Development of an algorithm for the transition from regression models to optimization models in the differential diagnosis of iron deficiency anemia and β-thalassemia in children in the Republic of Artsakh

HARUTYUNYAN IREN VAZGENI
Department of Natural Science, Communications and Information Technologies
Shushi University of Technology, 375000 Stepanakert, V.Vagharshyan 7
REPUBLIC OF ARTSAKH

AYDINYAN LUSINE GARNIKI
Department of Agronomy, Veterinary Medicine and Agricultural Mechanization
Shushi University of Technology, 375000 Stepanakert, V.Vagharshyan 7,
REPUBLIC OF ARTSAKH

LALAYAN ERNEST LAVRENTI
CJSC Medical Association "Arevik", 375000 Stepanakert, A.Mkrtchyan, 4
REPUBLIC OF ARTSAKH
GALSTYAN GYULCHORA ANUSHAVANI
Department of Natural Sciences, Communications and Information Technologies
Shushi University of Technology, 375000 Stepanakert, V.Vagharshyan 7
REPUBLIC OF ARTSAKH

HARUTYUNYAN ROBERT MISAKI
Department of Natural Sciences, Communications and Information Technologies,
Shushi University of Technology, 375000 Stepanakert, V.Vagharshyan 7
REPUBLIC OF ARTSAKH

Abstract: Mathematical tools for studying anemia course have not yet been fully developed. The main clinical diagnostic and anamnestic indicators of a sick child are the dominant point of view. In order to better predict anemia, we aim to comprehend the role played by mathematical tools in these studies. This paper presents a methodology for analyzing anamnestic, clinical, and laboratory data that designed for assessing the fundamental mathematical techniques using a variety of theoretical foundations such as the probability theory, mathematical statistics, and other mathematical forecasting methods using specialized application packages. The paper highlights some key aspects of predicting childhood anemia in the Republic of Artsakh.

Keywords: mathematical model, registration model, optimal model, IT, Mentzer Index, coefficient of determination, β − thalassemia, adjusted coefficient of determination.

Received: June 21, 2022. Revised: June 4, 2023. Accepted: July 24, 2023. Published: August 28, 2023.

1. Introduction

A. Differential diagnosis of IDA and β – thalassemia. Therefore, it is important to create parameters that are simple to survey and useful for differentiating between β-thalassemia and anemia in children. The study is aimed at clarifying the diagnostic value of the Mentzer index in predicting anemia in children compared with the results of the HGB electrophoresis study. The data were taken from 130 patients with anemia in the «Arevik» Medical Association CJSC in the Republic of Artsakh. In addition, the Mentzer index was calculated. The Mentzer index value was >13 (0.5–11.5, CI 95%). The Mentzer index can be used not only for the differential diagnosis of β-thalassemia and anemia, but also to predict the course of anemia in automatic blood cell counters, children, depending on the main markers of anemia.

The primary indicators for the diagnosis of anemia are based on blood parameters obtained with which traditionally measured the parameters of hemoglobin (HGB), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), erythrocyte distribution width (RDW), mean corpuscular volume . hemoglobin concentration (MCHC) and erythrocyte count (RBC). Numerous studies have looked at indicators of diagnostic precision with varying results, and none of these indicators has demonstrated 100%.

2. Problem Formulation

2.1. Diagnosis of anemia and β −thalassemia

The problem of iron deficiency states in pediatrics is urgent due to its widespread prevalence as well as the serious negative effects iron deficiency has on children's health. Numerous organs and bodily systems have been shown to malfunction as a result of iron deficiency. This is because iron is a component of numerous proteins, such as hemoglobin, myoglobin, cytochromes, iron seroproteins, oxidases, hydroxylases, and superoxide dismutases, which support cellular and systemic aerobic metabolism as well as overall redox homeostasis in the body. It is known that low iron levels in the body negatively impact metabolic processes, which in turn cause disruptions in the functioning of various organs and systems.

The differential diagnosis of β-thalassemia requires genetic testing using a polymerase chain reaction (PCR) device or hemoglobin (HGB) electrophoresis using electrophoresis or High Performance Liquid. Chromatography device (HPLC), available only in some hospitals

2.1.1. Statistical Analysis

The clinical prediction model can be used in a variety of complex clinical situations, inclu-ding the screening of people who are highly susceptible to asymptomatic disease, predicting future outcomes like the course of a disease or death, and supporting clinical decision-making. Although clinical prediction models have a positive effect on practice, prediction modeling is a complex process that necessitates careful statistical analysis and thoughtful clinical debate. Although there isn't a unanimous agree-ment on the best method for creating and valida-ting models, various scenarios and suggestions have been taken into account. In this review, we summarize the five steps for creating and validating a clinical prediction model, including planning ahead, selecting a data set, working with variables, creating the model, and evaluating and validating the model.

Good models can be prepared for use in practice once they have been developed and carefully tested under the right circumstances, possibly with usefulness evaluation and fine tuning. By reviving the use of predictive models or predictive studies to forecast anemia in children, we anticipate that this framework will encourage their active implementation in actual clinical practice.

3. Problem Solution

3.1. Why regression analysis is important
The data, which are the numbers and statistics that actually define the subject area, are what make regression analysis so important. Regression analysis has the advantage of allowing significantly reduce scores (based on optimization techniques), which will aid in improving our judgment when determining how anemia will develop in the present and the future.

The regression analysis method involves the study of relationships between data points, which can help us:

✓ ***To build regression models of the course of iron deficiency anemia in children.***

✓ ***Understanding the role of informative indicators (determination of control and controlled variables at this stage) in the diagnosis of anemia.***

✓ ***Determine the child's condition (according to anamnesis, clinical, laboratory and biochemical parameters) and subsequently determine the severity of the child's condition with anemia.***

✓ ***Observe and record how changes in various variables (indicators and signs) affect the course of the disease.***

✓ ***After determining the governing and governed variables, develop an optimal control model (with a certain probability) for the course of the disease, the severity of the condition of the baby, and conduct a differential diagnosis of IDA and beta-thalassemia.***

3.1.1.How regression analysis involves the study of relationships between data points, which can help us:

Suppose we want to predict future values of informative indicators, which are the main markers for diagnosing anemia in children, hemoglobin, mean corpuscular hemoglobin, erythrocyte distribution width average, corpuscular volume, hemoglobin concentration, and red blood cell count and .

Then some of the acquired informative indicators will turn out to be dependent variables, because they "depend" on the condition of the child and the course of the disease, which is an independent variable. (The independent variable is the one against which you measure something for comparison in this case, the main informative markers of anemia.) We must determine how closely these dependent and independent variables are related. Listed in Fig.1 is a matrix of correlation coefficients.

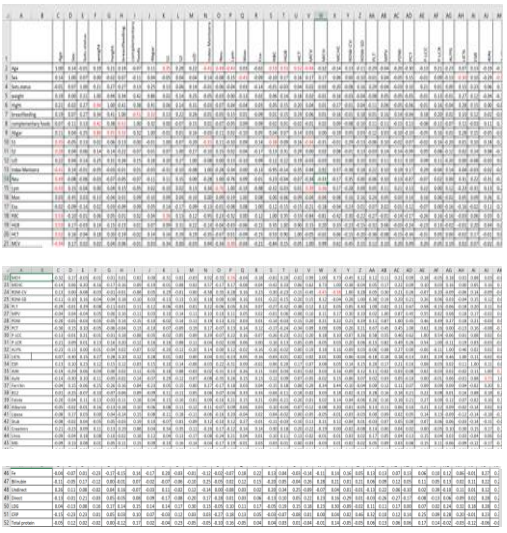

Figure 1. Matrix of correlation coefficients

The criterion for differentiation between anemia and other hematological diseases in children who often coincide the main markers of diseases is the Mentzer index, first described in 1973, is the mean corpuscular hemoglobin ratio to the number of red blood cells. They say that it helps to distinguish iron deficiency anemia from beta-talassemia. The index is calculated based on the results of a general blood test.

In our study, we continued to identify the role of the Mentzer index as a discriminatory test for differentiating iron deficiency anemia from a sign of beta thalassemia. The high-risk group can then be subjected to definitive diagnostic tests. This can lead to better patient compliance and cost-effectiveness. To determine the risk group, we used the apparatus of regression analysis. This apparatus has also been used to predict the values of the Mentzer index.

Below are the results of applying the regression analysis.

3.2. Regression analysis results

Dependence of the Mentzer index on Age, Abgar and Diagnosis

Regression Analysis: M versus Age; Abgar; Diagnosis

Multiple regression results (Step 2)
Dependent variable.:M
Multiple R = ,99 F = 1886
R^2= ,98 DW = 2,96
Number of observations.: 130
Adjusted R^2= ,975 p = 0,000
Standard error of estimation: 2,83

Abgar бета=,528 Diagnosis бета=,463

Table 1. Dependent variable regression totals M

R= ,99 R^2= ,975 Adjusted R^2= ,98 F=1886,4 p=0.000

	Beta	Standart Error - Beta	B	St.Er. - B	t(96)	p-signif.
Abgar	0,528	0,098	1,26	0,233	5,403	0,000
Diagnosis	0,463	0,098	7,99	1,687	4,735	0,001

Table 2. Analisys of Variance

	Sum of Squares	Standart deviation	F	p-signif.
Regression	30192,20	15096,10	1886,370	0,00
Remains	768,26	8,00		
Total	30960,46			

The model presented below reflects the relationship between the Mentzer index (as a control variable), on the one hand, and platelets and hemoglobin, on the other (at this stage of identifying control variables: they are control variables).

Results multiple regression (Step 2)
Dep.var. :M
Multiple R = ,99 F = 2547

R^2= ,98 DW = 2,103
Number of Observations: 130
Adjusted R^2= ,98 p = 0,000
Standard error of estimation: 2,52
MCH бета=,510 HGB бета=,485

Table 3. Dependent variable regression totals M

R= ,99 R^2= ,98, Adjusted R^2= ,98 F=2547,0 p=0.000

	Beta	St.Er. - Beta	B	St.Er. - B	t	p-signif.
MCH	0,51	0,074	0,343	0,049	6,93	0,000
HGB	0,485	0,074	0,078	0,012	6,59	0,000

Table 4. Analisys of variance

	Sum of Squares	Standard deviation	F	p-signif.
Regression	32267,08	16133,54	2546,98	0,00
Remains	652,44	6,33		
Total	32919,52			

Table 5. Dependent variable regression totals M

R= ,96, R^2= ,92 Adjusted R^2= ,92 F=260,93 p=0.000

	Beta	St.Er. - Beta	B	St.Er. - Beta	t	p-signif.
P-LCC	0,42	0,08	0,15	0,028	5,29	0,0001
Lym	0,37	0,08	0,13	0,029	4,53	0,0024

Table 6. Analysis of Variance

	Sum of Squares	Standart Deviation	F	p-signif.
Regression	20644,26	6881,42	260,93	0,000
Remains	1846,09	26,37		
Total	22490,34			

To build an optimal model, we use the adequate models obtained above (they reflect real processes with a probability much higher than 0.5). We have obtained linear models, where the control variables describe the control variable with high confidence.

Our objective function: the value of the Mentzer index will tend to the maximum, i.e. as a result, we should get a value of the Mentzer index greater than 13. For the diagnosis of β-thalassemia, we will consider the optimal model developed by us as a solution to the dual problem of linear programming.

Below we present the objective function with restrictions:

$$M = 0{,}51x_1 - 0{,}49x_2 \rightarrow max$$

$$\begin{cases} 1{,}26x_1 + 7{,}99x_2 \geq 13 \\ 0{,}42x_1 + 0{,}37x_2 \geq 13 \\ x_1, x_2 \geq 0 \end{cases}$$

MCH and HGB are presented as control variables in the objective function.

In restrictions, respectively: in the first restriction - x_1-value from the Abgar scale, x_2 −diagnosis, which changes its value depending on the following: anemia is a concomitant disease or the main one. In the second constraint, we obtained P-LCC and Lym as control variables. That is, the identified informative variables (at this stage of the development of the optimal model, they act as control variables) will have values greater than zero. They will change with age. But the value of the Mentzer index does not take into account age characteristics.

All developed models are adequate and of good quality: the coefficient of determination of models, more precisely, the adjusted coefficient of determination of models is above 85%.

	Number of observations	Mean	Confidence interval 95 %	Confidence interval 95%	Geometric Mean	Median	Minimum	Maximum	StDev
Diagnosis	130	1,02	0,99	1,048	1,014	1,00	1,00	2,0000	0,14
Abgar	130	7,42	7,26	7,57	7,37	7,00	6,00	9,0000	0,80
Lym	130	41,18	37,16	45,20	31,55	44,91	0,64	78,6167	20,69
RBC	130	4,26	4,18	4,33	4,24	4,30	3,25	5,5800	0,38
HGB	130	109,21	107,44	110,98	108,85	108,00	93,25	142,0000	9,14
MCV	130	73,56	72,38	74,74	73,31	73,63	56,03	87,9333	6,09
MCH	130	25,86	24,88	26,83	25,52	25,75	16,70	67,5833	5,04
PLT	130	303,69	286,94	320,43	288,72	303,50	27,40	569,0000	86,52
P-LCC	130	46,05	42,35	49,75	43,44	42,25	15,00	95,0000	16,19
M	130	17,5	16,96	18,03	17,29	17,21	12,00	24,4769	2,74

Fig. 2. Confidence interval for infor-mative indicators and main markers of diseases

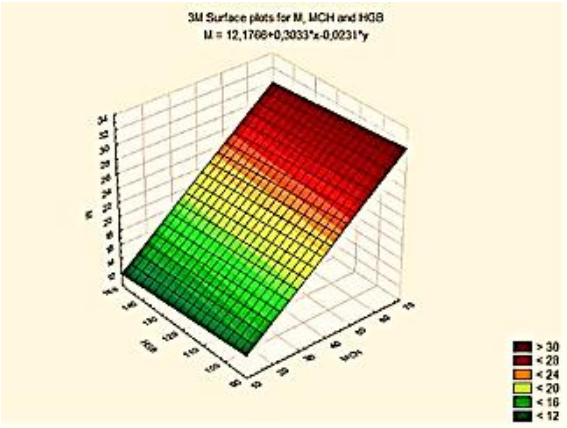

Fig. 2. 3M surface plots for M: dependent on MCH and HGB

4. Conclusion

There is no single optimal marker or combination of tests for the differential diagnosis of anemia. The physician's knowledge and experience, requiring appro-priate hematological and biochemical analy-zes associated with a preliminary diagnosis, play an important role in the diagnosis of anemia. Algorithms are recommended as a tool for predicting the course of anemia to reduce the number of laboratory tests and accurately diagnose the underlying cause(s) in patients.

Over the past decade, significant progress has been made in methods and algorithms for predicting the course of anemia in children. A complete blood count is the main method for studying anemia. At the first stage, the percentage of microcytic erythrocytes is taken into account. The second step is to check the number of MCV, MCH, RBC, Lym, HGB and PLT. The third step is to determine the value of the Mentzer index and predict it. The main purpose of this study was to predict the course of anemia in children, based on the main values of the Mentzer index (the main criterion for differentiating anemia and β-thalassemia).

References

[1]. Aivazyan S.A., Mkhitaryan V.S. The theory of probabilities and applied statistics.Vol. 1,/Applied statistics. Funda-mentals of econometrics. Textbook for higher educational institution, M., 2001, 656pp.

[2]. Balashova E.L., Mazur L.I. Modern approaches to the diagnosis of iron defi-ciency anemia in children. Russian Bulletin of Perinatology and Pediatrics. 2015; 60(4):31-36pp.

[3]. Korovina N.A., Zaplatnikov A.L., Zakharov I.N., Iron deficiency anemia in children. M., 1999.

[4]. 4.Gordeeva O.B., Botvineva V.V. Modern concepts of iron deficiency anemia in children and optimization of treatment. M., Press. 2004, 216pp.

[5]. Botvineva V.V., Gordeeva O.B., Namazova-baranova L.,S., Prospects for the diagnosis and treatment of various types of anemia in children. Pediatric pharmacology, 2012, 9(5): 35-40pp.

[6]. Rumiancev A.,G., Korovina N.A., et al. Diagnosis and treatment of iron deficiency anemia in children. M., 2004. 48pp.

[7]. Ebru Dundar Yenilmez and Abdullah Tuli. Laboratory Approach to anemia. December 20th 2017.

[8]. www.stat-nkr.am