Identifying Factors Affecting the Incidence of Chronic Diseases in Algeria

Hachem Amel [1*] **Belabbes Fadila** [2] **Bekhtaoui Assia** [3]

[1]University Oran 2 Mohamed Ben Ahmed, Algeria

*** Corresponding author:** Hachem.amel@univ-oran2.dz

https://orcid.org/0000-0003-1197-329X

[2]University Saida Moulay El Taher, Algeria

belabbesfadila@gmail.com

[3]University Oran 2 Mohamed Ben Ahmed, Algeria

Bakhtaoui.assia@univ-oran2.dz

https://orcid.org/0009-0002-5720-0401

Received: 12/08/2023 **Accepted: 09/01/2024** **Published: 19/01/2024**

Abstract

In this article, we will shed light on chronic diseases which are one of the most common diseases in the world, as a result of the serious complications that result from them, such as motor and visual disability or sometimes even death. This intervention comes to provide a picture of the factors influencing the incidence of chronic diseases in Algeria through the use of a binary logistic regression model in order to extract the most important social, demographic and economic factors influencing the incidence of chronic diseases. In this research, we relied on the database of the Sixth Multiple Indicator Cluster Survey (2019).Therefore, The results of the study showed that social and economic factors have an impact on the incidence of chronic diseases, and gender, place of residence, and educational level are considered the most important demographic factors.

Keywords: chronic diseases, Logistic regression, Demographic factors, Economic factors, Social factors.

Tob Regul Sci.™ 2024;10(1): 408 - 426

DOI: doi.org/10.18001/TRS.10.1.28

1- Introduction

Chronic diseases are considered the more common diseases spread throughout the world. According to the World Health Organization, these diseases constitute the first cause of death(OMS, 2019), especially heart and arterial diseases for both sexes 16% (WHO report, 2019). No communicable diseases, which include heart disease, diabetes, cancer, and the

respiratory system, occupy seven of the top ten causes of death globally and which are only four of the top ten causes of death in Early 2000 (OMS, 2019)

During the last two decades, the world has witnessed many changes, some of which affected the field of health. On the one hand, the average lifespan has increased a person ages from approximately 67 years to 73 years as a result of the remarkable progress in reducing child mortality. After infectious diseases in children were the main cause of death, non-communicable diseases in adults have now become the main cause of this. Instead of hunger, over eating has become one of the most important and dangerous factors in contracting the disease, and has increased the rise and spread of chronic diseases, since most people live the last years of their lives in disability and illness, especially in developing countries. According to the World Health Organization, chronic diseases are constantly increasing, especially in developed and developing countries, as a result of the lifestyles that people live in these countries, which are widespread in terms of pressures resulting from social and economic conditions, and negative nutritional behaviors. These diseases, that is, non-communicable diseases, especially chronic diseases, have become the main cause of death or disability for many residents.

We also record a decrease for there are differences between developed and developing countries regarding the incidence of some chronic diseases such as diabetes and high blood pressure. According to United Nations statistics, chronic diseases were the cause of death. Nearly 9 million of the population (OMS, 2019).As a result of the change in dietary pattern, lack of physical activity, and the spread of the scourges of smoking, in addition to developments witnessed by societies, such as the high rate of urbanization, which has resulted in several problems such as anxiety and obesity, and are considered one of the causes of the spread of chronic diseases in all the countries of the world. Consequently, the world faces a heavy burden of chronic diseases and their severity, which constitute the greatest challenges to development.

Algeria, like the rest of the countries of the world, has recorded in recent years a shift in the health map from infectious diseases to non-communicable diseases, most notably chronic diseases, the most important of which are heart and arterial diseases, diabetes, and high blood pressure. What draws attention is the spread of these diseases among children and young people compared to what was the situation in the millenniums, which led specialists to sound the alarm. According to statistics from the Ministry of Health, Population and Hospital Reform and the World Health Organization, the rate of chronic diseases in Algeria is high, as the death rate from these diseases in 2016 was estimated at 76%, most notably heart disease ((OMS 2018: 17) compared to only 63% in 2010 of total deaths, with heart diseases accounting for 28% of these deaths (OMS 2011:17).

Studies have shown that 10% of the population are infected with chronic diseases. This percentage increases in urban than rural areas (55% and 45%, respectively). There are also

females more incidence of these diseases compared to males (60% vs. 40%). The age group is 60 years and above more infection (scanning Cluster 3). The high number of people infected with these diseases, especially among children and youth, constitutes a major obstacle to sustainable development, which attracts the attention of officials and makes us wonder what the socio-demographic and economic factors are affecting the incidence of chronic diseases?

2- Objectives of the study:

This study aims to:

- Identify the most important factors that affect the incidence of chronic diseases.
- Build a mathematical model that explains the relationship between infection with a chronic disease and socio-demographic and economic factors, which helps, first, to understand the causes of infection with one of these diseases, and second, to develop new strategies to eliminate or reduce its spread.

3- Methodology and study sample:

3-1- The Study Approach:

To achieve the objectives of the study, we relied on the descriptive and analytical approach. The descriptive aspect represented the use of frequency tables to describe the sample, while the analytical aspect represented the construction of a mathematical model represented by binary logistic regression. It is one of the methods of multivariate analysis that seems to us the most appropriate in our study, after checking the quality of data matchmaking.

The descriptive approach allows us to identify the characteristics of people with chronic diseases and test the relationship between disease incidence and socio-demographic and economic variables. To implement this analysis, we set the error percentage atα= 5%, from which we take the decision. First, we will discuss the methodological framework by first presenting the study variables and secondly, introducing the source of the data used, which is cluster survey MICS6, and we will of course discuss the quality of the data.

3-2-The study sample:

The study sample included 107537 participants aged 15 years or older, with 49.25% females. The percentage of people with chronic diseases represents one fifth of the sample (20%), i.e. 20345persons.

It included the sixth cluster survey 5Forms: Form data has been used to study chronic diseases one Represented in the household file database, which enabled us to obtain the socio-demographic and economic characteristics of the sample that included 20345persons as we

mentioned before, suffering from at least one chronic disease of both sexes and their ages range 15 And 85 year.

3-3- Study variables:

It is limited to the dependent variable and the independent variables related to this study.

A- The dependent variable:

The dependent variable in this study is the presence of a chronic disease, which is a variable with two characteristics: the person has a chronic disease or does not have it. It is a binary qualitative variable; this means it contains two answers or two characteristics. Either the person suffers from a chronic disease and we have indicated it with the symbol"1", or it is not infected and is symbolized by "0".

B- Independent variables:

The sixth cluster survey database included MICS 6 is a group of variables that may directly or indirectly affect a person's infection with a chronic disease, that is, the dependent variable under study. Some of them are quantitative, such as age, and are qualitative, such as the wealth index, gender and place of residence, educational level, civil status, individual status, and geographical region.

4- Introducing the data source MICS 6:

The data source for this study is the recent Multiple Indicator Cluster Survey MICS 6, conducted between December 2018 and January 2019 in Algeria. It is considered one of the most important sources due to the availability of its own database, which contains the study variable first and a group of variables related to the primary variable, which is the incidence of a chronic disease. It is the only source at this time that enables us to carry out an analytical study of the situation of chronic diseases in Algeria.

– **What is a multiple indicator cluster survey MICS 6:**

It is a household-based multiple indicator survey. It was first conducted in the mid-1990s by UNICEF, with the aim of assessing progress achieved against statements and development plans for children (kindergarten and child health), in New York in September 1990.

As for Algeria, the first multiple indicator cluster survey was conducted in 1995 (MICS 1). MICS 2 came after it a year later in 2000,then MICS 3 in 2006 and after that survey the fourth MICS4 during 2012-2013; moreover, the last survey, which is the subject of the study 6MICS between the end of 2018 and the beginning of 2019.

The goal of this latest survey MICS 6 aims to collect information to determine the progress made in achieving the Millennium Development Goals (ODD), adopted in 2015 and global programs for children, and this survey was conducted on a sample of 31325 households distributed across all five national health regions, and only 29919 were questioned, representing a response rate of 96.7%.

Five questionnaires were used in this survey, including a questionnaire directed to the family, which is the subject of our study because it includes data related to the characteristics of people with chronic diseases who are over 15 years old.

5- Data quality assessment:

It is one of the most important steps since it is considered necessary before starting any analysis. It is important to evaluate the quality of the data. This process allows us to estimate the internal and external consistency of the data, since the reliability of the results of statistical analysis requires that the data used must be of high quality, which means that the non-response rate must be less than 10%. To evaluate the quality of the data in our study, we calculated the non-response rate for the various study variables which gave good results. Generally, except for the educational level variable, the Place of residence and wealth index, so that it will exceed the percentage of the Missing values more than 10%.

Schedule1: Distribution of the percentage of missing values for study variables

Variables	Declared values	Missing value	Non-response rate%
Suffer from at least one chronic disease	107655	118	0,11
Gender	107537	00	00
Age	107537	00	00
Educational level	107537	22889	21,28
Wealth index	107537	73380	68,24
Individual status	107537	0	00
Place of residence	107537	86382	80,33
Geographic Region	107537	00	00
Civil status	107537	00	00

Source: Calculation by researchers through the database of the cluster survey MICS6

To apply logistic analysis, it is required that there will be no missing values, so we first deleted the missing values from the file of people whose ages range from 15 years and above. Thus, after filtering the file, the sample became estimated in 107537 persons. Then, we deleted the variables that exceeded the percentage of non-answers, that is, the percentage of missing values 10%. This deletion included the educational level variable, the place of residence variable, and the wealth index, despite their importance in explaining the incidence of chronic disease. It must be noted that this survey was not conducted for a demographic purpose, But for health purpose.

6- Binary logistic regression:

The use of binary logistic regression aims to analyze the relationship between the explanatory independent variables and the binary dependent variable, which is whether the person has or does not have a chronic disease, based on a database 6 MICS, including building a mathematical model that explains this relationship.

6-1-The concept of Binary logistic regression:

Regression analysis in general is a model that analyzes and interprets the relationships that exist between the dependent variable and a group of independent variables explained by linking these variables to a mathematical equation that may be linear (linear regression), or it may be non-linear (regression non- linear) (Rico, 2009). After determining the form of the relationship, we estimate the parameters of the model that express the amount of influence of the independent variables on the dependent variable, for the purpose of interpretation or prediction according to the nature of the study.

Logistic regression is nothing but a type of regression in which the dependent variable is specific and may take two values (binary logistic regression), or it may take more than two values, and it is called multiple logistic regression.(Tuffery,2010).

It should also be noted that in logistic regression, our goal is not to explain the change in the values of the dependent variable, But explaining the probability of the occurrence or non-occurrence of the phenomenon under study. It represents a logistic regression equation as follows:

$$Pi = E\,(yi/Xi) = \frac{e^{(B0+B1X1+\ldots+BkXk)}}{1+e^{(B0+B1X1+\ldots+BkXk)}} \qquad (1)$$

where:

Pi = probability of occurrence of category or attribute I in the dependent variable.

6-2-Representation of model features:

Through the equation, it is clear that the relationship between the dependent variables and the independent variables is non-linear, through the use of logic transformation (logit). We make the relationship linear, so the previous equation 1 becomes as follows:

$$L = \log \frac{Pi}{1-Pi} = B_0 + B_1X_1 + \ldots + B_KX_K \quad (2)$$

So the amount $(\frac{Pi}{1-Pi})$ represents the odds ratio, which is considered the cornerstone in interpreting the model parameters and is interpreted according to the type of the explanatory (independent) variable.

6-3-Estimate model parameters:

The parameters of the logistic regression model are estimated using the likelihood method the greatest (Maximum of) until we obtain the appropriate equation, then we solve these equations numerically through iterative methods.(Adel ben Ahmed Babtain,2009).

6-4-Evaluation of the Sample:

Before relying on the model estimation results for the purpose of interpretation or prediction, this model must pass several statistical and standard tests to ensure validity for use. These tests can be divided into two types:

- Overall evaluation of the model:

This is done through Match making quality standards and tests of the total significance of the model.

6-5- Match making quality standards:

They are statistical criteria that measure the explanatory power of the model, and they are considered alternatives to the coefficient of determination for linear regression, so they are called quasi-coefficients of determination, which are calculated by comparing the explanatory power of the model without explanatory variables with the model after introducing the explanatory variables. There are several parameters, the most common of which are: R-two of Cox and Snell و of Nagelkerke R-two.

6-6- Testing the overall significance of the model:

This test aims to know the overall significance of the model parameters, which means whether all the parameters of the explanatory variables are equal to zero or is there at least one parameter different from zero. To do this test, we use the same principle in the case of ordinary regression, that is, comparing the expected values in the case of the model without independent variables

with the expected values in the model that contains the independent variables, and this test is called Hosmer test and Lemeshou.

6-7- Classification tables:

They are tables consisting of the observed classification of cases and the classification generated or produced by the model. Through these tables, we can know the percentage of correct classification and the percentage of incorrect classification. The higher the correct classification ratio, the better predictive the model. These tables are often used if the purpose of building the model is predictive rather than explanatory.

6-8- Significance tests of model parameters:

These tests aim to determine the statistical significance of each variable separately. A test is usually used Wald or odds ratio test in order to determine statistical significance. Each of these tests has advantages and disadvantages.

In this study, we denote the dependent variable by and for independent variables by X_k, where k takes the numbers from 0 to N. The nature of the dependent variable y is two types of binary attributes, symbolized by 0 and1.We give the value 1 to the characteristic that is the focus of the study and 0 if it is not, in the case of this study "suffering from a chronic disease. ' take (1) and "not infected" (0). The independent variables may be quantitative or qualitative.

So, if (P) is likely to be the person 'who has a chronic disease', then (P-1) is the probability that the event will not occur, which means that he will not suffer from a chronic disease.

7- The results of applying binary logistic analysis:

7-1-Sample properties of the study:

Schedule 2: Distribution of sample members according to gender and infection with one chronic disease at most

Gender	Status				Total	%
	Suffering from a chronic disease		Uninfected			
	Number	%	Number	%		
Males	8264	40,62	46344	53,15	54608	50,78
Woman	12081	59,38	40848	46,85	52929	49,22
Total	20345	18,92	87192	81,08	107537	100

Source: Prepared by the researchers through database analysis MICS6

Through table 2, we find that the incidence of chronic diseases is approximately 19%, which is much higher than the percentage declared by the Ministry of Health, Population and Hospital Reform. The results also showed that women are more affected by chronic diseases compared to males, as their percentage reached 59.38% compared to 40.62%.%; a difference between them of approximately 20 points.

Schedule3: Distribution of sample members according to age and infection with at least one chronic disease

Age	the condition				Total	%
	Suffering from a chronic disease		**Uninfected**			
	Number	%	Number	%		
15-29	1310	6,44	35473	40,68	36783	34,20
30-59	9579	47,08	44493	51,03	54072	50,28
60 and more	9456	46,48	7226	8,29	16682	15,51
Total	20345	18,92	87192	81,08	107537	100

Source: Prepared by the researchers through database analysis MICS6

Table 3 shows that people between the ages of 30 and 59 are more likely to suffer from chronic diseases, reaching 47.08%, slightly one point ahead of older people (60 years and over 46.48%), the later is much higher than those who are not affected. In the last place (third) are young people between the ages of 15 and 29 in terms of infection with 6.44%.Moreover, this distribution applies to females, i.e. the age groups most affected by at least one of the chronic diseases are 30 to 59 years, but for males it is different. Males aged 60 and over have the highest percentage (50.99%).

Schedule 4: Distribution of people with chronic diseases according to the type of disease and gender

Type of chronic disease	Males		Woman		Total	
	Number	%	Number	%	Number	%
High blood pressure	2729	33,02	5179	42,87	7908	38,87
Diabetes	2325	28,13	2689	22,26	5014	24,64

Type of chronic disease	Males		Woman		Total	
	Number	%	Number	%	Number	%
Heart disease	593	7,18	586	4.85	1179	5,80
Asthma	732	8,86	813	6,73	1545	7,59
Joint diseases	297	3,59	778	6,44	1075	5,28
Cancer types	104	1,26	201	1,66	305	1,50
Kidney failure	118	1,43	94	0,78	212	1,04
Neurological and psychological disorders	837	10,13	591	4,89	1428	7,02
Genetic diseases (hemophilia, thalassemia, celiac disease)	40	0,48	74	0,61	114	0,56
Other chronic diseases	489	5,92	1076	8,91	1565	7,69
Total	8264	100	12081	100	20345	100

Source: Prepared by the researchers through database analysis MICS6

It is clear from the data of the table above that high blood pressure disease occupies the front in both sexes with a percentage of 3887%, with a difference not exceeding eight points between males and females. Diabetes comes in the second place approximately 25% with a slight difference between the sexes in favor of males.

7-2- The building of a binary logistic regression model:

We previously mentioned that in this study we will use logistic regression to study the relationship between the explanatory variables, which are represented by variables such as socio demographic and economic variables, and the dependent variable having at least one chronic disease of binary adjective, by constructing a model that explains this relationship.

7-2-1- The display of the study variables:

In this model, we used five independent variables in addition to the dependent variable, which are as the followings:

Dependent variable: “Having at least one chronic disease,” which is a nominal variable consisting of two adjectives (infected and uninfected).

Independent changes:

X_1: The gender variable represents a nominal variable consisting of two characteristics, male and female.

X_2: The variable represents the age group, which is a categorical variable and takes three characteristics: 15-29 years old, 30-59 years old, 60 years old and above.

X_3: Civil status variable, which is a nominal variable that takes 5 characteristics: married, divorced, separated, and widowed.

X_4: represents the geographic region variable, a nominal variable with 4 characteristics: central, south, east, and west.

X_5: Single case variable, which is a nominal variable that holds eight characteristics: office holder, unemployed, unemployed, other, housewife, retiree, student/pupil, service nationalism.

Schedule 5: Dependent variable symbols

Code	Dependent variable
1	Suffer from at least one chronic disease
0	Uninfected

The table above shows the codes for the dependent variable, where the code for infection with at least one chronic disease is 1 and the number for non-infection is 0.

7-2-2- The results of the logistic regression model:

The model results are presented in two parts. The first part includes results for evaluating the model statistically in order to determine the strength of the model in explaining the studied phenomenon. As for the second part, which is considered the most important, we devote it to presenting the results that help us explain the relationship that exists between the dependent variable and the explanatory variables.

Estimating a binary logistic regression model:

- The overall evaluation of the model:

To determine the explanatory power of a model, the results of a model that includes only the constant term (i.e., no independent variables) are often compared with a model that may help for explaining (independent) variables.

Schedule6: Frequencies and maximum potential function value for the model without explanatory variables

Iterations steps	-2-log-vraisemblance The value of the maximum potential function	Constant parameters
1	105095,487	-1.243
2	104323,628	1,443-
3	104321,071	1,455-
4	104321,071	1,455-

Source: Prepared by the researchers through database analysis MICS6

This table represents the value of the maximum potential function in the case of the model that contains the fixed limit only, i.e. without the explained variables, where the value of the maximum potential function stabilized at 104321.071 at the fourth stage or step, where the change in value became less. We will compare this value of 0.01 with the model involving independent variables. If the introduction of independent variables leads to a decrease in this value, we say that these variables affect the dependent variable.

Thus, we move on to examining the regression equation table before introducing the independent variables.

schedule7: Equation Regression without explanatory variables

	Constant value	The errorStandard	Value Wald test	Degree of freedom	indication	Odds ratio
Block 0 fixed	- 1,455	0,008	34935,582	1	0,000	0.233

Source: Prepared by the researchers through database analysis MICS6

The results show that there is statistical significance, which indicates the model's ability to predict the dependent variable before introducing the explanatory variables.

Table 8: Correct classification without independent variables

		Expected		
		chronic diseases		
		0	1	%the correct
Watching chronic diseases	0	87192	0	100,0
	1	20345	0	,0
Total ratio				81,1

Source: Prepared by the researchers through database analysis MICS6

The overall correct classification rate before introducing the explanatory variables was estimated at 81.1% (Table 8).

Table 9: The effect of independent variables before entering them into the model

		Score	ddl	Sig.
Step 0	Age	18724,723	1	,000
	Gender	1036,590	1	,000
	civil status	11070,304	1	,000
	Individual status	5976,369	1	,000
	Geographic region	11,716	1	,001
	Total	21445,707	5	,000

Source: Prepared by the researchers through database analysis MICS6

Table 9 shows the significance of the independent variables that were excluded from the model and their impact if they were added to the model, in the sense that measures the significance of each independent variable before entering it into the model, whether this variable, if it is entered into the model, will improve or not and therefore whether it is worth entering it or not. It has been shown that all variables are statistically significant since the level of Sig is less than the level of significance 0.05, which means that these variables, if introduced into the model, will improve its ability to interpret and predict the dependent variable, which is the incidence of at least one of the chronic diseases. After this comes the stage of presenting the model with the independent variables.

schedule10: Frequencies and likelihood function value for the model including the explanatory variables

Repeating steps	-2 log of Likelihood	Constant	X_4 Individuel status	X_1 gender	X_3 civil status	X_2 Age group	X_4 Geographic Region
1	88220,899	-3,331	,108	,048	,172	,766	-,017
2	82464,821	-5,069	,178	,099	,211	1,273	-,024
3	81955,887	-5,799	,208	,135	,223	1,479	-,025
4	81949,122	-5,898	,212	,142	,224	1,505	-,025
5	81949,120	-5,899	,212	,142	,224	1,505	-,025
6	81949,120	-5,899	,212	,142	,224	1,505	-,025

Source: Prepared by the researchers through database analysis MICS6

The parameters of the logistic regression model are presented in an iterative manner, meaning that the calculation is repeated until the parameter values converge at a specific allowable error and at the smallest value of the maximum potential function. From this table, the following can be observed:

First: The parameters of the model are stabilized at the sixth step; they are similar to the fifth step, where we notice the values of the fifth step are the smallest value compared to the values of the previous steps.

Second: The estimation process stopped at the smallest value of the probability function (81949,120), and we note that this value is much smaller than the value of the greatest possibility function in the case of the model containing the fixed limit only (104321,071), this indicates a relationship between the dependent variable and the interpreted variables, and also indicates that these independent variables that were introduced into the model contributed to improving its ability to interpret and predict the dependent variable.

Schedule11:It shows the Chi-square test to test the significance of the model with the degrees of freedom and their statistical significance in the model.

		CHi-square	Ddl	Sig.

Step 1	Pass	22371,951	5	,000
	Bloc	22371,951	5	,000
	Model	22371,951	5	,000

Source: Prepared by the researchers through database analysis MICS6

Table 11 tests the significance of the model using the chi-squared test to test the morale. We find that the value of the chi-square from the table is equal to 22371,951 and that the p-value of the Chi squared statistic squared is 0.000, which is therefore significant at the level of statistical significance (0.05), and therefore we reject the null hypothesis (zero) and accept the alternative hypothesis, that is, the model is significant and represents the data well

schedule12: Matchmaking quality coefficients

-2 log of vraisemblance	R-square de Cox and Snell	R-square of Nagelkerke
81949,120	0.188	0.302

Source: Prepared by the researchers through database analysis MICS6

This table represents 12 coefficients that express the explanatory power of the model and therefore we note that the value of these coefficients is average, but they are considered acceptable in such models.

We then examine the extent to which the model matches the study data through the Housmer and Lemshaw test, which is a non-parametric test that is also used to ensure the quality of the model's matchmaking, as it depends on the chi-squared value of the hypothesis test.

Null hypothesis H0 The observed cases are equal to the expected cases, and this means that the model represents the data well.

Alternative hypothesisH1: The observed cases are not equal to the expected cases and therefore the model does not represent the data well.

Schedule13: Results of the Homser and Lamshow test for For the quality of model matching of fit of the model (Hosmer - Lemeshow)

Step	Chi squared	Moral level ddl	Significance level sig.
1	257,960	8	0,000

Source: Prepared by the researchers based on database analysis MICS6

The table above shows the results of the Hosmer and Lemshow test to ensure the quality of the matching model using the chi-square statistic.Through schedule 13, we find that the chi-square value is 257,960 at 8 degrees of freedom and the probability value of the chi-square statistic is 0.000, which is smaller than the significant value of 0.05, and therefore we reject the null hypothesis and accept the alternative hypothesis, meaning that the model does not represent the data well.

Table n°14: Shows the observed and expected values of the Hosmer and Lemshow test

		Uninfected		Suffering from a chronic disease		Total
		To watch	Expected	To watch	Expected	
Step 1	1	9942	9945,484	232	228,516	10174
	2	10362	10381,897	320	300,103	10682
	3	10576	10676,403	593	492,597	11169
	4	9404	9038,817	409	774,183	9813
	5	6957	7006,722	867	817,278	7824
	6	9302	9348,640	1256	1209,360	10558
	7	8609	8784,717	2267	2091,283	10876
	8	7592	7727,249	2559	2423,751	10151
	9	7999	7823,156	2756	2931,844	10755
	10	6449	6458,917	9086	9076,083	15535

Source: Prepared by the researchers through database analysis MICS6

Table 14 represents the distribution of the observed values in the sense of realism and expected values i.e. predicted by the model. From these results, we can see small differences between the two values.

Schedule15:Shows regression parameters, standard error, and statisticsWald with their degrees of freedom and their significance for the model in the presence of explanatory variables.

Explanatory	Logit	Standard	Father	Degree of	Calculated significance	Exp(B)	95% confidence

variables	value	error		freedom	level	Odds ratio	interval	
							lowest	the above
Age group	1,505	0,017	7945,927	1	0,00	4,506	4,359	4,658
Individual Status	0,212	0,007	980,713	1	0,00	1,236	1,220	1,253
Gender	0.142	0.022	41,707	1	0,00	1,152	1,104	1,203
civil status	0,224	0,011	452,063	1	0,00	1,251	1,225	1,277
Geographic region	- 0.025	0,011	5,521	1	0,019	0.976	0.956	0,996
Constant	- 5,899	0,050	14069,747	1	0,00	0,003		

Source: Prepared by the researchers through database analysis MICS6

Table 15 shows the optimal model parameters that we obtained in the fifth cycle and their estimates, in addition to the standard error for each parameter and the statistics Wald for each parameter along with the number of degrees of freedom, its significance statistics. The table shows the significance of all variables and their impact on the incidence of at least one chronic disease. We also find that the geographical area variable has a negative impact on the incidence of at least a chronic disease, while gender, age, individual status and civil status positively affect the incidence of at least a chronic disease.

Therefore, the binary logistic regression equation for the model can be written as follows:

$$\log \frac{Pi}{1-Pi} = - 5,899 + 0.212 \text{ Individual status} + 0.224 \text{ Civil status} + 0.142 \text{ Gender} + 1,505 \text{ Age group} - 0.025 \text{ Geographical aregion.}$$

And the order of the variables according to the degree of their influence on the dependent variable is as follows:

Rank	1	2	3	4	5
variable	Age	Individual status	civil status	Gender	Geographic region

Through this table, age is considered to have the most influence on the dependent variable, followed by the individual status variable, and then come the other variables.

The Interpretation of the odds ratio (Exp(B)):

- Regarding age: The age variable ranked first in terms of its influence on the dependent variable of having at least one chronic disease. The probability of people aged 60 years and over being infected with at least one chronic disease is four and a half times (4,506) times the probability of contracting people aged less than 60 years.
- Regarding gender: The probability of females being infected with chronic diseases is more than twice that of males.
- As for the individual case, the probability of a person with a profession contracting a chronic disease is one time (1.36) greater than that of a person who does not work.
- Geographical region: The probability of residents of the South being infected with chronic diseases is approximately one percent less than that of residents of the North.
- Regarding civil status: Married people are one time (1,152) more likely to be afflicted with chronic diseases than other groups.

Schedule16: Correct classification of the model.

		Does he suffer from a chronic disease?		Correct percentages
	Expected	Uninfected	Suffering from a chronic disease	
Step 1	**Uninfected**	81603	5589	93,6
	Suffering from a chronic disease	12132	8213	40,4
	General percentages			**83,5**

Source: Prepared by the researchers through database analysis MICS6

Table 16 represents the correct classification ratio of at least chronic disease and non-infection. The results showed that the model was able to correctly classify 93.6% of uninfected people, while only 40.4% of people with a chronic disease were properly classified. The overall correct rating ratio is estimated at 83.5%, which is considered a good percentage and expresses that the model has a good ability to predict the dependent variable.

Conclusion:

Chronic diseases are among the problems of the present era and one of the leading causes of death. The cultural development witnessed by societies and the behavioral changes of individuals, especially with regard to nutrition, and the results of this development are considered among the main reasons for the spread and rise of these chronic diseases such as smoking, anxiety, lack of exercise, modern nutrition, obesity, etc... and among the basic goals of the state's health policy. Reducing, controlling and preventing these diseases due to the financial cost they cause, as well as the social cost, which harms the country's economy in general.

References in Arabic:

[1] United Nations General Assembly, September 16 report2011.

[2] Ministry of Health, Population and Hospital Reform, Government Policies in the Field of Health, September 2015.

[3] Ghanem Adnan Andcome Awni Farid Khalil (2011),Use technology Binary response logistic regression in studying the most important economic determinants Andsocial "For the adequacy of family income," applied to a random sample of families in a governorate Damascus, Damascus Journal of Economic Sciences Andlegal Volume 27 Issue the first.

[4] **In language Foreign:**

[5] MSPRH (2014); Rapport MICS 4 (2012-2013)

[6] MSPRH, National Multisector Integrated Strategic Plan to control risk factors for non-transmissible maladies 2015-2019, page 19-22-27-31-39.

[7] OMS (2006), Prévention des maladies chronicles un vital investissement, 2006, page 39.

[8] OMS,PLan d'action 2013-2018 pour la strategy mondiale de lutte contre les maladies non-transmissibles.

[9] Rico Racotomolala (2009), Pratique de la regression logistique. Lumière University Lyon.

[10] TUFFERY Stephane (2012). Data mining and accurate statistics.Edition TECHNIP. Paris

[11] WHO (2014), Non-communicable diseases progress monitor;Geneva2015, p. 236

[12] WHO (2015), non communicable diseases 2014. Geneve;who;2015;p280.

[13] WHO(2018), noncommunicable diseases 2016. Geneve;who;2018;p280.

[14] WHO(2011), noncommunicable diseases 2010. Geneve;who;2011; p28.