Proteomics & Bioinformatics: Current Research
PBCR, 1(1): 17-28
www.scitcentral.com

Original Research Article: Open Access

Genotype-Phenotype Correlation and the Severity Index for β-Thalassemia

Mohammad MS Al-Haggar[1], Mohammad Emad Rasmy[2], Mohammad AM Islam[2], Darwish A[3] and Balkis A Khair-Allah[4*]

[1]*Genetics Unit, Pediatrics Department, Faculty of Medicine, Mansoura University, Egypt*
[2]*Systems and Biomedical Department, Faculty of Engineering, Cairo University, Egypt*
[3]*Hematology Unit Pediatrics Department, Faculty of Medicine, Mansoura University, Egypt*
[*4]*Biomedical Engineering Department, Mansoura University Children Hospital, Egypt.*

Received August 17, 2018; Accepted September 02, 2018; Published January 12, 2019

ABSTRACT

β-thalassemia is a well-known and major public health problem. There are many factors contributing to the diagnosis and management of the disease. The advance of high throughput techniques utilizing the biomedical informatics (BMI) give chances to prepare the current century for the development of translational and personalized medicine using decision support systems (DSSs). In this article, quantitative analysis of the interdependency between these factors is presented, using the Multi-Criteria Decision-Making (MCDM) techniques. Consequently, a severity index for βeta thalassemia is developed. Eventually, classes for the iron overload risk factor for blood transfusion dependent patients are estimated.

Keywords: Translated bioinformatics, Decision support system (DSS), Multi-criteria decision-making (MCDM), Decision making trial and evaluation laboratory (DEMATEL), Analytical network process (ANP), βeta thalassemia, Genotype/phenotype correlation, Severity index, Iron overload

INTRODUCTION

Success in life sciences obliges us to adopt advances in informatics. Biomedical informatics (BMI) is the scientific field that deals with the storage, retrieval, sharing and optimal use of biomedical data. The field of BMI applied to biomedicine for problem solving and decision making in order to improve human health [1-3].

Bioinformatics (molecules and cells), imaging informatics (tissue and organs), clinical informatics (individuals or patients) and public health informatics (populations and society) are the aspects of biomedical informatics related to translational medicine [4]. Translational Bioinformatics involves the development and use of computational methods with life science data being collected and stored for the purpose of creating new tools for medicine. While bioinformatics methodologies have been used to enable biological discoveries, here the end product has to be translational or applying to human health and disease [5,6]. Translational informatics is now a promising methodology that can drive the translation of laboratory data at the bench to health gains at the bedside. Also, as personalized medicine stated, information about individual and management guidelines help to enable delivery of the right treatment to the right patient, at the right time.

βeta thalassemia (β-thalassemia) is a well-known inherited hemoglobin disorder, as the number of babies born every year with inherited hemoglobin disorders is about 320,000 babies, causing a global public health problem [7]. According to the statistics, 80% of the previous number appears in the third world countries. On the other hand, several studies have proved that the β-thalassemia alone constitute about 3-4% of the major hemoglobin disorders with an estimate of around 8,000-10,000 new births with the most severe form of diseases each year [8]. In addition, β-thalassemia is being carried with a global frequency of 1.5% [7]. In Egypt, a high rate of carriers has been reported ranging from 4-5% and reaching up to 9-10% [9]. The accumulation of β-thalassemia defected genes in families is due to the high rate of consanguineous marriage in Egypt [9] **(Figure 1)**.

Corresponding author: Balkis A Khair-Allah, Biomedical Engineering Department, Mansoura University Children Hospital, Egypt, E-mail: balkisahmed60@yahoo.com

Citation: Al-Haggar MMS, Rasmy ME, Islam MAM, Darwish A & Khair-Allah BA. (2019) Genotype-Phenotype Correlation and the Severity Index for β-Thalassemia. Proteomics Bioinformatics Current Res, 1(1): 17-28.

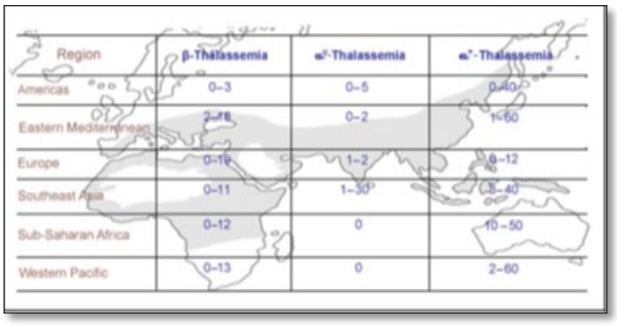

Figure 1. Epidemiology of thalassemia: Carrier frequencies of thalassemia alleles (%) [9].

Taking into consideration the general six activities of medicine: screening, diagnosis followed by severity assessment, treatment, prognosis, monitoring and follow-up; making the decisions in the medical field is considered as a very complex and sophisticated process. High epidemic and mortality rate, expensive tests, time consuming, confusion, incomplete information and requirement of special experience, are the main reasons behind the medical decision-making problems [10]. Therefore, solving the previous problems shall improve the medical diagnosis and the disease management. Furthermore, these solutions will enhance the patient care by limiting the follow-up time and identifying new targets for therapies.

Traditional medicine uses neither genotype information, nor decision support systems, nor severity index. The integration of genomic information in the patients' electronic health records (EHRs) can confirm or roll out the existence of the disease [10-12]. In addition, the use of automated decision support systems is perfect for solving the complexity problems in the medical diagnosis, due to its ability to deal with multiple factors or variables [10-12]. In this context, applying the Multi-Criteria Decision-Making (MCDM) approaches that depend on small data sets and experts reduce the need for huge datasets required to develop models for diagnosis. In addition, the nature of decision-making trial and evaluation laboratory (DEMATEL) and analytic network process (ANP) MCDM approaches are suitable for studying the correlation between the interrelated factors. Moreover, having a general scoring system for severity assessment can help in selecting the appropriate treatment and follow-up procedures. Finally, dividing the risk of iron overload for β-thalassemia patients who receive blood transfusion into classes will facilitate adjusting the iron chelation dose.

The previous works concerning the development of decision support systems for β-thalassemia usually depended on huge datasets, different methods of learning techniques and data mining [13-19]; including statistical methods [20,21]. In addition, it was discovered that the most common mutations in Gaza have led to some difference in the biochemical and hematological features of β-thalassemia [20]. A phenotypic severity scoring model for the non-transfusion dependent thalassemia (NTDT) patients has been formulated before [21]. Typically, the researchers used to study the relationship between the factors that contribute to the disease by determining the significance of each factor. On the other hand, the current study is the first study to discriminate between the cause and effect factors for the β-thalassemia disease. The factors considered include genetic modifiers, patient clinical features and patient history criteria.

The work presented in the current article does not seek to replace experts or clinicians as a patient's regular medical specialist. Instead, a decision support system is developed in order to minimize errors and to save time during β-thalassemia diagnosis. This system includes as an assessment scoring system for the βeta Thalassemia severity and an estimator of the iron overload risk classes resulting from the blood transfusion. The system applies MCDM methods, namely, DEMATEL and ANP. The goal of applying the DEMATEL method is to discover the relationship between the βeta Thalassemia severity factors and to estimate the degree of influence each factor has on the others. In addition, the ANP method estimates the relative contribution of different factors to the severity. The above contributes to the design of the severity index estimator, along with the iron overload classes.

The rest of this article illustrates the steps of the integration of DEMATEL and ANP as a hybrid MCDM approach; it also gives the structure of our proposed decision support system and illustrates its application to β-thalassemia; and also includes the detailed discussion of the implementation of our work and its results; lastly, presents the conclusions regarding the proposed severity index and the iron overload classes.

THE MCDM APPROACHES

The approaches of MCDM are part of the decision-making processes; a well-known branch which is characterized by multiple and conflicting factors [22]. MCDM started in the 1960s as a sub-discipline of operational research (O.R) [22-27]. The operational research targets the practical problems in marketing, manufacturing, transportation, management, engineering science, information technology (IT) and other fields [28,29]. Through the use of the MCDM mathematical algorithms, it is easy to transform the qualitative measurements into quantitative data, through a computational model for the evaluation of decision alternatives [22]. The following sections illustrate a brief overview of classical DEMATEL and classical ANP approaches used in this work.

The decision-making trial and evaluation laboratory (DEMATEL) approach

The DEMATEL approach is one of the newest versions of the MCDM mathematical algorithm. It first appeared in the Geneva Research Centre of Battelle Memorial Institute, and

it was designed to solve the interrelation in real world problems [30]. The DEMATEL approach differentiates the variables into two featured groups, one for the cause group and the other for the effect group [30]. The DEMATEL approach can be illustrated as follows [31-33]:

Step 1: Calculating the initial average matrix or the direct relation matrix (A): This matrix reflects the experts average estimate of the degree of the direct influence of each criterion/sub-criterion i (factor i) on criterion/sub-criterion j (factor j). This estimate is element aij of matrix [A]. The influence is estimated using a scale of 0 to 4, starting from 0 for "no influence", and gradually ascending to 4 for "very high influence". The general form of the initial average matrix is shown in **Figure 2**.

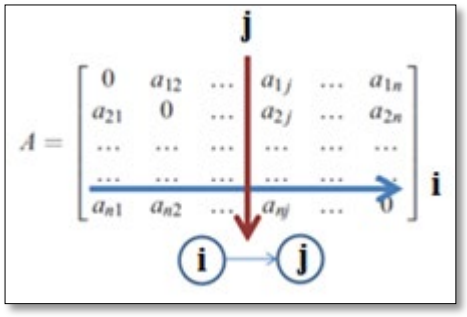

Figure 2. The resultant matrix [A].

Step 2: Calculating the normalized initial direct relation matrix (D): Elements of matrix A are divided by the maximum total influence exerted by any criteria (factor), as given by Equations (1) and (2).

$$S = \frac{1}{\max_{1\le i\le n}\sum_{i=1}^{n} a_{ij}, \max_{1\le i\le n}\sum_{j=1}^{n} a_{ij}} \quad (1)$$

$$D = A \cdot S \quad (2)$$

Step 3: Deriving the total (direct and indirect) influence matrix or the total relation matrix (T), T can be obtained using Equation (3).

$$T = D(I - D)^{-1} \quad (3)$$

Where, $T=[t_{ij}]_{n\times n}$, for i, j=1, 2, …, n and (I) is the identity matrix.

Step 4: Calculating the prominence and relation of different factors: Define the two vectors R and C, whose elements are given by equations 4 and 5, respectively. The values of r_i and c_j reflect the sum of the total effects, both direct and indirect, each factor has on, and received from, the other factors, respectively. In other words, r_i is the i^{th} row sum in the matrix T and displays the sum of the direct and indirect effects dispatching from factor F_i to the other factors. Similarly, c_j is the j^{th} column sum in the matrix T and depicts the sum of direct and indirect effects that factor F_j is receiving from the other factors.

$$r_i = \sum_{j=1}^{n} t_{ij} \quad (4)$$

$$c_j = \sum_{i=1}^{n} t_{ij} \quad (5)$$

Thus, when i=j, (r_i+c_i) refers to the degree of importance that factor i plays in the system while $(r_i - c_i)$, refers to the net relation that factor i contributes to the system. When $(r_i - c_i)$ is positive, factor i is a net causer, while, when $(r_i - c_i)$ is negative, factor *i* is a net effector.

Step 5: Reducing of the system's complexity, by removing the elements of matrix T whose influence is below a certain threshold α, obtaining T_α. This threshold is the average of the elements in matrix T [34].

Step 6: Building the impact relation map (IRM): This map visualizes the relationship between different factors, by signing all coordinates sets (r_i+c_i) the horizontal axis and $(r_i - c_i)$ the vertical axis on the diagram. Finally, the arrows of the diagram represent the elements of the reduced complexity matrix T_α.

The analytic network process (ANP) approach

Recently, the ANP approach has become one of the most famous and powerful methods of the MCDM techniques [35,36]. Through network features between the elements in each decision level [37,38], the ANP approach solved the complex decision-making problems with dependencies [35,39,40]. The following lines describe the steps related to the ANP method:

Step 1: Constructing network structure: The creation of the network means defining network elements and relations among these elements. These elements are expressed by criteria (clusters) and sub-criteria (nodes). In addition, relations represent the influences between elements. The network structure can be obtained by decision-makers through brainstorming or other appropriate methods [35].

Step 2: Applying pairwise comparisons between elements: After obtaining the network structure pairwise comparisons are performed to determine the relative importance weights (relative priorities/priority eigenvector) of compared elements using Saaty's method [41,42]. The resultant comparison matrices are termed by the relative importance matrices. To obtain the previous matrices, a group of experts were asked to provide two sets of pairwise comparisons: Node comparisons and cluster comparisons. The first set compares nodes within the clusters based on their influences on a node in another cluster where they were linked. The second set compares the clusters themselves with respect to their contribution to the goal. These comparisons were based on Saaty's scale ranging between 1 (the equal importance) to 9 (the extreme importance) [43].

Before proceeding to Step 3 for each comparison matrix the consistency of expert's opinions must be checked. The consistency test will be performed based on the consistency ratio (C.R) [35,41,42]. So, the CR is a measure of consistency of each expert. If CR ≤ 0.1, then the estimate is

accepted; otherwise, a new comparison matrix is repeated until CR ≤ 0.1.

Step 3: Forming the unweighted and weighted super matrices: The first type of these matrices is the unweighted super matrix, by using relative importance weights (relative priorities) that are derived from the nodes pairwise comparisons obtained from the previous step. The description of the general form of the super matrix is shown in Equation (6):

$$W = \begin{array}{c} \begin{matrix} \quad C1 & \quad C2 & \cdots & Cm \\ e11 \ldots e1n_1 & e21 \ldots e2n_2 & \cdots & em1 \ldots e2mn_n \end{matrix} \\ \begin{matrix} C1 \\ \vdots \\ C2 \\ \vdots \\ Cm \end{matrix} \begin{matrix} e11 \\ e12 \\ \vdots \\ e1n_1 \\ e21 \\ e22 \\ \vdots \\ e2n_2 \\ \vdots \\ em1 \\ em2 \\ \vdots \\ emn \end{matrix} \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1m} \\ M_{21} & M_{22} & \cdots & M_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mm} \end{bmatrix} \end{array} \quad (6)$$

Whereas, C_m represents the m^{th} cluster, e_{mn} indicates the m^{th} node in the m^{th} cluster.

Each column in matrix M_{ij} is the relative importance weights (relative priorities) of elements, which are compared in pairs between the i^{th} and j^{th} clusters. The second type of matrices is the weighted super matrix *W'*. The weighted super matrix is obtained by multiplying all the elements in a component of the unweighted super matrix by the corresponding cluster weight [36].

Step 4: In this final step, the weighted super matrix is multiplied by itself for K times using Equation (7) until the values in this matrix stabilized enough to obtain the overall priorities in a limiting super matrix W", which is considered as the third type of super matrix.

$$W'' = \lim_{k \to \infty} = \sum_{k=1}^{N} (w')^k \quad (7)$$

BETA THALASSEMIA DECISION SUPPORT SYSTEM

A hybrid DEMATEL-ANP MCDM system was built. The system estimates the severity of β-thalassemia, and the risk of iron overload. The DEMATEL method divides the β-thalassemia severity factors into the cause-and-effect groups. The ANP is a feedback method which replaces the hierarchy with a network when dealing with inner relations [44].

In general, the proposed decision-making model depends on four major phases as follows: (1) identifying the target, stating the problem, defining the model elements, and constructing the model; (2) building a description of the clusters' relations by using the DEMATEL; (3) calculating the total weights by applying the ANP steps through the IRM being constructed in (2); and (4) obtaining the severity index and classes of iron overload risk factors.

Definitions and system construction

The elements of the proposed model were identified and classified into multiple criteria and sub-criteria. All of these elements were suggested based on expert's opinion, in addition to literature and guidelines [44-48]. Three main criteria along with their ten sub-criteria for the assessment of severity and the risk of iron overload were derived and are listed in **Table 1**.

Table 1. The proposed criteria and sub-criteria.

Criteria (P)	Sub-Criteria (e)
Genotype Modifiers (P1)	Common Mutation Type (e1)
	β-Genotype (e2)
	α-Genotype (e3)
	HbF Response (e4)
Phenotype Factors (P2)	Initial Presentation (e5)
	Initial Hb (e6)
	Blood Transfusion Dependency (e7)
Patient History (P3)	Consanguinity (e8)
	Family History (e9)
	Age @ 1st Clinical Features (e10)

On the other hand, eight alternatives (categories of β-thalassemia severity) are assessed. These alternatives are: Severe Class (I), Severe Class (II), Severe Class (III), Severe/Moderate, Moderate, Moderate/Mild, Mild and Silent. Class (I) reflects high risk of iron overload, Class (II) reflects moderate risk, and Class (III) reflects low risk.

According to **Figure 3**, the developed decision system integrates two forms of structure, the hierarchal structure and the network structure. The hierarchal structure consists of four levels. The top level is the main goal of the decision problem to obtain β-thalassemia severity index (SI). The three main criteria and the ten sub-criteria are included in the intermediate levels, where the main criteria contribute to the goal. The bottom level includes the eight alternatives evaluated in terms of the sub-criteria. The concept of the hierarchal structure assumes that the previous levels are independent. The network structure appears in the intermediate levels assuming interdependency occurs between criteria and sub-criteria.

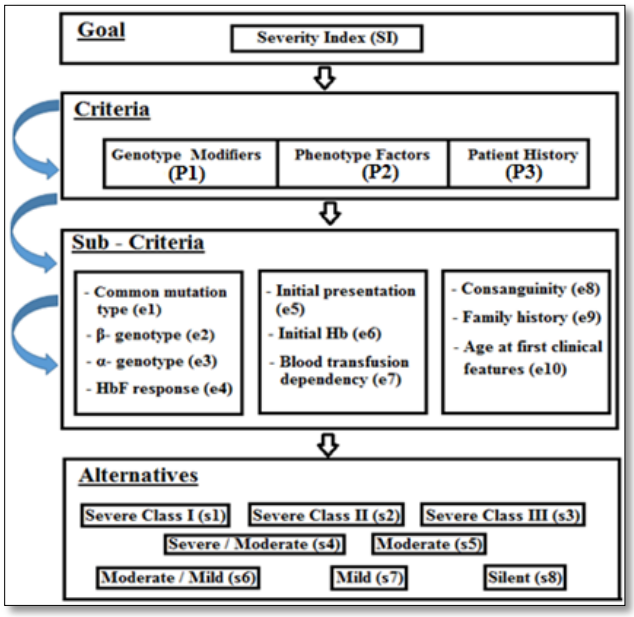

Figure 3. Hierarchy and network of the general model.

Implementation of DEMATEL and ANP approaches

The DEMATEL steps were implemented using Excel sheet. Accordingly, in DEMATEL approach, six experts were able to express their opinion of effects (direction and intensity) among factors of severity assessment with more control. Also, the DEMATEL approach was used to determine the Influential Relation Matrix and the Impact Relationship Map (IRM) by implementing equations from 1 to 6. In the proposed integrated approach, the ANP network structure through factors 'mutual influence is obtained by experts using results from DEMATEL by implementing equations 6 and 7.

Evaluation of alternatives and final index

The β-thalassemia severity index was derived by the evaluation of the alternatives using the desirability index (Di) approach [49-52]. The desirability index (Di) approach integrates all weights resulted from different types of pairwise comparisons to obtain the final priority weight of the whole model. Desirability index Di for the alternative (i) is defined as follows:

$$D_i = \sum_{j=1}^{J} \sum_{k=1}^{Kj} P_i * \left(A^{D}_{kj}\right) * \left(A^{I}_{kj}\right) * S_{ijk} \qquad (8)$$

Where, K_j is the index set of sub-criteria for criterion j (J is the index set for criterion j).

In equation (8), P_j is the weight results of pairwise comparisons among criteria with respect to the goal which is taken from the cluster matrix results. A^{D}_{kj} is the results of pairwise comparisons among sub- criteria with respect to their control criteria, which is taken from the weighted supermatrices results. A^{D}_{kj} is the stabilized importance weight of the sub-criterion k of criterion j for the interdependency (I) relationships, which is taken from the limiting super matrix results. S_{ijk} is the results of pairwise comparisons among alternatives with respect to the sub-criteria. Finally, the severity index (SI) can be obtained from the following equation:

$$SI = \frac{D_i}{\sum D_i} \qquad (9)$$

The final decision is then made based on the SIs for the eight alternatives.

RESULTS

The aim of this study was to construct a quantitative index and a scoring model in order to predict the clinical severity and risk of the iron overload of βeta thalassemia, using results from studying the correlation between genotype, phenotype and patient history. The proposed model is considered to be the first scoring system for evaluating the risk of the iron overload of transfusion dependent β-thalassemia patients.

Results from the gene test

In the Hematology Unit of Mansoura University Children Hospital (MUCH). Gene testing was performed on blood samples from 150 different patients, looking for the three most common mutations of βeta thalassemia in Egypt. The three mutations, namely (IVS 1-1, IVS 1-6, and IVS I-110) were detected in 146 (97.3%) patients out of 150 **(Figure 4)**.

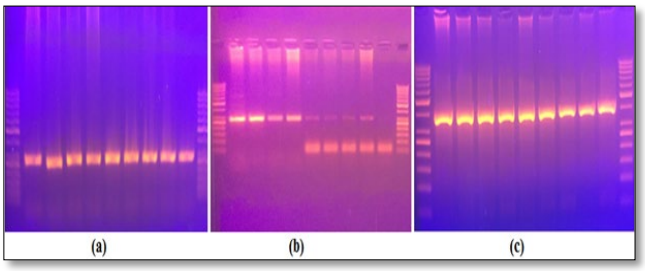

Figure 4. Cases from 1 to 9 for (a) IVS 1-1, (b) IVS 1-6, and (c) IVS 1-110.

Results from the DEMATEL approach

Applying the DEMATEL steps mentioned above via six experts, the total relation matrix *T* for criteria is obtained and shown in **Table 2**. The genotype modifiers have the highest value of *R*=2.194059629. On the other hand, phenotype factors have the highest value of *C*=1.964640372. The total sum of effects as received by each criterion lead to obtain the prominence and relevance results represented by (*r*+*c*) and (*r* - *c*), respectively as shown in **Table 3**.

Table 2. Total relation matrix **T** of criteria.

Total Relation Matrix	Phenotype Factors	Genotype Modifiers	History	R	C
Phenotype Factors	0.309698376	0.300696	0.603619	1.214013376	1.964640372
Genotype Modifiers	0.875359629	0.503016	0.815684	2.194059629	0.85243887
History	0.779582367	0.04872687	0.293643636	1.121952873	1.712946636
Average				4.530025878	4.530025878

Table 3. Prominence and relation results for each criterion.

	r+c	Prominence	*r - c*	Relation
Phenotype Factor	3.178653748	1	-0.750626996	Effect
Genotype Modifiers	3.046498499	2	1.341620759	Cause
History	2.834899509	3	-0.590993763	Effect

In **Table 3**, the phenotype factor has highest prominence value. The genotype modifiers have the second highest prominence value. The least prominent factor was the patient history. On the other hand, genotype modifiers were the only factor with positive relevance. This means that the genotype modifies are cause, while the phenotype and patient history are both effect.

Applying previous results from **Tables 2 and 3**, the Impact Relationship Map (IRM) and the causal diagram of criteria are presented in **Figures 5a and 5b**, respectively. The direction of the arrows shows the direction of influences. These two figures represent the most important effects in **Table 2**, after removing the "below average" effects. Based on these figures, the total (direct and indirect) influences matrix T_α can be represented in **Table 4**.

Table 4. The final total relation matrix T_α of criteria.

Total Relation Matrix (Tα)	Phenotype Factors	Genotype Modifiers	History
Phenotype Factors	0	0	0.603619
Genotype Modifiers	0.875359629	0.503016	0.815684
History	0.779582367	0.504872	0

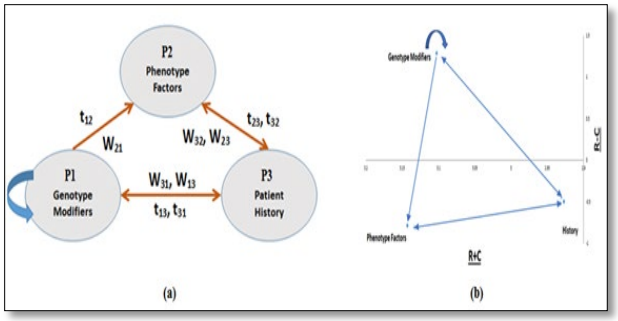

Figure 5. Graphical relations between criteria: (a) The Impact-Relations-Map, (b) Causal Diagram.

Results from the ANP approach

The ANP steps were implemented using the super decision software "Super Decision **2.3.0**". In this way, conditions for constructing structural relations among factors in the system are created in **Figure 6**.

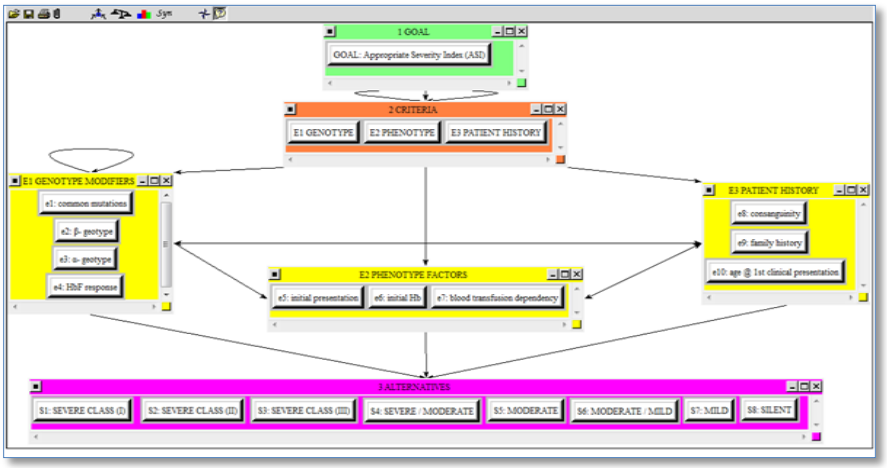

Figure 6. The super decision ANP model.

Applying Equations 6 and 7 to the final relationship matrix in **Table 4**, we obtained the unweighted supermatrix (W). Consequently, the relative weights for sub-criteria were calculated **(Table 5)**.

Table 5. The sub-criteria weights.

Criteria	Sub-Criteria	Limiting Weights
Genotype modifiers	e1: Common Mutation Type	0.01025
	e2: β-Genotype	0.127561
	e3: α-Genotype	0.014983
	e4: HbF Response	0.031046
Phenotype factors	e5: Initial Presentations	0.033695
	e6: Initial Hb	0.025003
	e7: Blood Transfusion Dependency	0.184986
Patient history	e8: Consanguinity	0.021611
	e9: Family History	0.018438
	e10: Age @ First Clinical Features	0.102502

Results from evaluation

According to Equation (8) and (9), the desirability index (Di) was obtained; and the final severity index (SI) was produced. As shown in **Table 6**, we divided the severities into the following ranges.

Table 6. Severity ranges.

Range	Severity
SI ≥ 0.3	SEVERE CLASS (I)
0.2 ≤ SI<0.3	SEVERE CLASS (II)
0.15 ≤ SI<0.2	SEVERE CLASS (III)
0.1 ≤ SI<0.15	SEVERE/MODERATE
0.07 ≤ SI<0.1	MODERATE
0.05 ≤ SI<0.07	MODERATE/MILD
0.03 ≤ SI<0.05	MILD
SI<0.03	SILENT

DISCUSSION

Due to the growing needs and expectations of all decision makers especially in the medical field, continuous improvement process is considered a necessary precondition to success in organizations. The keyword for the improvement process in health care institutions is quality of services given to patients. In general, optimal disease management is characterized by its three main pillars high speed, low cost and high accuracy. This optimization is highly dependent on clinician's gained knowledge and experience. Confirmation of disease existence and severity assessment will enhance patient care and limiting follow-up time. Also studying correlation among factors related to the disease and patient are considered one of the most essential directions in tailoring regimens and follow up. For these reasons the cycle of disease management using conventional methods calls for using high throughputs models.

Blood transfusion plays a pivotal role in the management of transfusion dependent thalassemia (TDT) patients; however, it causes significant iron overload. This iron overload puts the TDT patients at greater risk due to the iron-related complications that should be treated. The iron chelation is the primary treatment for transfusional iron overload in these patients. The iron chelation does not start in children before the age of 2 years old to avoid toxicity. The clinical indicators for the adjustment of the chelation therapy are: transfusion requirement, severity of iron overload and treatment goal [12]. Therefore, it is necessary to have a special quantitative index for β-thalassemia severity assessment and an estimator of the iron overload risk classes resulting from the blood transfusion. According to this index the therapeutic dose of iron chelation will be prescribed.

In most Egyptian children hospitals and clinics, β-thalassemia diagnosis and severity assessment are based on the apparent symptoms (phenotype factors). Clinically, most hemoglobin disorders are similar in symptoms leading to clinical misdiagnosis. In addition, β-thalassemia includes three main forms: Thalassemia Major (TM), Thalassemia Intermedia (TI) and Thalassemia minor or carrier. Thalassemia Major associated with severe microcytic anemia, mild jaundice and hepatosplenomegaly. Thalassemia Intermedia (TI) similar symptoms as in Thalassemia Major but with milder clinical findings [12]; Thalassemia minor or carrier with no or mild anemia [12]. So, to overcome this differentiating in diagnosis the genotype modifiers have been included in this work. Genotyping is the only confirmatory technique for the clinical diagnosis of β-thalassemia.

β-thalassemia is a single gene disorder, where the detection of one type of mutation in the HBB gene, the disease will appear. Each country has its specific common mutations, in Egypt there are 10 common mutations [12]. So, genotyping clarifies the gene mutations that are responsible for β-thalassemia disease. β-thalassemia alone constitute about 3-4% of the major hemoglobin disorders with an estimate of around 8,000-10,000 new births with the most severe form of diseases each year. According to these facts, the gene investigation of the three most common mutations in Egypt namely (IVS 1-1, IVS 1-6 and IVS I-110) was detected in 146 (79.3%) patients out of 150.

In order to obtain an index for β-thalassemia severity assessment and an estimator of the iron overload risk classes, an empirical study based on the concept of MCDM approaches have been demonstrated. These approaches deal with real world problems and mainly depend on experts. Severity assessment and estimated iron overload classes have been obtained by using the proposed hybrid MCDM framework: DEMATEL technique and ANP approach. This hybrid MCDM framework was proposed for solving the

problem of interrelations and uncertainty during judgment in order to obtain adequate final index. In the next sections the analytic results of applying the proposed model will be discussed:

As mentioned before DEMATEL was employed to determine the interdependencies among severity factors. Based on the analytic results in **Tables 2 and 3**, an impact relation maps (IRM) and the causal diagrams were constructed **Figure 7**. This figure was a useful visual model that aids clinicians in the decision-making process, because it gave an overview on the relationship between the different influence factors.

In **Table 2**, the genotype modifiers have the most effect given with the highest value of R and assigned in cause group with positive value of (R-C). On the other hand, phenotype factors have the most effect received with the highest value of C and assigned in effect group with negative value of (R-C). Also, phenotype factors were the highest value of (R+C). This indicates that the phenotype factors are the most important severity assessment criteria and also has the most important interactions with the other factor. Furthermore, the genotype modifiers with the highest value of (R-C) mostly affect other criteria. In fact, the genotype modifiers have a critical role in β-Thalassemia severity assessment. This means that desirable goals through other criteria will not necessarily be achieved without due consideration of genotype modifiers factors. All these results are according to the expert's practice. Reasons to obtain these results are: From the clinical perspective, the phenotype factors criterion ranked first in importance. As a routine work, the consulted expert's evaluation for the severity is usually based on the phenotype factors and the patient history criteria only, without considering the genotype modifiers that we added, although the genotype is the main cause for β-Thalassemia existence due to its autosomal features.

According to the ANP result, which is presented in Table 4, the most two important sub-criteria for severity assessment are e7 and e2. Blood Transfusion Dependency related to "Phenotype" criteria with limiting weight=0.184986 and β-Genotype related to "Genotype" criteria with limiting weight=0.127561, respectively. These results confirmed the DEMATEL results that the phenotype factors and genotype modifiers are the most important criteria in severity assessment.

Therefore, based on experts' viewpoints these two sub-criteria have critical roles to play for β-thalassemia severity assessment. Hence, it is recommended that clinicians should pay more attention to these factors due to their direct effects on the assessment of severity. Taking into consideration that the identification of genetic modifiers of disease severity promises to improve patient management and treatment. **Figure 7** gives an overview on the proposed model.

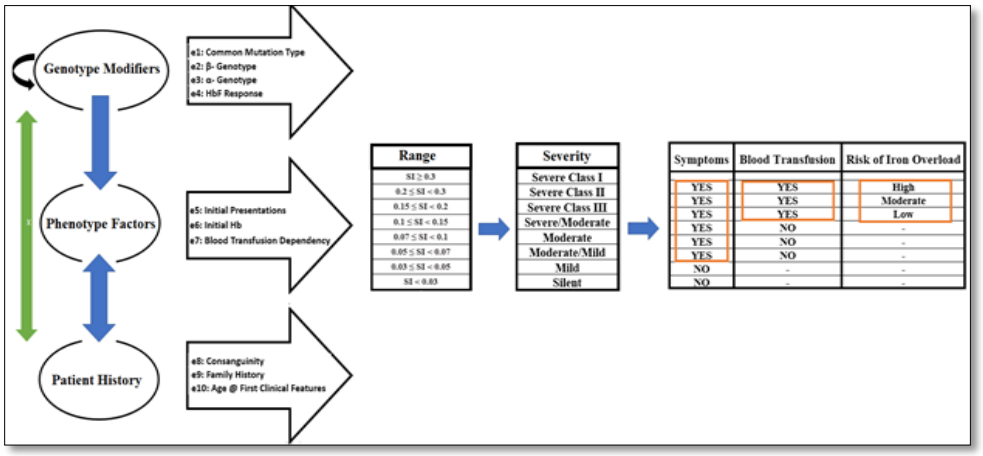

Figure 7. Flow diagram of the severity scoring system.

CONCLUSION

Benefits obtained from the concluded results when using DEMATEL-ANP approach are: First, the causal relationships between criteria gave clinicians effective information on how the risk factors of iron overload should be treated. Second, the weight of severity assessment factors which presents the priority of factors contributes to obtain the severity index. On the other hand, the severity assessment basically depends on the genotype that influences the other contributing factors. From the clinical perspective, the phenotype factors criterion ranked first in importance. As a routine work, the consulted expert's evaluation for the severity is usually based on the phenotype factors and the patient history criteria only, without considering the genotype modifiers that we added, although the genotype is the main cause for βeta Thalassemia existence due to its autosomal features. The proposed model offers the following advantages:

- The model predicts the disease's severity in an early stage of life, improving our knowledge about the disease's

pathophysiology, phenotype diversity and genotypic heterogeneity. This in turn helps at the early intervention and limiting the time of the follow-up.

- Evaluating the severity of the disease supports clinicians to decide the most suitable regimens for the patients, for example, for the transfusion dependent beta thalassemia patients, whether to start a transfusion program or not for the supportive treatment, and also whether to perform a bone marrow transplantation or not.

REFERENCES

1. Miller JB, Shan G, Lombardo J, Jimenez-Maggoria G (2018) Biomedical informatics applications for precision management of neurodegenerative diseases. Alzheimers Dement (NY) 4: 357-365.
2. Ohno-Machado L (2018) Clinical research informatics: A growing sub specialization of biomedical informatics. J Am Med Informatics Assoc 25: 223.
3. Ohno-Machado L (2018) Clinical decision support: Informatics interventions for better patient care. J Am Med Informatics Assoc 25: 457.
4. Chen J, Qian F, Yan W, Shen B (2013) Translational biomedical informatics in the cloud: Present and future. BioMed Res Int 2013: 1-8.
5. Altman RB (2012) Translational bioinformatics: Linking the molecular world to the clinical world. Clin Pharmacol Ther 91: 994-1000.
6. Shameer K, Badgeley MA, Miotto R, Glicksberg BS, Morgan JW, et al. (2017) Translational bioinformatics in the era of real-time biomedical, health care and wellness data streams. Brief Bioinform 18: 105-124.
7. Hossain MS, Raheem E, Sultana TA, Ferdous S, Nahar N, et al. (2017) Thalassemias in South Asia: Clinical lessons learnt from Bangladesh. Orphanet J Rare Dis 12: 93.
8. Sura D, Chakravorty R (2016) Prevalence of hemoglobinopathies and thalassemia carriers in women of reproductive age group especially the prospective mothers - A single center study at West Bengal. J Hematol 5: 99-102.
9. Shawky RM, Kamal TM (2012) Thalassemia intermedia: An overview. Egypt J Med Hum Genet 13: 245-255.
10. Esfandiari N, Babavalian MR, Moghadam AME, Tabar VK (2014) Knowledge discovery in medicine: Current issue and future trend. Expert Systems with Applications 41: 4434-4463.
11. Uzokaa FME, Obotb O, Osujid J (2011) Clinical decision support system (DSS) in the diagnosis of malaria: A case comparison of two soft computing methodologies. Expert Syst Appl 38: 1537-1553.
12. Capellini MD, Taher A (2014) Guidelines for the management for transfusion of thalassemia dependent (TDT). 3rd Edn. Thalassemia International Federation (TIF).
13. Ntaios G (2007) Discrimination indices as screening tests for β-thalassemic trait. Ann Hematol 86: 487-491.
14. Amendolia SR (2003) A comparative study of K-nearest neighbor, support vector machine and multi-layer perceptron for thalassemia screening. Chemometrics and Intelligent Laboratory Systems 69: 13-20.
15. Wongseree W (2007) Thalassemia classification by neural networks and genetic programming. Information Sciences 177: 771-786.
16. Dogan S, Turkoglu I (2008) Iron-deficiency anemia detection from hematology parameters by using decision trees. Int J Sci Technol 3: 85-92.
17. Piroonratana T (2009) Classification of hemoglobin typing chromatograms by neural networks and decision trees for thalassaemia screening. Chemometrics and Intelligent Laboratory Systems 99: 101-110.
18. Damrongrit S (2012) Classification of complete blood count and hemoglobin typing data by a C4.5 decision tree, a naive Bayes classifier and a multilayer perceptron for thalassemia screening. Biomed Signal Process Control 7: 202-212.
19. Paokanta P (2012) Parameter estimation of binomial logistic regression based on classical (maximum likelihood) and Bayesian (MCMC) approach for screening β-thalassemia. International Journal of Intelligent Information Processing 3: 1
20. Yassin MM (2013) Genotype-phenotype characteristics of β thalassemia children in the Gaza Strip, Palestine. J Genet Disord Genet Rep 2: 1-6.
21. Cappellini MD, Porter JB, Musallam KM, Kattamis A, Viprakasit V, et al. (2016) Development of a new disease severity scoring system for patients with non-transfusion-dependent thalassemia. Eur J Intern Med 28: 91-96.
22. Doumpos M, Zopounidis C (2011) Preference disaggregation and statistical learning for multi-criteria decision support. Eur J Operational Res 209: 203-214.
23. Lin RH, Chuang CL (2010) A hybrid diagnosis model for determining the types of the liver disease. Comput Biol Med 40: 665-670.
24. Uzokaa FME, Obotb O, Barkerc K, Osujid J (2011) An experimental comparison of fuzzy logic and analytic

hierarchy process for medical decision support system. Comput Methods Programs Biomed 10: 10-27.

25. Suner A, Çelikoğlu CC, Dicle O, Sökmen S (2012) Sequential decision tree using the analytic hierarchy process for decision support in rectal cancer. Artif Intell Med 56: 59-68.
26. Triantaphyllou E, Shu B, Sanchez SN, Ray T (1998) Multi-criteria decision making: An operations research approach. Encyclopedia of Electrical and Electronics Engineering 15: 175-186.
27. Zopounidis C, Doumpos M (2002) Multicriteria classification and sorting methods: A literature review. Eur J Operational Res 138: 229-246.
28. Aruldoss M, Lakshmi TM, Venkatesan VP (2013) A survey on multi criteria decision making methods and its applications. Am J Inform Syst 1: 31-43.
29. Lertprapai S (2013) Review: Multiple criteria decision-making methods with applications. International Mathematical Forum 8: 347-355.
30. Gabus A, Fontela E (1972) World problems, an invitation to further thought within the framework of DEMATEL. Battelle Geneva Research Center: Geneva, Switzerland.
31. Chen WC, Chang HP, Lin KM, Kan NH (2015) An efficient model for NPD performance evaluation using DEMATEL and fuzzy ANP - Applied to the TFT-LCD touch panel industry in Taiwan. Energies 8: 11973-11981.
32. Yang CL, Yuan BJC, Huang CY (2015) Key determinant derivations for information technology disaster recovery site selection by the multi-criterion decision making method. Sustainability 7: 6149-6188.
33. Si SL, You XY, Liu HC, Zhang P (2018) DEMATEL technique: A systematic review of the state-of-the-art literature on methodologies and applications. Mathematical Problems in Engineering 2018: 1-33
34. Yang JL, Tzeng GH (2011) An integrated MCDM technique combined with DEMATEL for a novel cluster-weighted with ANP method. Expert Syst Appl 38: 1417-1424.
35. Saaty TL (2005) Theory and applications of the analytic network process: Decision making with benefits, opportunities, costs and risks. RWS Publications: Pittsburgh, PA, USA.
36. Saaty TL (2016) Decision making in complex environments: The analytic network process (ANP) for dependence and feedback. McGraw-Hill Company, New York.
37. Yang YPO, Shieh HM, Leu JD, Tzeng GH (2008) A novel hybrid MCDM model combined with DEMATEL and ANP with applications. International Journal of Operational Research 5: 160-168.
38. Zheng X, Xu F, Feng L (2017) Analysis of driving factors for extended producer responsibility by using interpretative structure modeling (ISM) and analytic network process (ANP). Sustainability 9: 540.
39. Chemweno P, Pintelon L, van Horenbeek A, Muchiri P (2015) Development of a risk assessment selection methodology for asset maintenance decision making: An analytic network process (ANP) approach. Int J Prod Econ 170: 663-676.
40. Fazli S, Mavi RK, Vosooghidizaji M (2015) Crude oil supply chain risk management with DEMATEL-ANP. Oper Res 15: 453-480.
41. Chang CW, Wu CR, Lin CT, Lin HL (2007) Evaluating digital video recorder systems using analytic hierarchy and analytic network processes. Information Sciences 177: 3383-3396.
42. Cabała P (2010) Using the analytic hierarchy process in evaluating decision alternatives. Operations Research and Decisions, no. 1.
43. Saaty TL (1996) Decision making with dependence and feedback: The analytic network process. RWS Publications, Pittsburgh.
44. Shieh JI, Wu HH, Huang KK (2010) A DEMATEL method in identifying key success factors of hospital service quality. Knowledge-Based Systems 23: 277-282.
45. Sripichai O, Makarasara W, Munkongdee T (2008) A scoring system for the classification of beta-thalassemia/Hb E disease severity. Am J Hematol 83: 482-484.
46. Danjou F (2014) A genetic score for the prediction of beta thalassemia severity. Haematologica.
47. Temtamy S, Aglan M (2012) Consanguinity and genetic disorders in Egypt. Middle East J Med Genet 1: 12-17.
48. Thein SL (2005) Pathophisiology of β-thalassemia – A guide to molecular therapies. Am Soc Hematol 2005: 31-37.
49. Meade LM, Sarkis J (1999) Analyzing organizational project alternatives for agile manufacturing processes: An analytic network approach. Int J Prod Res 37: 241-261.
50. Meade LM, Sarkis J (1999) Analyzing organizational project alternatives for agile manufacturing processes: An analytical network approach. Int J Prod Res 37: 241-261.

51. Lee H, Lee H, Park Y (2009) Selection of technology acquisition mode using the analytic network process. Mathematical and Computer Modeling 49: 1274-1282.

52. Sumrit D, Anuntavoranich P (2013) An analytic network process modeling to assess technological innovation capabilities: Case study for Thai automotive parts firms. Am Trans Eng Appl Sci, pp: 189-211.