

Forecasting mortality patterns of thalassaemia major patients in Iraq by using VAR model and reasons for this mortality

Rana Sabeeh Abood Alsudani^{1,2}, Jicheng Liu¹, Zahrah Ismael Salman²

1- School of Mathematics and Statistics Huazhong University of Science and Technology
Wuhan city, Hubei province, China

2- School of Mathematics
College of Basic education Maysan University-Republic of Iraq
rana19792015@hotmail.com

ABSTRACT

The vector autoregression model (VAR) is a natural extension of the univariate autoregressive model dynamic multivariable time series. It is one of the most successful, flexible, and easy to use models for the analysis of multivariable time series. The VAR model has proved to be particularly useful describing the dynamic behaviour of economic and financial time series and forecasting. Often it provides superior forecasts to those of time-series models and univariate and detailed forecasts, based on the theory of simultaneous equation models. Expectations of VAR models are very flexible because they can be conditioned on possible paths for the future in the form of specific variables. In addition to describing the data and forecasting, the VAR model is used to deduce structural and policy analysis. This study used the VAR model for forecasting the number of deaths in patients with thalassaemia in Maysan province in southern Iraq, and also addressed the causes of these deaths. There was a strong relationship between mortality in thalassaemia patients and an increase in the proportion of iron and the highest number of deaths was recorded for patients who had a very high proportion of iron. It was 'the most important cause of mortality (Cardiac disease, infections, the liver, the spleen).

Keywords

Forecasting; mortality for thalassaemia; model (VAR).

1. INTRODUCTION

About 60,000 individuals are born annually with thalassaemia.²¹ It is a severe inherited anemia arising from the failure of hemoglobin synthesis.¹⁶ Symptoms of thalassaemia do not appear at birth immediately, because the appearance of a child with thalassaemia is not different to that of other newborns, but over time, especially after the first six months, the child begins to suffer from anemia, and prominence in the face bones appears and the child becomes susceptible to infections. The blood hemoglobin level decreases and build, so the child will need periodic blood transfusion operations, and the long-term survival of such patients depends on regular blood transfusion.¹⁴ This treatment can result in some major complications such as iron overload,¹⁷ cardiac disorders,¹ and endocrine disorders.²⁰ Iron overload is the primary cause of mortality and morbidity in thalassaemia major patients despite advances in chelation therapy.¹¹ Accumulating levels of iron in the body generated from hundreds of blood transfusions, together with increased absorption of iron from the diet, cause critical organs and glands in patients with thalassaemia to experience serious iron-induced toxicity and early death.²² Several publications provide evidence that the heart is unquestionably the most critical organ affected by the iron, jeopardizing the survival of thalassaemia patients; 36% of patients between the ages of 15 and 18 showed detectable cardiac iron toxicity.²² The life of patients with thalassaemia has improved both in terms of duration and quality in industrialized countries, but complications are still common and include heart disease, growth retardation, pallor, jaundice, poor musculature, hepatosplenomegaly, leg ulcers, and the development of masses from extramedullary haematopoiesis, and skeletal changes (due to bone marrow expansion) are found in untreated or poorly transfused individuals with thalassaemia major.⁷ This study aimed to develop a standard model for forecasting the number of deaths in patients with thalassaemia from Maysan province in southern Iraq in the coming years (2016-2020). The study concluded that the vector autoregression (VAR) model is one of the most successful, flexible, and easy to use models for the analysis of multivariate time series. It is a natural extension of the univariate autoregressive model for dynamic multivariate time series. The VAR model has proven to be especially useful for describing the dynamic behaviour of economic and financial time series and for forecasting. It often provides superior forecasts to those from univariate time series models and elaborate theory-based simultaneous equation models. Forecasts of VAR models are quite flexible because they can be made conditional on the potential future paths of specified variables in the model. Vector autoregressive (VAR) models have a long tradition as tools for multiple time series analysis.¹⁵ Being linear models, they are relatively easy to work in both theory and practice. Although the related computations are relatively straightforward, they are sufficiently involved to make applied work cumbersome before powerful computers were in widespread use. VAR models became popular for economic analysis when Sims (1980) advocated them as alternatives to simultaneous equation models. The later models have used extensively since the 1950s. The availability of longer and more frequent observed time series emphasized the need for models that focused on the dynamic structure of the variables, however. Sims also criticized the exogeneity assumptions



for some of the variables in simultaneous equation models as ad hoc and often not backed by fully developed theories. In contrast, in VAR models often all observed variables are treated as a priori endogenous. Statistical procedures rather than subject matter theory are used for imposing restrictions. Our study also included a precise scanning of our data for most complications caused by thalassemia by selecting the most causative one for mortality.

2- Methodology: VAR model :(Vectorial Autoregressive)

Proposed as a model by Sims in 1981,¹⁹ vector autoregression (VAR) is an econometric model used to capture the linear interdependencies among multiple time series. VAR models generalize the univariate autoregressive model (AR model) by allowing for more than one evolving variable. All variables in a VAR enter the model in the same way: each variable has an equation explaining its evolution based on its own lags and the lags of the other model variables. VAR modelling does not require as much knowledge about the forces influencing a variable, as do structural models with simultaneous equations: the only prior knowledge required is a list of variables that can be hypothesized to affect each other.⁸ General models can be written as:

$$\phi(B)Y_t = \varepsilon_t$$

Where:

Y_t : A random context dimensioned n stable

$\phi(B)$: Polynomial matrix, grade P, slow time factor B

$$\phi(B) = \phi_0 - B\phi_1 - B^2\phi_2 - \dots - B^p\phi_p$$

ϕ_0 : Matrix monocrystalline grade n

ε_t : White noise

Var model can also be written as follows:¹⁸

$$Y_{1t} = \phi_{11}^{(1)} y_{1,t-1} + \dots + \phi_{11}^{(p)} y_{1,t-p} + \dots + \phi_{1n}^{(1)} y_{n,t-1} + \dots + \phi_{1n}^{(p)} y_{n,t-p} + \varepsilon_{1,t}$$

$$Y_{nt} = \phi_{n1}^{(1)} y_{1,t-1} + \dots + \phi_{n1}^{(p)} y_{1,t-p} + \dots + \phi_{nn}^{(1)} y_{n,t-1} + \dots + \phi_{nn}^{(p)} y_{n,t-p} + \varepsilon_{n,t}$$

2-1 constructing the model (VAR):

- 1- Using a stationary time series, this does not contain the root of unity.
- 2- Specify the number of periods of slow time.
- 3- Study the causal relationship between the variables. Equations.¹⁸

2-1-1 – Stationarity:¹⁰

With autocovariance functions, we can define the covariance stationarity, or weak stationarity. In the literature, usually stationarity means weak stationarity, unless otherwise specified. The time series $(x_t, t \in Z)$ (where z is the integer set) is said to be stationary if:

- a. $E(X_t^2) < \infty \forall t \in Z$.
- b. $EX_t = \mu \forall t \in Z$.
- c. $\gamma_X(s, t) = \gamma_X(s + h, t + h) \forall s, t, h \in Z$

In other words, a stationary time series $\{X_t\}$ must have three features: finite variation, constant first moment, and that the second moment $\gamma_X(s, t)$ only depends on $(t - s)$ and does not depend on s or t. In light of the last point, we can rewrite the auto covariance function of a stationary process as

$$\gamma_X(h) = \text{Cov}(x_t, x_{t+h}) \text{ for } t, h \in Z.$$

Also, when x_t is stationary, we must have

$$\gamma_X(h) = \gamma_X(-h).$$

When $h=0$, $\gamma_X(0) = \text{Cov}(x_t, x_t)$ is the variance of x_t so the autocorrelation function for a stationary time series $\{x_t\}$ is defined as

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}$$

Dickey –Fuller test



In statistics, the Dickey–Fuller test, tests the null hypothesis of whether a unit root is present in an autoregressive model. The alternative hypothesis is different depending on which version of the test is used, but is usually stationary or trend-stationary. It is named after the statisticians David Dickey and Wayne Fuller, who developed the test in 1979.⁵ A simple AR (1) model is

$$y_t = py_{t-1} + u_t$$

Where y_t is the variable of interest, t is the time index p is a coefficient, and u_t is the error term. A unit root is present if $p=1$. The model would be non-stationary in this case. The regression model can be written as

$$\nabla y_t = (p - 1)y_{t-1} + u_t = \delta y_{t-1} + u_t$$

Where ∇ is the first difference operator. This model can be estimated and testing for a unit root is equivalent to testing $\delta = 0$ (where $\delta \equiv p - 1$). Since the test is done over the residual term rather than raw data, it is not possible to use standard t-distribution to provide critical values.

There are three main versions of the test:

- 1. Test for a unit root:

$$\nabla y_t = \delta y_{t-1} + u_t$$

- 2. Test for a unit root with drift

$$\nabla y_t = a_0 + \delta y_{t-1} + u_t$$

- 3. Test for a unit root with drift and deterministic time trend

$$\nabla y_t = a_0 + a_1t + \delta y_{t-1} + u_t$$

Each version of the test has its own, critical value which depends on the size of the sample. In each case, the null hypothesis is that there is a unit root, $\delta = 0$. The tests have low statistical in that they often cannot distinguish between true unit-root processes ($\delta = 0$) and near unit-root processes (δ is close to zero). This is called the “near observation equivalence” problem. The intuition behind the test is as follows. If the series y is stationary (or trend stationary), then it has a tendency to return to a constant (or deterministically trending) mean. Therefore large values will tend to be followed by smaller values (negative changes), and small values by larger values (positive changes). Accordingly, the level of the series will be a significant predictor of the next period's change, and will have a negative coefficient. If, on the other hand, the series is integrated, then positive changes and negative changes will occur with probabilities that do not depend on the current level of the series, in a random walk.⁶

2-1-2- determine the time periods a slowdown:

When Sims gave his model did not give any limitation with respect to the length of the time period of a slowdown that can be applied. There are several quantitative criteria can be used to determine the time periods in a slowdown:

1. Final Predictor Error Criterion (FPE) :³

Akaike's Final Prediction Error (FPE) criterion provides a measure of model quality by simulating the situation where the model is tested on a different data set. After computing several different models, you can compare them using this criterion. According to Akaike's theory, the most accurate model has the smallest FPE. If you use the same data set for both model estimation and validation, the fit always improves as you increase the model order and, therefore, the flexibility of the model structure. Akaike's Final Prediction Error (FPE) is defined by the following equation:

$$FPE = \det \left(\frac{1}{n} \sum_1^N e(t, \theta_N) (e(t, \theta_N))^T \right) \left(\frac{1+d/N}{1-d/N} \right)$$

where:

- N is the number of values in the estimation data set
- e(t) is a ny-by -1 vector of prediction errors
- θ_N represents the estimated parameters

2. AKAIKE Information Criterion (AIC)³

This computes the following relationship:

$$AIC(P) = \text{Log} \left(\det \Omega(p) + 2 \frac{n^2 p}{N} \right)$$

Where:



- Ω : Matrix variations and changes of the estimated remainders
- n : The number of internal variables
- N : Total views

After that we choose P_0 that achieve the equation

$$AIC(P_0) = \underset{p=1}{\text{Min}} AIC(P)^k$$

3. Bayesian Information Criterion (BIC):¹²

Computes the following relationship:

$$BIC(P) = \text{Log}(\det\Omega(p)) + \left(\frac{n^2 \cdot p \cdot \log N}{N} \right)$$

Of the calculate the number of extended slowdown

$$BIC(P_0) = \underset{p=1}{\text{Min}} BIC(P)^k$$

4. Hannan & Quinn Information Criterion (HQIC)¹²

It is calculated from the following relationship:

$$HQIC(P) = \text{Log}(\det\Omega(p)) + \left(2n^2 P_c \frac{\log \cdot \log N}{N} \right)$$

Where:

C : strength indicator benchmark and we consider it equal to 2 in the practical application

Of the calculate the number of extended slowdown

$$HQIC(P_0) = \underset{p=1}{\text{Min}} HQIC(P)^k$$

Note: we can apply the different outcomes of these criteria in practice in this case we select the time slowdown for which we got the largest number of criteria

2-1-3 Causality test: ¹²

- Estimates the following equation by using the least squares method:

$$Y_t = \phi_1(B) \cdot Y_t + \phi_2(B) \cdot X_t + \varepsilon_t$$

Where

$$\phi_1(B) = \sum_{i=1}^p \phi_{1i} \cdot B^i \quad \text{and} \quad \phi_2(B) = \sum_{i=1}^q \phi_{2i} \cdot B^i$$

We calculate the total deviations of actual values about estimated symbolized her SCR1.

- Estimate the following equation

$$Y_t = \phi_1(B) \cdot Y_t + \varepsilon_t$$

We calculate the total deviations of actual values about estimated symbolized her SCR2.

- calculate statistically the test F_c of the relationship

$$F_c = \frac{(SCR2 - SCR1)/P}{SCR1/(M - N)}$$

Where

$M = T - \text{Max}(p, q)$ and $N = p + q + 2$

T : number of views

P : The number of time slowdowns of internal variables



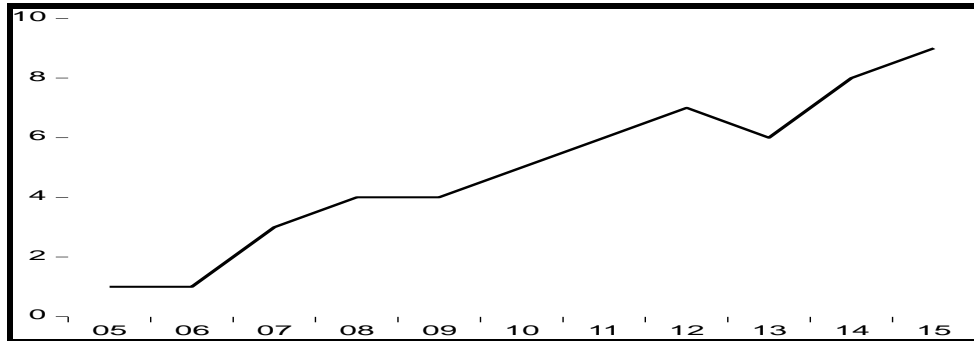
q: The number of time slowdowns of external variables.

3 - Data analysis:

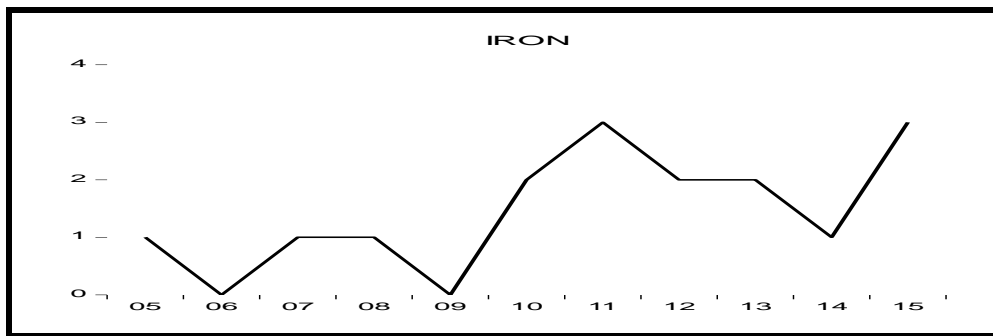
We obtained time series stretching from 2005 to 2015, showing the number of deaths in patients with thalassemia.

3-1 Showing time series:

Graph (1) shows clearly the instability of the series, but does not indicate whether the instability is due to the presence of unit root or not? Therefore we must test the unit root



Graph (1) Time series of mortality for patients with thalassemia



Graph (2) Time series for the preparation of cases of death due to high iron for the years 2005-2015

3-2 Stationarity test: ¹⁰

Augmented Dickey & Fuller (ADF)

Because the instability in time series is often due to the existence of unit root, the Dickey–Fuller test has been proposed to detect whether unit root is present or not. Assuming that we have random context in x_t , we calculate the regression as follows:

$$\Delta x_t = \varphi x_{t-1} + \sum_{i=1}^p a_i (x_{t-i} - x_{t-i-1}) + \varepsilon_t$$

$H_0: \varphi = 0$ Acceptance of this hypothesis implies acceptance that there is unit root instability

$H_0: \varphi < 1$ Accepting this implies acceptance of the hypothesis of stability and the lack of unit root.⁴

We have the test result shows us that the time series is not stable and contains the root of the unit we convert it to the first application of a stable candidate variances series $\Delta = (1 - B)$ and After which test the resulting string.⁹

Table (1) Dickey-Fuller test enlarged

Augmented Dickey –Fuller test statistic		t-statistic	Prob.
		-0.438295	0.8662
Test critical values	1% level	-4.297073	
	5% level	-3.212696	
	10% level	-2.747676	



We find from Table (1) that the absolute value of (DAF) is calculated (0.838611) smaller than the absolute values at different levels of significance. If we accept the null hypothesis we accept the existence of unit root in the time series. Applying the unit root test again to the first series of difference we find that the results confirm the time series in table (2):

Table (2) Expanded Dickey Fuller test after the first time series difference

Augmented Dickey –Fuller test statistic		t-statistic	Prob.
		-3.785640	0.0273
Test critical values	1% level	-4.582648	
	5% level	-3.320969	
	10% level	-2.801384	

To get rid of the unit root we apply again the first candidate difference to the first series of difference and we can obtain a stable series as shown in Table (3)

Table (3) Expanded Dickey Fuller test after the second difference of the time series

Augmented Dickey –Fuller test statistic		t-statistic	Prob.
		-3.845741	0.0024
Test critical values	1% level	-2.937216	
	5% level	-2.006292	
	10% level	-1.598068	

Table (4) Dickey-Fuller test enlarged Iron series

Augmented Dickey –Fuller test statistic		t-statistic	Prob.
		-1.784308	0.3657
Test critical values	1% level	-4.297073	
	5% level	-3.212696	
	10% level	-2.747676	

We find from Table (4) that the absolute value of (DAF) is calculated (1.784308) smaller than the absolute values at different levels of significance. If we accept the null hypothesis we accept the existence of unit root in the time series. Applying the unit root test again to the first series of difference we find that the results confirm the series in Table (5):

Table (5) Dickey-Fuller test enlarged Iron series, after taking the first differences of the series

Augmented Dickey –Fuller test statistic		t-statistic	Prob.
		-3.632122	0.0296
Test critical values	1% level	-4.420595	
	5% level	-3.259808	
	10% level	-2.771129	



To get rid of the unit root we apply again the first candidate difference to the first series of difference and we can obtain a stable series as shown in Table (6).

Table (6) Dickey-Fuller test enlarged Iron series, after taking the second differences of the series

Augmented Dickey –Fuller test statistic		t-statistic	Prob.
		-3.132381	0.0072
Test critical values	1% level	-2.937216	
	5% level	-2.006292	
	10% level	-1.598068	

3-3 Specify the amount of extended slowdown time

To specify the amount of extended slowdown time we use the criteria that we have advanced in advance and that appear in the following Table (7).

Table (7) Criteria determine the number of periods of slowdown of time

VAR Lag Order Selection Criteria						
Endogenous variables: DTHALASSEMIA DIRON						
Exogenous variables: C						
Sample: 2005 2015						
Lag	LogL	LR	FPE	AIC	SC	HQ
0	-24.44796	NA*	2.557079	6.611991	6.631851	6.478041
1	-18.62741	7.275690	1.747198	6.156853	6.216434	5.755002
2	-10.99232	5.726322	1.005012*	5.248079*	5.347381*	4.578327*
* indicates lag order selected by the criterion						
LR: sequential modified LR test statistic (each test at 5% level)						
FPE: Final prediction error						
AIC: Akaike information criterion						
SC: Schwarz information criterion						
HQ: Hannan-Quinn information criterion						

We find from Table 4 that the three criteria (FPE), (AIC),(SC) and (HQ) refers to taking 2 lag

3-4 Causality test

Table (8) shows the iron variable causes deaths for patients with thalassemia with 2 time lags at a level of significance of 5%.

Table (8) Granger causality test



Pairwise Granger causality tests			
Sample:2005-2015			
Lags :2			
Null Hypothesis	obs	F –statistic	Prob
Diron does not Granger cause Dthalassemia	8	4.31177	0.1311
Dthalassemia does not Granger cause Diron		0.23802	0.8018

Table 8 shows that the variable iron causes a thalassemia mortality variable at the level of 5%.

3-5 Estimates VAR model:

Given the causality test results in Table (8), and values of slowdown extended standards in Table (7), in order to reconcile these we choose 2 lags when estimating the model (VAR), see Table (9)

Table (9) The transactions estimates of the model

Vector Autoregression Estimates		
Sample (adjusted): 2008 2015		
Included observations: 8 after adjustments		
	Dthalassemia	D iron
DTHALASSEMIA(-1)	-0.395164	0.277190
	(0.33018)	(0.76185)
	[-1.19681]	[0.36384]
Dthalassemia(-2)	0.926277	0.049635
	(0.24567)	(0.56685)
	[3.77041]	[0.08756]
Diron (-1)	0.784854	0.018613
	(0.33239)	(0.76695)
	[2.36124]	[0.02427]
Diron (-2)	-0.421989	-0.415146
	(0.27399)	(0.63219)
	[-1.54017]	[-0.65668]
C	3.012226	0.553650
	(0.58070)	(1.33988)
	[5.18728]	[0.41321]
R –squared	0.958945	0.275973
Adj .R-squared	0.904205	-0.689396
Sum sq.resids	0.815967	4.344161



S .E.equation	0.521526	1203351
F –statistic	17.51820	0.285873
Log likelihood	-2.220217	-8.909072
Akaike AIC	1.805054	3.477268
Schwarz SC	1.854705	3.526919
Mean dependent	5.375000	1.500000
S.D. dependent	1.685018	0.925820
Determinant resid covariance (dof adj)		0.380596
Determinant resid covariance		0.053521
Log likelihood		-10.99232
Akaike information criterion		5.248079
Schwarz criterion		5.347381

VAR (2) model is estimated:

$$DTHALASSEMIA = - 0.395164233577 * DTHALASSEMIA (-1) + 0.926277372263 * DTHALASSEMIA (-2) + 0.784854014599 * DIRON (-1) - 0.421989051095 * DIRON (-2) + 3.01222627737$$

$$DIRON = 0.277189781022 * DTHALASSEMIA (-1) + 0.0496350364964 * DTHALASSEMIA (-2) + 0.0186131386861 * DIRON (-1) - 0.415145985401 * DIRON (-2) + 0.553649635036$$

Can be written by Layout the matrix:

$$\begin{bmatrix} D(THALASSEMIA)_t \\ D(IRON)_t \end{bmatrix} = \begin{bmatrix} -0.3951 & 0.7848 \\ 0.2771 & 0.0186 \end{bmatrix} \cdot \begin{bmatrix} D(THALASSEMIA)_{t-1} \\ D(IRON)_{t-1} \end{bmatrix} + \begin{bmatrix} 0.9262 & -0.4219 \\ 0.0496 & -0.4151 \end{bmatrix} \cdot \begin{bmatrix} D(THALASSEMIA)_{t-2} \\ D(IRON)_{t-2} \end{bmatrix}$$

3-6 The residual tests:

In order to validate the estimated model we must make sure that it undergoes the residuum normal distribution and that it is linked to itself

1. The probability distribution of the residual: use the test Jarque –Bera²

Table (10) Normal distribution test of residuals

VAR Residual Normality Tests			
Orthogonalization : cholesky(lutkepohl)			
Null hypothesis: residuals are multivariate normal			
Sample: 2005-2015			
Included observations:8			
Component	Jarqur- Bera	df	Prob.
1	0.431051	2	0.8061
2	0.834086	2	0.6590
Joint	1.265137	4	0.8673



The test indicates that the null hypothesis should not be rejected for each of the residuals for the first and second equation and that. Not to reject hypothesis of normal distribution of residuals at the 5% significance level should not be rejected.

2. Autocorrelation of residuals test: use the Ljung-Box test¹³

Table (11) Autocorrelation of residuals

VAR Residual portmanteau Test for Autocorrelations					
Null Hypothesis: no residual autocorrelations up to lag h					
Sample: 2005-2015					
Included observations :8					
Lag	Q-stat	Prob	Adj Q-stat	Prob	Df
1	6.237209	NA	7.128239	NA	NA
2	11.25064	NA	13.81281	NA	NA
3	14.76834	0.0052	19.44113	0.0006	4
4	16.24727	0.0390	22.39899	0.0042	8
5	17.89003	0.1191	26.77968	0.0083	12
6	18.78529	0.2800	30.36072	0.0162	16

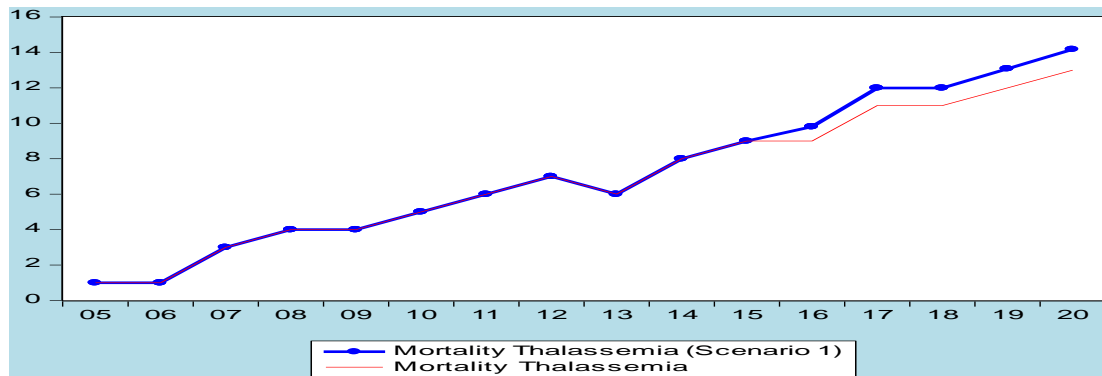
The autocorrelation test indicates that the null hypothesis should be rejected, not which means a lack of self-link when the level of significance is 5%.

4- Forecasting:

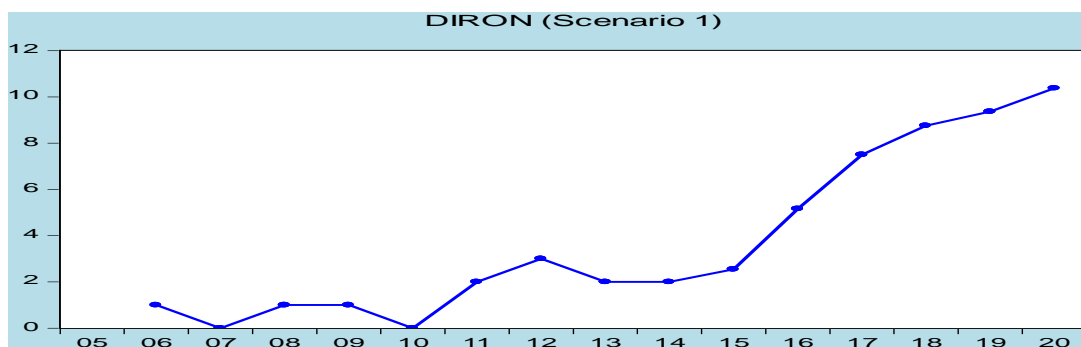
Table (12): Forecasting thalassemia mortality and iron , for the years 2016-2020

Year	¹ L Mortality thalassemia	² U mortality thalassemia	L iron	U iron
2016	11.84799	12.79583	5.157853	5.261010
2017	14.84125	16.02855	7.502795	7.652851
2018	16.02047	17.30211	8.758657	8.933830
2019	16.87972	18.23010	9.367777	9.555132
2020	17.85826	19.28692	10.37876	10.58634

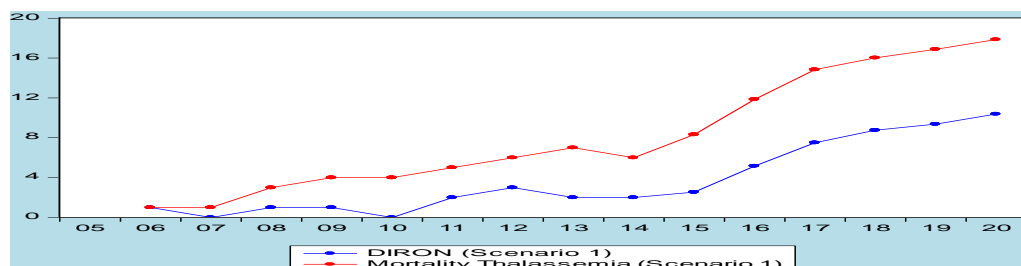
- (1) Minimum of forecasting
- (2) Upper limit of forecasting



Graph (3) Graph of the values of forecasting of thalassemia mortality



Graph (4) Graph of the values of forecasting of iron



Graph (5) Graph of the values of forecasting of thalassemia mortality and iron

Result:

Through a graphs and tables we can forecast the numbers of mortalities for the coming years where there is a very clear increase in the number of mortalities as well as a very clear correlation between the deaths and the increase of iron in the blood. The highest recorded cases of death are for 2020 (**19.28692**) and it is clear that the number of deaths will increase in the coming years, and whenever iron is increased the proportion of patient will increase.

5- Cause of mortality for patients with thalassemia:

The accumulation of iron levels in the body resulting from hundreds of blood transfusions,

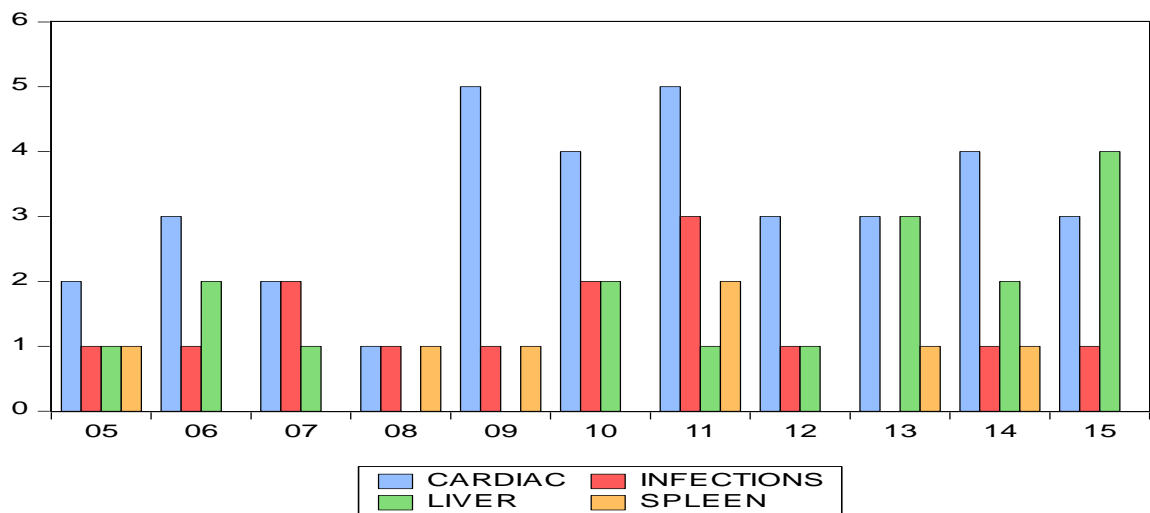
along with the increase in the absorption of iron from food, cause vital organs and glands in patients with thalassemia to experience induced early mortality. In this study, examined which diseases represent the greatest that cause of mortality. We took the cases of death of thalassemia patients in the province of Maysan for the period 2005-2015 and the number of mortalities during this period, which was 73 as well as information from the thalassemia center in the province and based on the data collection we found that the primary causes of these deaths were:

1. Cardiac disease
2. Infections
3. The liver
4. The spleen



Table (13) the mortality cause for thalassemia patients for the years 2005-2015

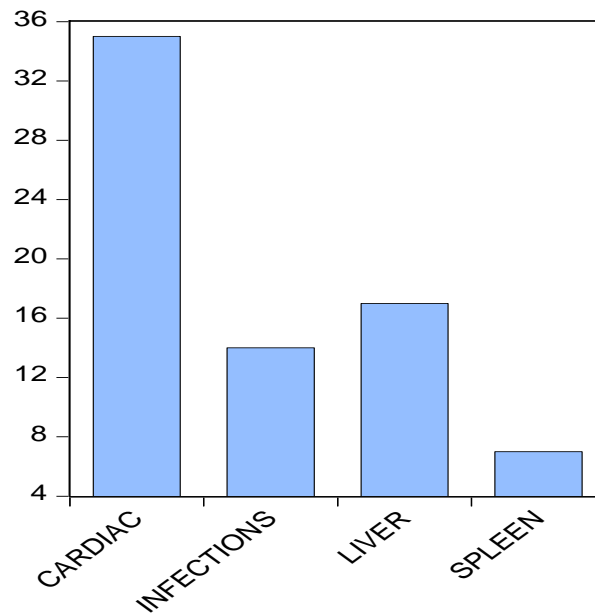
The mortality cause	Cardiac disease	Infections	The liver	The spleen
Years				
2005	2	1	1	1
2006	3	1	2	0
2007	2	2	1	0
2008	1	1	0	1
2009	5	1	0	1
2010	4	2	2	0
2011	5	3	1	2
2012	3	1	1	0
2013	3	0	3	1
2014	4	1	2	1
2015	3	1	4	0
Sum	35	14	17	7
Percentage	48%	19%	23%	10%



Graph (6) the graph for the diseases causing the death, distributed according to the years 2005-2015



Sums



Graph (7) The percentage of diseases that cause the death

Result:

Through a table and graphs we found that the main cause of death was heart disease. There were 35 deaths due to heart disease (48%), followed by disease of the liver (17 cases, 23%), infections (14 cases, 19%), and disease of the spleen (7 cases, 10%).

Conclusions:

1. The VAR (vectorial autoaggressive) model has the ability to detect a causal relationship between thalassemia and iron in the long term.
2. VAR models are flexible and are relatively easy to work with both in theory and practice.
3. In general thalassemia patients live with more risks than people without thalassaemia.
4. Despite the use of iron chelating therapy in recent years in Iraq, serious complications can result in deaths who suffer from high iron levels.
5. High levels of iron in the blood cause several diseases as a result of deposition of iron in the body's organs.
6. The leading cause of death among people with thalassemia is related cardiac issues.

Recommendations:

1. The model should be used in the prediction of other diseases, and the adoption of forecasts given by the study put of many of the problems solutions of thalassemia patients to reduce the burden of this disease.
2. Thalassemia is a reality in our country. It must be accepted as a public health problem in the long term and awareness campaigns should be initiated that make people more familiar with the seriousness of the disease in an attempt to reduce its prevalence by conducting tests before marriage.
3. Cardiac-related complications are the leading cause of death among people with thalassemia, and it is therefore very important for patients who do not have cardiac issues to take steps to prevent these complications from developing, and for patients who already do have cardiac issues to work to reverse them or prevent them from getting worse.

References:

1. Aessopos A, Farmakis D, Hatziliami A, Fragodimitri C, Karabatsos F, Joussef J, et al: "Cardiac status in well-treated patients with thalassemia major", *Eur J Haematol*. 2004; 73: 359-366.
2. Bera A K, and Jarque C M: "An efficient large Sample test for normality of observations and regression residuals", Working paper in *Econometrics* No 40, Australian National University, Canberra; 1982.
3. Cromwell J B Hannan M J, Labys W C and Terraza M: *Multivariate Tests for Time Series Models*. SAGE Publications, Inc California; 1994. pp. 73-75.
4. Dickey D. and Fuller W: "The likelihood ratio statistics for autoregressive time series with a unit root", *Econometrica*, 1981; 1057-1072
5. Dickey D. and Fuller W "Distribution of the estimators for autoregressive time series with a unit root", *Journal of the American Statistical Association*, 1979; 74: 427- 431.



6. Dickey and Fuller W A "Distribution of the estimators for autoregressive time series with a Unit Root", *Journal of the American Statistical Association*, 1979; **74** (366): 427–431. doi:10.2307/2286348. JSTOR 2286348.
7. Galanello R Origa R: "Beta-thalassemia", *Orphanet Journal of Rare Disease* 2010; 5.(11): 1750-1172 (electronic).
8. Hacker R S and Hatemi-J A: "Optimal lag-length choice in stable and unstable VAR models under situations of homoscedasticity and ARCH", *Journal of Applied Statistics* 2008; 35(6): 601– 615.
9. Henin P Y: Bilans et essais sur la non-stationnarite des series macroeconomiques revue d' economie politique; 1989; 5: 661-691.
10. Kirchgassner G and Wolters J: *Introduction to Modern Time Series Analysis*. springer-Verlag, Berlin Heidelberg;2007; pp. 13-14.
11. Lal A, Porter J, Sweeters N, Ng V, Evans P, Neumayr L, et al: "Combined chelation therapy with deferasirox and deferoxamine in thalassemia",. *Blood Cells Mol Dis*. 2013; 50; 99-104.
12. Lardic S et Mignon V: *Econometrie des Series Temporelles Macroeconomiques et Financieres*. Ed. Economica. Paris; 2002 pp. 97.
13. Ljung G M, and Box G E P: "on a measure of the lack of fit in time Series models", *Biometrika*, 1978; 65: 297-303.
14. Prati D: "Benefits and complications of regular blood transfusion in patients with beta thalassaemia major", *Vox Sang*. 2000; 79; 129-137.
15. Quenouille MH: *The Analysis of Multiple Time-Series*. Griffin, London, 1957.
16. Rund D and Rachmilewitz E: "Beta-thalassemia", *N Engl J Med*. 2005; 353: 1135-1146.
17. Shander A, Cappellini MD and Goodnough LT: "Iron overload and toxicity: the hidden risk of multiple blood transfusions", *Vox Sang* 2009; 97; 185-197.
18. Shumway RH and Stoffer D S: *Time Series Analysis and its Applications* Springer, New York; 2006; pp. 303-304.
19. Sims C A: "Macroeconomics and Reality", *Econometrica*, 1981; 48: 1- 48.
20. Toumba M, Sergis A, kanaris C and Skordis N: "Endocrine complications in patients with thalassemia major", *Pediatr Endocrinol rev*, 2007;5: 642-648.
21. Wetherill DJ and Clegg JB: *The Thalassaemia Syndromes*. Blackwell Science 2001.
22. Wood JC, Origa R, Agus A, Matta G, Coates TD and Galanello R: "Onset of cardiac iron loading in pediatric patients with thalassemia major", *Haematologica*, 2008; 93(6) 917 – 920. doi:10.3324/haematol.12513