.41**	98.45	**100.00**	**96.75**	**96.74**
SVM	96.91	90.74	91.80	79.71	96.43	97.83	95.73	88.00
RF	96.22	91.12	90.74	79.45	97.56	99.54	95.65	87.60
MLP	**97.81**	**97.65**	93.21	**93.28**	**99.84**	**99.73**	**95.72**	96.43

Note. Precision & recall indices were calculated for thalassemia carrier

Selection and Evaluation of the Effectiveness of AI Models in Screening for Thalassemia

Experimental result with general CBC and iron status database

According to **Table 3**, when using the four most important indicators in thalassemia screening instead of using them all in CBC and iron status database, accuracy and precision indices decreased but not much.

With several models, the result was even more improved. For example, with SVM model, accuracy and precision indices were increased (95.00% and 91.25% versus 94.50% and 90.12%).

KNN model using only four important features and training with the general CBC and iron status database (pregnant women and their husbands) had an accuracy of 97.00%, a precision of 93.75%, a recall of 98.68%, and an F1-score of 98.15% for original data and an accuracy of 92.50%, a precision of 85.90%, a recall of 98.68%, and an F1-score of 90.91% for SMOTE; MLP model had an accuracy of 97.00%, a precision of 92.68%, a recall of 99.23%, and an F1-score of 96.20% for original data and an accuracy of 94.50%, a precision of 88.24%, a recall of 98.68%, and an F1-score of 93.17% for SMOTE. SVM and RF models showed similar, but slightly lower results for some indices (**Table 3**).

Experimental result with CBC and iron status database of pregnant women

With the data filtered for pregnant women, the training and testing are similar to the full one that include pregnant women and their husband. The results were shown in **Table 4**.

The results showed that among four training models, KNN and MLP models were the best with 96.93% and 96.91% accuracy, 93.22% and 93.25% precision, 98.21% and 100.00% recall when using all CBC and iron status features. Wherein, with only four most important features, the accuracy of KNN and MLP models was impressively increased from 96.93% to 97.53% and 96.91% to 97.81%. The accuracy of SVM model also increased from 95.68% to 96.91% (**Table 4**). One remarkable issue with this database was that after balancing the dataset by SMOTE technique, the result was improved for some indexes.

Experimental result with CBC and iron status database of husband

Similar to CBC and iron status data of pregnant women, the models trained were not disturbed by sex. With full CBC and iron status, RF model showed the highest accuracy at 97.55%, precision index at 95.46%, and recall index at 100.00% (**Table 5**). With using only four important features, results were not changed, except for MLP model whose accuracy was increased significantly to 97.44%. One remarkable issue with this database was that after balancing dataset by SMOTE technique, results were decreased for some indexes (**Table 5**).

Evaluation of Effectiveness of Expert System in Screening for Thalassemia and Comparation with AI Models

To build an ES in screening for thalassemia, the rules in **Table 6** were applied. The effectiveness of ES in the same CBC, HPLC, and iron status database that was used to test AI models was evaluated, including 396 pregnant women and husbands who met the inclusion criteria (presence of thalassemia gene by PCR). However, the result showed that 323 cases were determined the risk of having thalassemia by the software and 73 cases that were not. The reason for 73 cases was that

Table 5. Comparing results of models using four most important features with models using all features in CBC & iron status database of husbands

Model	Accuracy (%)		Precision (%)		Recall (%)		F1-score (%)	
	Original data	SMOTE	Original data	SMOTE	Original data	SMOTE	Original data	SMOTE
All CBC & iron status indices								
KNN	92.31	92.31	95.00	94.72	90.48	90.23	92.68	92.23
SVM	94.87	94.62	95.24	94.94	94.52	94.31	95.36	95.18
RF	**97.55**	92.31	**95.46**	95.46	**100.00**	90.48	92.67	92.76
MLP	97.91	97.44	94.46	95.46	96.55	91.92	97.68	97.68
Only four most important indices (HGB, MCV, MCH, & RDW)								
KNN	91.44	90.84	94.86	94.54	90.12	89.87	91.64	91.52
SVM	94.45	94.16	94.83	94.52	93.96	93.72	95.04	94.98
RF	96.44	84.87	95.32	94.22	99.87	91.24	97.67	87.24
MLP	97.44	94.18	96.42	95.44	100.00	95.95	97.62	95.18

Note. Precision & recall indices were calculated for thalassemia carrier

Table 6. Table on top of a page

Evaluation	Indices	Values	Conclusion	Prediction results
Anemia	HGB	Pregnant women≥110 & Husband≥130	Non-anemia	Insufficiency of data
		Pregnant women<110 & Husband<130	Anemia	
Size of red blood cell	MCV	<85	Microcytic erythrocyte	
		85-100	Normal-size RBC	
		>100	Macrocytic erythrocyte	
Amount of hemoglobin	MCH	<28	Hypochromic erythrocytes	
		28-32	Normochromic erythrocytes	
		>32	Hyperchromic erythrocytes	
Iron deficiency	Ferritin	<13	Iron deficiency	
		≥13	Non-iron deficiency	
Risk of thalassemia carrier if patient performed CBC test & ferritin blood test	HGB, MCV, MCH, RDW, & Ferritin		-Possible types of thalassemia -Recommendation for performing genetic testing searching for thalassemia mutations	High risk of thalassemia
Risk of thalassemia carrier if patient performed CBC test, ferritin blood test, & hemoglobin variant analysis test	HGB, MCV, MCH, RDW, Ferritin, HbA2, & HbF			
History of hydrops fetalis or having children or family members diagnosed with thalassemia			Yes or no	Low risk or high risk of thalassemia

Table 7. Evaluation of effectiveness of expert system

	HR	LR	U (n=73)	Indicators	
Patients with thalassemia	113	0	36	True positive	113
				True negative	205
Patients without thalassemia	5	205	37	False positive	5
				False negative	0

Note. HR: High risk; LR: Low risk; & U: Unknown

Table 8. Table on top of a column (font size: 9)

	ES	AI model			
		KNN	SVM	RF	MLP
Accuracy (%)	98.45	97.00	94.50	96.00	97.00
Sensitivity (%)	100.00	-	-	-	-
Specificity (%)	97.62	-	-	-	-

Note. ES: Expert system & AI: Artificial intelligence

patients had performed only CBC test, neither serum ferritin nor hemoglobin variant analysis and no related history, which caused a lack of indicators to predict. 323 cases left were evaluated and calculated shown in **Table 7** and **Table 8**. There were five incorrect cases in risk assessment (false positive) and no false negative cases, which meant no thalassemia patients were left out. Comparing ES with AI model, the accuracy of ES was 98.45%, while the KNN and MLP models was 97.00%, RF model–96.00%, and SVM model–94.50% when using only the four most important features (**Table 8**).

DISCUSSION

The present study analyzed the prevalence of different thalassemia types among pregnant women and their husbands who came to National Hospital of Obstetrics and Gynecology. The prevalence of different thalassemia types in fetuses of parents with established, according to PCR tests, thalassemia genes was also evaluated.

The prevalence of alpha-thalassemia in pregnant women, husbands, and fetuses was higher than in beta-thalassemia carriers (**Table 1** and **Table 2**).

This may have been due to the fact that women whose babies had hydrop details in the previous pregnancy or the fetuses that were alpha thalassemia would show clinical manifestations so they would go for a check-up and perform prenatal screening tests to do the treatment or prevent thalassemia for the next pregnancy. This was the reason why the prevalence of alpha-thalassemia carriers of pregnant women who were screened for thalassemia was also higher than beta-thalassemia at the prenatal diagnostic center in Central Obstetrics Hospital in the [22].

In our study, 8,841 pregnant women were screened by CBC, HPLC, and iron status test, the prevalence of alpha-thalassemia carriers was 6.40%, the prevalence of beta-thalassemia carriers

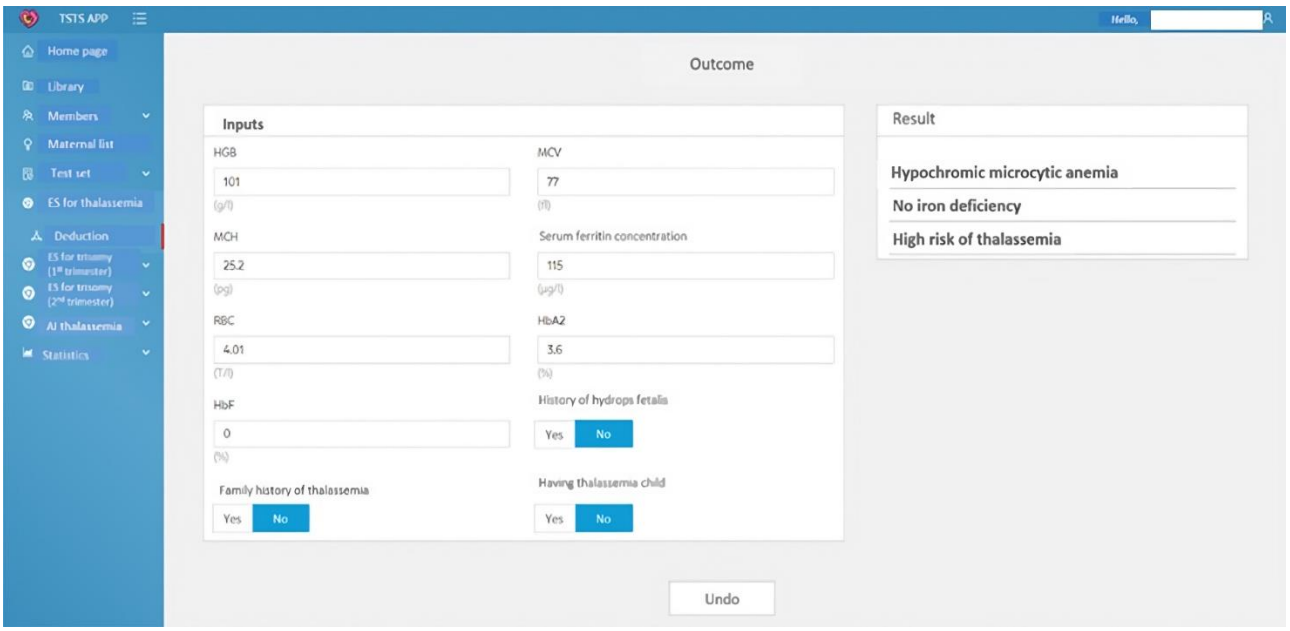

Figure 3. Screenshots of web-based ES after submission of laboratory data (Source: Authors' own elaboration)

was 1.32%, the prevalence of co-inheritance of alpha and beta-thalassemia was 0.08% (**Table 1**), which were higher than those in [23]. In this study, the percentage of pregnant women carrying the alpha-thalassemia carriers, beta-thalassemia carriers and co-inheritance of alpha, and beta-thalassemia were 1.11%, 0.52% and 0.07%, respectively. However, in both studies, the percentage of pregnant women and pregnant women with alpha-thalassemia carriers was the highest (6.4%, 61.85% and 1.11%, 71.30%) [23]. However, our results were lower compared to the research of [24] (8.10% and 3.40%) and [25] (alpha gene carrier thalassemia is 9.80%). The reason for these differences can be explained by the difference in size, region, and country.

Nowadays, one of the most important components in family planning and pregnancy is the detection and prevention of hereditary pathologies in future offspring. In this regard, genetic tests and consultations are conducted for future parents, based on which final decisions are made. CDSS, which is currently being widely implemented in various fields of medicine, can undoubtedly provide significant assistance to physicians in the diagnosis and prognosis of hereditary pathologies [26].

Thalassemia is a complicated disease without any cures. Currently, the best strategies for thalassemia management are screening for thalassemia based on family history, cell blood count, serum iron and Ferritin. However, this is a challenge for primary health care facilities and ethnic minority areas due to the limited understanding of thalassemia not only by the population but also by the grassroots medical staffs. To solve this problem, the authors conducted a CDSS aiming to support physicians in screening and diagnosing thalassemia. CDSS is a system comprising two different systems: ES and AI-based CDSS.

ES is a knowledge-based CDSS encoding the experts' knowledge into an automated system [27]. The operating principle of ES is the simulation of the procedure of diagnosis and screening for thalassemia done by medical physicians. It aims to support doctors deal with complicated cases, especially at the commune health centers where they are facing a shortage of qualified health workers. Like other ESs, it is composed of three main parts: The knowledge base, the inference engine, and the EHR front-end interface. The knowledge base includes a set of rules in the form of if-then rules built based on RBC indices, ferritin, and hemoglobin electrophoresis results by using the guideline for prenatal screening for thalassemia of WHO and the Vietnam Ministry of Health (**Table 8**).

The guideline comprises recommendations proposed by experts and used by medical physicians to screen and diagnose thalassemia in clinical practice, thus it is reliable and suitable for the racial characteristics of Vietnamese people. The system also includes a knowledge update interface, with which the experts can enrich the knowledge base by updating their clinical experience and new medical knowledge directly. This is an important part to ensure the accuracy and the update of ES because the knowledge of thalassemia can be updated and changed over time. The inference engine is an essential part of ES, which applies the if-then rules of the knowledge base to the patient's clinical data to create an inference. Development of the inference engine is an important step when building ES, in the research, to increase the accuracy of the system, the three-layer model was applied.

The EHR frontend interface of the model has two sections: the input section and the output section (**Figure 3**). The input section includes nine boxes to enter patients' clinical information including cell blood count test results, plasma iron, and Ferritin. The output section, which includes three boxes shows the conclusion drawn by the inference engine. The conclusions drawn by ES predict the risk of thalassemia whether it is high risk or low risk, therefrom the physicians can identify if this case requires diagnostic genetic testing or not.

The advantage of ES is that it is possible to explain how the system makes the recommendation, and due to this, it has high reliability and accuracy. Another advantage of our ES is that it is deployed on the Internet, thus it is possible for users to access ES at any computer and at any time with an internet connection. The evaluation of the effectiveness of the developed ES in thalassemia screening in the present study showed a high accuracy of 98.45%.

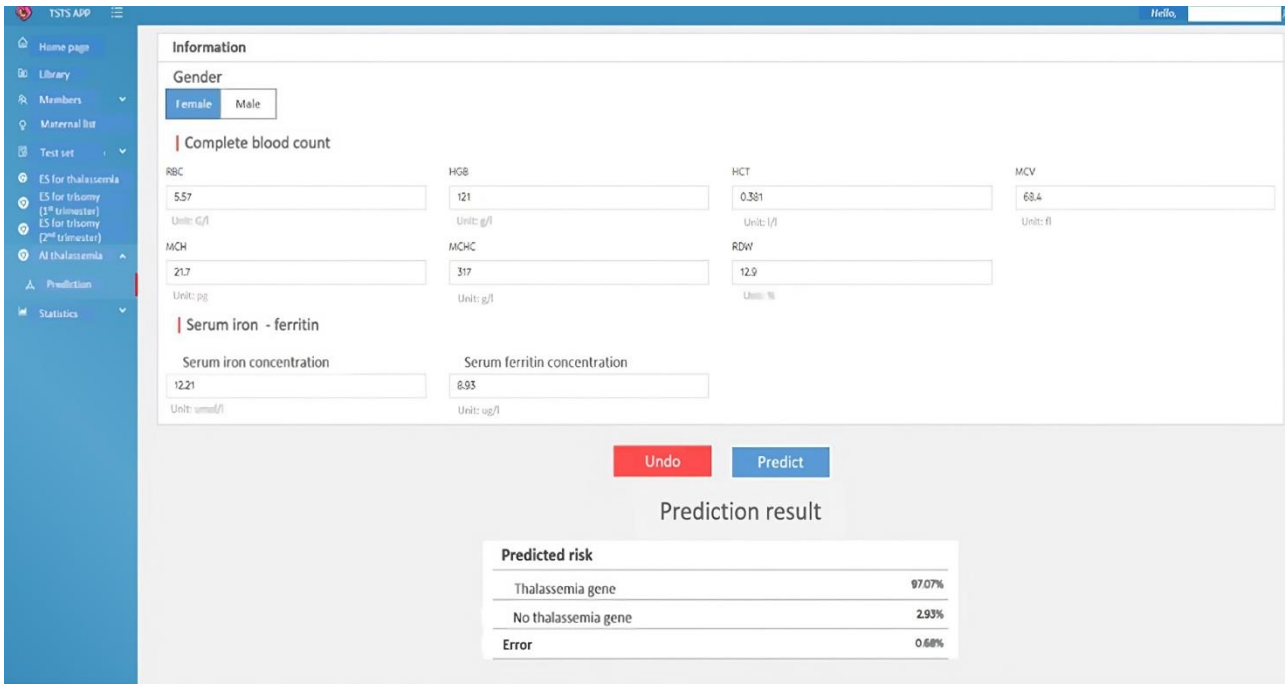

Figure 4. Screenshots of web-based AI model after submission of laboratory data (Source: Authors' own elaboration)

Unlike ES, AI-based CDSS does not use the knowledge base, instead, it uses a form of AI called ML, which allows computers to learn from past experiences and/or find patterns in clinical data to make decisions, thus it does not require writing rules. ML describes the use of computer algorithms to determine patterns in very large datasets. Over the past years, ML has been applied in a wide range of medical fields and demonstrated impressive results, especially in clinical decision support, patient monitoring, and management [28]. There have been many ML techniques applied in building CDSS in which KNN, SVM, RF, and MLP are the most common techniques [29].

To optimize CDSS, before constructing CDSS, the authors conducted to determine the features of the CBC, HPLC and iron status needed to screen for thalassemia. After using MID and MDA algorithms, it was found that four parameters included HGB, MCV, MCH and RDW were the most important ones in screening for thalassemia from the database containing over 10,000 cases having CBC, HPLC and iron status results. This is consistent with the recommendation of WHO in screening for thalassemia, in fact, HGB, MCV and MCH are the parameters currently used in screening for thalassemia in clinical practice.

In the research, the authors built an AI-based CDSS including two main components: AI algorithm and the EHR frontend interface.

To determine most appropriate AI algorithm for constructing AI-based CDSS, the authors conducted to train the dataset on four models SVM, KNN, MLP, and RF then evaluated them on four indices: accuracy, precision, recall, and F1-score to choose the most appropriate one.

The obtained result showed that the MLP model was the most stable one regardless of the training database. Particularly, when training with the general CBC and iron status database for four most important features, it had an accuracy of 97.00%, a precision of 92.68%, a recall of 99.23%, an F1 score of 96.20% for original data, and an accuracy of 94.50%, a precision of 88.24%, a recall of 98.68%, and an F1-score of 93.17% for SMOTE. When training with the database of pregnant women, it had an accuracy of 97.81%, a precision of 93.21%, a recall of 99.84%, an F1-score of 95.72% for original data and an accuracy of 97.65%, precision of 93.28%, recall of 99.73%, F1-score of 96.43% for SMOTE. Especially, with only four important features selected above, the results of the model were extremely impressive when the accuracy of MLP increased from 96.91% to 97.81%. Another remarkable thing about this database is that after balancing data with SMOTE, the results had a significant difference between the models. Thus, it is possible to temporarily draw the conclusion that with only the data of the pregnant woman, the model is no longer confounded by the cell blood count data of the husband. When training with the database of husbands, it had an accuracy of 97.44%, a precision of 96.42%, a recall of 100.00% and an F1 score of 97.62% for original data and an accuracy of 94.18%, a precision of 95.44%, a recall of 95.95% and an F1-score of 95.18% for SMOTE.

With such high accuracy, it is entirely possible to apply MLP to construct an AI-based CDSS to predict the risk of thalassemia. In fact, this has been proven by previous research. The study [30] proposed an AI model for thalassemia prenatal screening built by training and evaluating three models including KNN, NB (navie Bayes), and MLP, among which the MLP model got the highest accuracy with 99.73%. The study [31] proposed to apply MLP to build an AI model for thalassemia classification. In the research, the accuracy of the MLP was 98.11%. Like ES, EHR fronted interference of AI-based CDSS also has two sections: input section and output section (**Figure 4**). The difference between the two systems is that ES concludes the type of anemia and whether patients are at high risk or low risk while AI-based predicts the risk of thalassemia in the percentage form.

Despite its advantage, the adoption of AI in medicine is rife with challenges, including the impossibility of explaining the logic that ML uses to make an inference (black box). In AI-based CDSS, the users and researchers can only know the inputs and outputs, but it is challenging to understand how it works inside. Therefore, the accuracy of AI-based CDSS is questionable [17]. To deal with this problem, the authors compared it to ES by

testing both ES and AI-based CDSS with 396 cases of thalassemia. The result showed that when testing on the same dataset, AI-based CDSS got an accuracy of 94.5%, 96%, and 97% (depending on the algorithm used) when using four important features, which is slightly lower than ES with 98.45%. This proves that it is possible to apply AI-based CDSS in screening for thalassemia.

It should be noted that the proposed AI-based CDSS for thalassemia screening is experimental. The advantage of algorithms built on deep ML over physician-based assessments requires more in-depth and comprehensive research.

According to a meta-analysis of publications on the use of AI in CDSS models for various diseases, no advantage of algorithms built on deep ML over physician estimates was noted [32]. It is noted that the effectiveness and safety of AI-based CDSS varies and is ambiguous: there are both successes and failures [33]. Regarding the proposed AI-based CDSSs, it is important to note that future advances in genetic diagnosis of thalassemia may require a significant revision of these CDSS and new studies to confirm the effectiveness and safety of such systems. AI-based CDSSs are an emerging but understudied field, requiring much effort before showing real progress.

CONCLUSION

Based on PCR tests, it was found that among pregnant women and their husbands who came to National Hospital of Obstetrics and Gynecology, the rate of patients with Alpha thalassemia is 10.73% (1,085 patients), the rate of patients with beta-thalassemia is 2.24% (227 patients), and 0.29% (29 patients) of patients carry both alpha-thalassemia and beta-thalassemia gene mutations. The authors successfully built expert and four AI-based CDSS for prenatal screening for thalassemia. ES developed based on WHO and Vietnamese Ministry of Health rules and guidelines for prenatal thalassemia screening showed an accuracy of 98.45%. Among AI-based CDSS developed, the MLP model was the most stable regardless of the training database (accuracy of 98.50% using all features and 97% using only the four most important features). When comparing ES with AI-based CDSS, comparable accuracy of ES and AI-based models was established. Thus, AI-based CDSS showed satisfactory results. Further development of such systems is promising with a view to their introduction into clinical practice.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: This study was supported by the state-level scientific research project: “Research on building an artificial intelligence system to support prenatal screening for some common abnormalities in Vietnam” under the program KC-4.0/19-25.

Acknowledgements: The authors would like to thank the colleagues from the National Hospital of Obstetrics and Gynecology for supporting us during carrying out the research.

Ethical statement: Authors stated that the study has been performed in accordance with the Declaration of Helsinki. The research was approved by the Ethics cCommittee of Vietnam Hospital of Obstetrics and Gynecology with approval number 54/5 of 2020-09-12.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

1. Sheth S, Thein SL. Thalassemia: A disorder of globin synthesis. In: Kaushansky K, Prchal JT, Burns LJ, Lichtman MA, Levi M, Linch DC, editors. Williams Hematology 10. New York, NY: McGraw-Hill Education; 2021. p. 12-3.
2. Gell DA. Structure and function of haemoglobins. Blood Cells Mol Dis. 2018;70:13-42. https://doi.org/10.1016/j.bcmd.2017.10.006 Mid:29126700
3. Viprakasit V, Ekwattanakit S. Clinical classification, screening and diagnosis for thalassemia. Hematol Oncol Clin North Am. 2018;32(2):193-211. https://doi.org/10.1016/j.hoc.2017.11.006 PMid:29458726
4. Petrakos G, Andriopoulos P, Tsironi M. Pregnancy in women with thalassemia: Challenges and solutions. Int J Womens Health. 2016;8:441-51. https://doi.org/10.2147/IJWH.S89308 PMid:27660493 PMCid:PMC5019437
5. Shang X, Xu X. Update in the genetics of thalassemia: what clinicians need to know. Best Pract Res Clin Obstet Gynaecol. 2017;39:3-15. https://doi.org/10.1016/j.bpobgyn.2016.10.012 PMid:27876354
6. Farashi S, Harteveld CL. Molecular basis of α-thalassemia. Blood Cells Mol Dis. 2018;70:43-53. https://doi.org/10.1016/j.bcmd.2017.09.004 PMid:29032940
7. Shah FT, Sayani F, Trompeter S, Drasar E, Piga A. Challenges of blood transfusions in β-thalassemia. Blood Rev. 2019;37:100588. https://doi.org/10.1016/j.blre.2019.100588 PMid:31324412
8. Bain BJ. Delta beta thalassemia. Am J Hematol. 2020;95(1):116. https://doi.org/10.1002/ajh.25623 PMid:31444804
9. Borgna-Pignatti C, Gamberini MR. Complications of thalassemia major and their treatment. Expert Rev Hematol. 2011;4(3):353-66. https://doi.org/10.1586/ehm.11.29 PMid:21668399
10. Thalassaemia International Federation. Response to the proposal for the inclusion of whole blood and red blood cells on the WHO essential medicines lists (EML). Thalassaemia International Federation; 2013. Available at: https://issuu.com/internationalthalassaemiafederation/docs/tif_magazine_issue_62 (Accessed: 13 December 2022).
11. Fucharoen S, Winichagoon P. Haemoglobinopathies in Southeast Asia. Indian J Med Res. 2011;134:498-506.
12. Nguyen KHH. Thalassemia and a model of prevention in Vietnam. Sydney: The University of Sydney; 2005.
13. Dossanova A, Lozovoy V, Wood D, Lozovaya Y. Reducing the risk of postoperative genital complications in male adolescents. Int J Environ Sci Educ. 2016;11(13):5797-807.
14. Thalassaemia International Federation. Guidelines for the management of transfusion dependent thalassaemia. Thalassaemia International Federation; 2021. Available at: https://thalassaemia.org.cy/publications/tif-publications/guidelines-for-the-management-of-transfusion-dependent-thalassaemia-4th-edition-2021/ (Accessed: 13 December 2022).
15. Dossanov B, Trofimchuk V, Lozovoy V, et al. Evaluating the results of long tubular bone distraction with an advanced rod monolateral external fixator for achondroplasia. Sci Rep. 2021;11(1):14727. https://doi.org/10.1038/s41598-021-94146-z PMid:34282216 PMCid:PMC8290032

16. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. BMJ. 2005;330:765. https://doi.org/10.1136/bmj.38398.500764.8F PMid:15767266 PMCid:PMC555881
17. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med. 2020;3:17. https://doi.org/10.1038/s41746-020-0221-y PMid:32047862 PMCid:PMC7005290
18. Muhiyaddin R, Abd-Alrazaq AA, Househ M, Alam T, Shah Z. The impact of clinical decision support systems (CDSS) on physicians: A scoping review. Stud Health Technol Inform. 2020;470-3.
19. Amendolia SR, Brunetti A, Carta P, et al. A real-time classification system of thalassemic pathologies based on artificial neural networks. Med Decis Making. 2002;22(1):18-26. https://doi.org/10.1177/0272989X0202200102 PMid:11833662
20. Masala GL, Golosio B, Cutzu R, Pola R. A two-layered classifier based on the radial basis function for the screening of thalassaemia. Comput Biol Med. 2013;43(11):1724-31. https://doi.org/10.1016/j.compbiomed.2013.08.020 PMid:24209918
21. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. J Artif Int Res. 2020;16(1):321-57. https://doi.org/10.1613/jair.953
22. Dang HTT. Screening for thalassemia in pregnant women who come for examination and treatment at the National Hospital of Obstetrics and Gynecology. Vietnam: National Hospital of Obstetrics and Gynecology; 2019.
23. Nguyen HHK. Research for alpha and beta thalassemia screening and diagnosis. Vietnam: Ho Chi Minh City University of Medicine and Pharmacy; 2013.
24. He S, Zhang Q, Li D, et al. Prevention and control of Hb barts disease in Guangxi Zhuang Autonomous Region, China. Eur J Obstet Gynecol. 2020:178:138-41. https://doi.org/10.1016/j.ejogrb.2014.03.034 PMid:24792538
25. Husna N, Handayani NSN. Molecular and haematological characteristics of alpha-thalassemia deletions in Yogyakarta special region, Indonesia. Rep Biochem Mol Biol. 2021;10(3):346-53. https://doi.org/10.52547/rbmb.10.3.346 PMid:34981010 PMCid:PMC8718782
26. Sarafidis M, Manta O, Kouris I, et al. Why a clinical decision support system is needed for tinnitus? Annu Int Conf IEEE Eng Med Biol Soc. 2021;2021:2075-8. https://doi.org/10.1109/EMBC46164.2021.9630137 PMid:34891697
27. Klar R, Zaiß A. Medical expert systems: Design and applications in pulmonary medicine. Lung. 1990;168(1):1201-9. https://doi.org/10.1007/BF02718262 PMid:2117122
28. Rowe M. An introduction to machine learning for clinicians. Acad Med. 2019;94(10):1433-6. https://doi.org/10.1097/ACM.0000000000002792 PMid:31094727
29. Shaikh F, Dehmeshki J, Bisdas S, et al. Artificial intelligence-based clinical decision support systems using advanced medical imaging and radiomics. Curr Probl Diagn Radiol. 2021;50(2):262-7. https://doi.org/10.1067/j.cpradiol.2020.05.006 Mid:32591104
30. AlAgha AS, Faris H, Hammo BH, Ala'M AZ. Identifying β-thalassemia carriers using a data mining approach: The case of the Gaza Strip, Palestine. Artif Intell Med. 2018;88:70-83. https://doi.org/10.1016/j.artmed.2018.04.009 PMid:29730048
31. Egejuru NC, Olusanya SO, Asinobi AO, Adeyemi OJ, Adebayo VO, Idowu PA. Using data mining algorithms for thalassemia risk prediction. Int J Biomed Sci Eng. 2019;7(2):33. https://doi.org/10.11648/j.ijbse.20190702.12
32. Liu X, Faes L, Kale AU, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. Lancet Digital Health. 2019;1(6):e271-97. https://doi.org/10.1016/S2589-7500(19)30123-2 PMid:33323251
33. Roshanov PS, Fernandes N, Wilczynski JM, et al. Features of effective computerised clinical decision support systems: metaregression of 162 randomised trials. BMJ. 2013;346:f657. https://doi.org/10.1136/bmj.f657 PMid:23412440