Estimation of the prevalence of diabetes and study of some socio-demographic determinants among the diabetic population in Salé, Morocco-2021

Abdelghani Asraoui[1,*], *Abdelmajid Soulaymani*[2] and *Chems Eddouha Khassouani*[3]

[1,2] Center for Doctoral Studies, Faculty of Sciences, Ibn Tofail University, Kénitra, Morocco
[3] Hospital of August, 20th, Ministry of Health, Ifrane, Morocco

Abstract. According to the World Health Organization, diabetes is the cause of 1.5 million deaths each year, 80% of which occur in low- and middle-income countries. It will become the principal cause of death by 2030. In Morocco, the situation is as alarming as at the global level. However, the prevalence of diabetes at the provincial level is unknown. This paper aims to present partial results obtained from a statistical study, carried out in February 2021, where we estimated the prevalence of diabetes and studied some socio-demographic determinants among the diabetic population living in Salé, Morocco. The investigation included a sample of 488 households selected according to a two-stage stratified probabilistic sampling plan (466 of which were surveyed (i.e. 1868 individuals) with a response rate of 466/488 =95.5%). The results showed that, in 2021, diabetes was affecting 5.5% (95% CI 4.5%, 6.6%) of the studied population (5.8% of women and 5.1% of men, 5.6% in urban areas and 4.5% in rural areas). The proportion of diabetics increases significantly with age (p-value=0.0001<0.05). It was 0.1% among young people (<18 years), 3.9% among individuals aged 18-59, and 26.9% among those aged 60 and above. A logistic model, making it possible to predict, with an accuracy of 91.7%, the risk of contracting diabetes among the population aged 18 and above (1,308 subjects, 7.8% of which were diabetic), was performed. The model included several socio-demographic determinants. Age, educational level, economic activity, average household income and smoking were significantly associated with diabetes. **Keywords:** Prevalence, Diabetes, Socio-demographic determinants, Logistic regression, Salé-Morocco.

1 Introduction

According to the World Health Organization (WHO), Noncommunicable diseases (NCDs), including heart disease, stroke, cancer, diabetes and chronic lung disease, are collectively responsible for almost 70% of all deaths worldwide. Almost three-quarters of all NCD deaths, and 82% of the 16 million people who died prematurely, or before reaching 70 years of age, occur in low- and middle-income countries [1].

The situation in Morocco is as alarming as it is worldwide. The epidemiological and demographic transition results in an increase in the burden of morbidity and mortality of NCDs, in particular cancers, diabetes, cardiovascular diseases, chronic respiratory diseases and chronic kidney failure, and a notable decline in communicable diseases [2]. Therefore, they are a public health problem and are increasingly burdensome because of their health, economic and psychosocial consequences.

Concerning diabetes, the WHO estimates that 422 million adults were living with diabetes in 2014. Diabetes is also the cause of 1.5 million deaths each year, 80% of which occur in low- and middle-income countries. It will become the leading cause of death by 2030 [3]. Like many

countries around the world, Morocco is experiencing a rapid increase in the prevalence of diabetes. Indeed, for only seven years, diabetes increased by 45.4%. Its prevalence increased from 3.3% in 2011 to 4.8% in 2018. In addition, 7.1% of the adult population (18 years and older) contracted the disease in 2018 [4]. As a major public health problem, diabetes was listed as a priority in the 2025 Health Plan and the National Strategy for the Prevention and Control of NCDs 2019-2029.

Certainly, the National Survey on Population and Family Health (NSPFH-2018) and the National Survey on Common Risk Factors for NCDs, STEPS 2017-2018 are rich sources of information on diabetes and other chronic diseases at the national and regional levels. However, at the provincial level, information remains almost absent, except for a few sporadic studies on specific populations such as cardiovascular risk estimation among consultants at primary health care facilities in Settat during 2018 [5]; among women of reproductive age (20-49 years) in Oujda during 2013 [6] or among hospital staff in Meknes during 2010 [7].

Thus, existing literature was based on small and specific populations. Therefore, the generalization of the results on the city's population will lead to erroneous and biased

---

* Corresponding author: abdelghani.asraoui@uit.ac.ma

conclusions. We found that the study of the prevalence and socio-demographic determinants of chronic diseases among the total population of a city in Morocco has never been conducted. As a result, we were required to conduct a statistical survey, on a representative sample of the total population of Salé, during February 2021, to estimate the prevalence of chronic diseases and study some socio-demographic determinants of those affected. The choice of this city was established because it belongs to the Rabat-Salé-Kénitra region, which has one of the highest proportions of people with at least one chronic disease in Morocco (23.5%, 5.0% were diabetics, 7.3% were contracting Hypertension) [4].

Therefore, this study was a contribution whose main objectives were: 1) to estimate the prevalence of diabetes in Salé; 2) to study some socio-demographic predictors of the risk of contracting diabetes. In addition, we intend to contribute to the enrichment of knowledge on the extent of diabetes and the socio-demographic characteristics of diabetics. We also believe that we can provide decision-makers, in this city, with clear and reliable information to help them make decisions and implement plans and programs to combat NCDs, including diabetes.

2 Methods

2.1 Household survey

It was a statistical household survey, via a questionnaire, carried out during February 2021, In Salé. The questionnaire was developed based on the literature (especially the NSPFH-2018 questionnaire) and the terms of reference used to construct items. It was divided into three sections; the first section referred to socio-demographic data, the second section to collect information on chronic diseases and the third section to measure household well-being.

The survey was carried out on a sample of households selected according to a two-stage stratified probabilistic sampling plan (first stage: cluster selection, second stage: household selection).

To determine the initial sample size n_0, the following formula was used, normality of the population distribution was admitted [8]:

$$n_0 = \frac{N.z^2.p(1-p)}{(N.\varepsilon^2 + z^2.p(1-p)).\bar{H}} \qquad (1)$$

With:

n_0: initial sample size;
N: total population in Salé;
$\bar{H}$: average household size in terms of number of persons;
z: set at 1.96 for a 95% confidence interval;
p: proportion of the population suffering from at least one chronic disease;
ε: margin of error allowed.

Once the initial sample size was calculated, it was then adjusted, based on population size, sampling design effect and response rate [8]. Thus, taking into account the population size N, the formula for calculating the sample size became:

$$n_1 = n_0 \frac{N}{N + n_0} \qquad (2)$$

Since the sample design was not a simple random, the sample size is adjusted to take into account the effect of the sampling design *Deff* by applying the formula:

$$n_2 = n_1.Deff \qquad (3)$$

Finally, it was necessary to consider **r** the response rate. Hence, the final sample size was:

$$n = \frac{n_2}{r} \qquad (4)$$

The sample size of the households to be selected was 444 households, with:

$N = 1,120,186$ [9];
$\bar{H} = 4.3$: average household size in the Rabat-Salé-Kénitra region [4];
$z = 1.96$: for a 95% confidence interval;
$p = 23.5\%$: proportion of the population suffering from at least one chronic disease in the Rabat-Salé-Kénitra region [4];
$\varepsilon = 2\%$: margin of error allowed;
$Deff = 1.109$: calculated under the NSPHF-2018 base [4];
$r = 98.9\%$: household response rate [4].

The sample size was determined in such a way as to remedy, as far as possible, sampling errors arising from the fact that data collection concerns only a part of the population. However, another type of error can affect the survey, which is observation error that results from the risks associated with the observation itself. It depends on the rate of coaching, the quality and training of the interviewers, the clarity of the questionnaire, etc. [10]. For greater accuracy of results, we have increased the sample size by 10% [11]. Thus, the final number of households to be surveyed was 444*(1+10%) = 488 households, distributed proportionately between urban and rural areas, 93% in urban areas (452 households) and 7% in rural areas (36 households) [12].

2.2 Inclusion probabilities

2.2.1 Probability of cluster inclusion

The selection of clusters was carried out inside of each considered strata. The probability of inclusion of G_{dhi} (Cluster "i" of the stratum "h" of the Commune/Arrondissement "d" in the sample of Clusters of size n_{dh}) was given by:

$$P_{dhi} = n_{dh} * (N_{dhi} / N_{dh}) \qquad (5)$$

With:

$N_{dh} = \Sigma i \, N_{dhi}$, N_{dhi}: means the size of the G_{dhi}
N_{dh}: means the size of the stratum "h" of the Commune/Arrondissement "d"

2.2.2 Probability of household inclusion

A systematic proportional probability sampling of households in each of the clusters considered (urban and rural areas) was conducted.

The probability of inclusion of the household in the sample was given by:

$$P_{dhij} = (P_{dhi}) * (M / S_{dhi}) \qquad (6)$$

With M = Number of selected households and S_{dhi} = total number of households in the cluster G_{dhi}.

2.3 A predictive model of the risk of contracting diabetes

In this paper, we measured the prevalence of diabetes in Salé during 2021 (the significance of the observed intergroup differences was verified using the famous Khi-2 homogeneity test). Using a logistic regression model, we also tried to determine the relationship between certain socio-demographic and economic variables and the probability of contracting diabetes.

2.3.1 Construction of the model

We performed a model to identify the variables that influence diabetes presence and most effectively predict the probability of contracting the disease (DIABETE; 1=yes, 0=no). So, we introduced the following variables [13, 14]: age (AGE), gender (SEXE; 1=male, 0=female), areas of residence (MILIEU, 1=urban, 0=rural), wealth index (QBE;1=poorest, 2=second, 3=medium, 4=fourth, 5=richest), educational level (EDUCATION; 1=none, 2=primary, 3=middle school, 4=secondary, 5=higher), marital status (ETAT; 1=single, 2=married, 3=widowed, 4=divorced), economic activity (ACTIVITE; 1=working, 2=unemployed, 3=retired, 4=student, 5=housewife, 6=unable to work), average monthly household income in MAD (REVENU; 1= less than 3,000, 2=3,000-4,999, 3=5,000-9,999, 4=10,000+), smoking (TABAGIMSE; 1=smoking, 2=former smoker, 3=non-smoker) and being or not head of the household (CHEF; 1=head of household, 0=no). Therefore, the result obtained from logistic regression is between 0 and 1. If the value is close to 0, the probability of reaching the disease is low, whereas if the value is close to 1, the probability is high [15]. Logistic regression coefficients are admitted when the predicted values are located as close as possible to the observed values [16]. Thus, the model to be tested was:

$$P = \frac{1}{1+e^{-X}} \qquad (7)$$

Where:
$X = b_0 + b_1*AGE + b_2*SEXE + b_3*MILIEU + b_4*QBE + b_5*INSTRUCTION + b_6*ETAT + b_7*ACTIVITE + b_8*REVENU + b_9*TABAGIME + b_{10}*CHEF + \varepsilon$;
P: is the probability of contracting diabetes;
e: is the basis of natural logarithms;
b_0 to b_{10}: represent the parameters of the model to be estimated;
ε: is the model specification error.

2.3.2 Assessing the model

The fit quality of the model was assessed using the log-likelihood [15]:

$$LL = \sum_{i=1}^{N}[Di * ln(P(Di)) + (1 - Di) * ln(1 - P(Di))] \qquad (8)$$

Where: $P(Di)$ is the probability that diabetes occurs for the ith individual.
Cox and Snell R^2CS [17] and Nagelkerke R^2N [18] were also used to assess the model fit:

$$R^2CS = 1 - e^{[-\frac{2}{n}(LL\,(model)-\,LL\,(base))]} \qquad (9)$$

$$R^2N = \frac{R^2CS}{1-e^{[\frac{2}{n}(LL\,(base))]}} \qquad (10)$$

Where n is the sample size.
To measure the contribution of each predictor, we used the Wald statistic [19], which is the equivalent of the t-test in linear regression.

$$Wald = \frac{b}{SE(b)} \qquad (11)$$

Where $SE(b)$ is the standard error of the coefficient b.

2.4 Data analysis

Data entry was conducted using the Census and Survey Processing System (CSPro version 7.1, 2018). Once cleared, the data were analyzed using Statistical Package for Social Sciences software (SPSS version 26.0, 2019) and Microsoft Excel 2016.

2.5 Ethical approval

The survey was ethically approved by the Center for Doctoral Studies of Ibn Tofail University. Data collection from households was «verbally» authorized by the Prefecture of Salé. Interviewees were informed about the purpose of the survey as well as the purpose of the results. The right to refuse to participate or to withdraw from the investigation, at any time, anonymity, confidentiality and data protection have all been guaranteed.

3 Results

3.1 Main socio-demographic characteristics

In total, out of the 488 households selected for this survey, 466 took part (N=1,868 individuals). Thus, the household response rate was 466/488=95.5%. The average age of the population in Salé was 32.63 (±0.95) years. The 18-59 age group accounted for 58.3% of the studied population. The sex ratio H/F was 0.94 and 92.3% of individuals lived in urban areas. 38.0% belonged to the richest quantile. For subjects aged 18 and over (N=1,308), 24.3% had no education, 61.7% were married, 15.4% were unemployed, 13.8% belonged to households with an average monthly income less than 3,000 DH, 14.6% were smokers at the moment of the survey and 35,6% were household heads (Table 1.).

Table 1. Main socio-demographic characteristics

		Number	%
Age group (years)	< 18	560	30.0
	18-59	1089	58.3
	60 & +	220	11.8
Gender	Male	907	48.5
	Female	962	51.5
Areas of residence	Urban	1724	92.3
	Rural	144	7.7
Wealth index	Poorest	56	3.0
	Second	91	4.9
	Medium	375	20.1
	Fourth	637	34.1
	Richest	709	38.0
For individuals aged 18 and over (N = 1308)			
Educational level	None	318	24.3

	Primary	235	17.9
	Middle School	285	21.8
	Secondary	251	19.2
	Higher	219	16.8
Marital status	Single	393	30.0
	Married	808	61.7
	Widower	71	5.4
	Divorced	37	2.8
Economic activity	Working	545	41.6
	Unemployed	202	15.4
	Retirement	89	6.8
	Student	89	6.8
	Housewife	364	27.8
	Unable to work	20	1.5
Average household income	<3,000	180	13.8
(MAD / Month)	3,000-4,999	428	32.7
	5,000-9,999	396	30.2
	10,000 & +	304	23.3
Smoking	Smoking	191	14.6
	Former smoker	113	8.7
	Non-smoker	1004	76.7
Head of household	Yes	466	35.6
	No	842	64.4
Total		**1868**	**100.0**

3.2 Prevalence of diabetes

The results (Table 2.) revealed that 5.5% (95% CI 4.5%, 6.6%) of the population in Salé had diabetes during 2021 (the disease was confirmed by a doctor for all cases). 89.9% of them were on regular treatment for the disease. Based on age, we found that among youth under the age of 18, the prevalence of diabetes was only 0.1%. The disease affected 3.9% of people aged 18-59 and 26.9% of those aged 60 and over. Subsequently, we limited the analysis to subjects aged 18 and older (N=1,308 individuals).

Table 2. Prevalence of diabetes by age group, Salé 2021

Age group (years)	(%) diabetes	95% CI		(%) confirmed by a doctor	(%) following regular treatment	Number
		Lower bound	Upper bound			
<18	0.1	0.0	0.8	100	100.0	5,60
18-59	3.9	2.9	5.2	100	89.2	1,089
60 & +	26.9	21.3	32.9	100	90.2	220
Total	**5.5**	**4.5**	**6.6**	**100**	**89.9**	**1,868**

Table 3. shows that 7.8% of the population aged 18 and over had diabetes. The proportion of diabetic women was 8.2% compared with 7.3% for men. The disease affected 7.9% of people living in urban areas and 6.5% of those in rural areas. The prevalence of diabetics increases with the household socio-economic level. It rose from 0.9% among the poorest individuals to 8.5% for the richest. The results also showed that the disease affected 15.0% of individuals with no educational attainment and 7.9% of those with primary education. According to marital status, diabetes was present among 27.4% of widows and 9.1% of married subjects. Concerning economic activity, we found that 25.7% of retirees had diabetes, while the disease was absent among students (0.0%). In addition, diabetes was present in 3.8% of individuals belonging to households whose average monthly income was less than 3,000 DH. This proportion ranged from 9.2% of subjects with average household income between 3,000 and 4,999 DH. Otherwise, 13.8% of former smokers reported having diabetes compared to 2.9% of current smokers and 8.0% of non-smokers. Similarly, diabetes was observed at 12.5% among heads of household.

Table 3. Distribution (in %) of diabetics by selected socio-demographic variables, Salé 2021

		(%) diabetes	P-value	Standard error	95% confidence interval		(%) confirmed by a doctor	(%) following regular treatment	Total workforce
					Lower bound	Upper bound			
Gender	Male	7.3	0.479	1.04	5.5	9.6	100	86.1	626
	Female	8.2		1.05	6.3	10.4	100	92.9	683
Areas of residence	Urban	7.9	0.707	0.78	6.5	9.5	100	91.2	1,207
	Rural	6.5		2.45	3.2	13.1	100	70.5	101
Wealth index	Poorest	0.9		1.36	0.0	5.3	100	100.0	46
	Second	4.2		2.58	0.7	10.3	100	100.0	60
	Medium	7.5	0.213	1.64	4.7	11.1	100	90.0	257
	Fourth	8.3		1.28	6.0	11.0	100	89.6	462
	Richest	8.5		1.27	6.2	11.2	100	89.2	484
Educational level	None	15.0		2.00	11.5	19.3	100	90.3	318
	Primary	7.9		1.75	4.8	11.6	100	96.6	235
	Middle School	5.8	0.001	1.38	3.4	8.7	100	80.7	285
	Secondary	4.8		1.35	2.6	8.0	100	84.2	251
	Higher	3.1		1.18	1.4	6.2	100	100.0	219
Marital status	Single	1.4		0.59	0.6	3.1	100	100.0	393
	Married	9.1	0.001	1.01	7.3	11.3	100	86.0	808
	Widower	27.4		5.29	18.7	39.3	100	100.0	71
	Divorced	7.5		4.32	2.3	20.1	100	100.0	37
Economic activity	Working	4.4		0.88	2.9	6.4	100	90.3	545
	Unemployed	6.0		1.68	3.3	9.8	100	94.8	202
	Retirement	25.7	0.001	4.63	17.6	35.6	100	85.2	89
	Student	0.0		-	-	-	-	-	89
	Housewife	10.6		1.61	7.8	14.2	100	89.6	364
	Unable to work	20.2		8.98	7.2	40.8	100	100.0	20
Average household income (MAD / Month)	<3,000	3.8		1.42	1.8	7.5	100	90.8	180
	3,000-4,999	9.2	0.442	1.40	6.7	12.1	100	90.3	428
	5,000-9,999	7.3		1.31	5.1	10.2	100	92.1	396
	10,000 & +	8.7		1.62	6.1	12.5	100	86.4	304
Smoking	Smoking	2.9		1.21	1.3	6.4	100	53.9	191
	Former smoker	13.8	0.002	3.25	8.7	21.5	100	89.2	113
	Non-smoker	8.0		0.86	6.4	9.8	100	92.4	1,004
Head of household	Yes	12.5	0.001	1.53	9.7	15.7	100	89.1	466
	No	5.1		0.76	3.8	6.8	100	90.8	842
Total		**7.8**	**-**	**0.74**	**6.4**	**9.3**	**100**	**89.8**	**1,308**

3.3 Socio-demographic determinants of diabetics

3.3.1 Basic model (null model)

The main objective was to determine the socio-demographic variables explaining the presence or not of diabetes among individuals aged 18 and over in Salé during 2021. The model used was defined in equation *(7)*. The basic model (Table 4.) did not include any predictors of diabetes risk. It made it possible to determine the chances (a priori) that an individual, taken at random, had diabetes. It included only the constant b_0, which equals -2.475. Hence: $(Di) = \frac{1}{1+e^{2.475}} = 7.8\%$.

Table 4. Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step0	Constant	-2.475	.103	573.893	1	.000	.084

Table 5. represents the values of the Score statistic for each predictor variable outside the null model. At this point, we found that age in years, educational level, marital status, economic activity, smoking, and whether or not being head of the household were variables that could significantly influence the risk of developing diabetes (p-value<0.05). Furthermore, gender, areas of residence, wealth index and average monthly household income did not seem to significantly affect the risk of reaching the disease (p-value>0.05).

Table 5. Variables not in the Equation

			Score	df	Sig.
Step0	Variables	Age	125,613	1	0.000
		Gender	0.385	1	0.535
		Areas of residence	0.259	1	0.611
		Wealth index	4.671	4	0.323
		Educational level	34,635	4	0.000
		Marital status	62,691	3	0.000
		Economic activity	65,254	5	0.000
		Average household income (MAD / month)	5.784	3	0.123
		Smoking	12.2	2	0.002
		Head of household	22,737	1	0.000
	Overall Statistics		164,943	25	0.000

3.3.2 The predictive model

In the logistic regression model, we introduced all predictors of diabetes risk. It allowed determining the chances (a posteriori) that an individual, taken at random, had diabetes according to its socio-demographic characteristics. Based on results (Table 6.), we concluded that, collectively, the predictors affected significantly the risk of developing diabetes (p-value <0.0001).

Table 6. Omnibus tests of model coefficients

		Chi-square	df	Sig.
Step 1	Step	160.187	25	.000
	Block	160.187	25	.000
	Model	160.187	25	.000

Values of the R^2CS (0.115) and R^2N (0.274) coefficients showed that, collectively, variables included in the model were able to explain between 11.5% and 27.4% of the variability in the risk of developing diabetes (Table 7.).

Table 7. Model Summary

Step1	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
	554.019	.115	.274

The classification table (Table 8.) shows that the model predicts the risk of diabetes correctly with an accuracy of 91.7%.

Table 8. Classification Table[a]

Step1			Predicted		
	Observed		Diabetic		Percentage Correct
			No	Yes	
Diabetic	Yes		1197	10	99.2
	No		99	2	2.3
Overall Percentage					**91.7**

a. The cut value is .500

Table 9. represents both the estimated parameters of logistic regression, the standard error, the associated Wald statistic and the level of significance (p-value=Sig.). As a result, age, educational attainment, economic activity, average household income, smoking and being head of the household had a significant impact on the risk of developing diabetes (p-value<0.05). On the other hand, sex, areas of residence, wealth index and marital status had no significant influence on diabetes (p-value>0.05).

Table 9. Variables in the Equation

Step1[a]	B	SE	Wald	df	Sig.
Age	0.06	0.011	28.372	1	0.000
Gender	0.46	0.431	1.144	1	0.285
Areas of residence	0.62	0.57	1.181	1	0.277
Wealth index	1.85	1,864	1.062	4	0.499
Educational level	-0.25	0.414	0.628	4	0.002
Marital status	0.56	0.714	0.882	3	0.445
Economic activity	2.61	0.849	9.473	5	0.002
Average household income (MAD / month)	2.55	1.155	9.029	3	0.029
Smoking	0.98	0.511	3,711	2	0.044
Head of household	0.58	0.359	2,606	1	0.106
Constant	-12.14	2.134	32.355	1	0.000

a. Variable (s) entered on step 1: Age, Gender, Areas of residence, Wealth index, Educational level, Marital status, Economic activity, Average household income (MAD / Month), Smoking, Head of household.

The model performed was as follows: $P = \frac{1}{1+e^{-X}}$ Where P is the probability that diabetes occurs for an individual and X is defined by:
X= -12.14 + 0.06*AGE + 0.46*SEXE + 0.62*MILIEU + 1.85*QBE - 0.25*INSTRUCTION + 0.56*ETAT + 2.61*ACTIVITE + 2.55*REVENU +0.98*TABAGISME + 0.58*CHEF.

4 Discussions

In this paper, we determined the prevalence of diabetes and studied some socio-demographic factors that

influence significantly the risk of contracting the disease in Salé during 2021.

The results revealed a prevalence of diabetes of 5.5% (95% CI 4.5%, 6.6%). This proportion exceeded the national average recorded in 2018 (4.8%) and that of the Rabat-Salé Kénitra region (5.0%) [4]. The results also showed that 7.8% of the population aged 18 and over had diabetes. The national prevalence of diabetes was 7.1% for the same age group, according to the NSPHF-2018 [4].

The predictive model used to predict the probability of diabetes among the population aged 18 years and above contained age, gender, areas of residence, wealth index, educational level, marital status, economic activity, average monthly household income, smoking and being or not the head of the household. These variables predicted, collectively, the probability of developing diabetes with a success rate of 91.7%. They also collectively explained between 11.5% and 27.4% of the variability in risk of contracting the disease.

The analysis showed that diabetes was strongly associated with age (p-value<0.05). The prevalence of the disease was much higher among older subjects. Indeed, 26.9% of people aged 60 and over was suffering from diabetes. This proportion exceeded the national prevalence which was 20.0% in 2018 [4] for the same age group. The performed logistic model confirmed this close relationship between age and the risk of contracting diabetes. It revealed that the risk of developing the disease increases significantly with age (p-value<0.05 and coefficient b1=0.06>0). It was consistent with the results of several previous studies that have shown that the probability of contracting diabetes increases significantly with age [13, 20, 21].

Our study has also shown that education level affects significantly the risk of developing diabetes. Indeed, the prevalence of diabetes decreases significantly when the individual educational level increases (p-value<0.05). It rose from 15.0% for individuals (18 years and older) with no educational attainment to 3.1% for those with attained higher education. By comparison, according to the NSPHF-2018, diabetes affected 10.8% of individuals (18 years and older) with no education [4]. The relationship between diabetes and educational attainment was also revealed by Sacerdote et al. in 2012 [22] and Borrell et al. in 2006 [23]. The inverse association between diabetes and education level found in our study has also been revealed in other studies [13, 24, 25]. This may be due to a healthier lifestyle among those with higher education level [26].

The results also showed that economic activity had a significant influence on the risk of contracting diabetes (p-value=0.002<0.05). Pensioners (25.7%) those unable to work (20.2%) and housewives (10.6%) were the most affected by diabetes. The average monthly household income affected significantly the risk of developing diabetes (p-value=0.029<0.05). By way of comparison with previous studies, we found that these results were in line with the work of Phelan et al. [27] where it was noted that the level of socio-economic resources a person has or does not have, namely money, education, status, power and social connections, either helps protect one's health or promotes sickness and premature mortality [14].

Furthermore, our study has shown that smoking was a risk factor for diabetes (p-value=0.044<0.05). This is consistent with other studies, including those of Carter et al. in 2015 [28].

On the other hand, gender (p-value=0.285>0.05) and areas of residence (p-value=0.277>0.05) had no significant influence on the risk of developing diabetes. Although we found that the prevalence of diabetes did not differ significantly between male and female and between urban and rural areas (p-value=0.479>0.05 and p-value=0.707 >0.05 respectively). Gale and Gillespi (2001) found that women of non-European origin and men of European origin were more likely to develop diabetes [29]. According to Aregbesola et al. (2017), the risk of diabetes is higher among men [30]. The risk of developing diabetes was higher among Caribbean women, according to another study conducted in 2015 [31]. In addition, another study, conducted in rural northern China in 2019, found a higher prevalence of diabetes among older women than among men [32].

Similarly, wealth index and marital status and whether being or not the head of the household did not affect significantly the risk of developing diabetes among the population aged 18 and over (p-value=0.499>0.05, p-value=0.445>0.05 and p-value=0.106>0.05 respectively).

5 Conclusions

Our study revealed that, during 2021, Salé had a prevalence of diabetes of 5.5% (7.8% among the population aged 18 and over). This proportion exceeded national and regional prevalence. It was also revealed that age, educational attainment, economic activity, average monthly household income, and smoking are significantly related to diabetes.

Therefore, we have pointed out that Salé faces a real health problem with several complications (retinopathy, diabetic foot, kidney failure, stroke, loss of vision and nerve damage, risk of fetal mortality and others) and expensive management. Our study calls for greater attention of decision-makers and health sector managers in Salé to come up with appropriate measures for the prevention and fight against diabetes. The insurance of easy and equitable access to basic and local diagnosis is promoted. The intervention can also include blood glucose control among diabetics by ensuring rigorous monitoring in primary healthcare facilities, promoting a healthy lifestyle (diet, physical activity, blood pressure and lipid control, tobacco control, etc.). The involvement of civil society can play a crucial role, particularly in raising public awareness of the danger of diabetes.

As limitations, our study was based on the interviewees' statements. While an exact prevalence of diabetes can only be measured through blood glucose testing. A survey using the STEPWISE approach, recommended by the WHO, can remedy this problem. In addition, sophisticated laboratory tests to distinguish the types of diabetes (type 1 and type 2) are promoted. The predictive model was performed to predict the probability of contracting diabetes based on some socio-demographic factors. However, other variables may be more significant and

serve as better predictors of the risk of contracting diabetes, namely: daily consumption of tobacco, alcohol consumption, food hygiene, physical activity and physical and biochemical measurements.

References

1. World Health Organization web site, https://www.who.int/health-topics/noncommunicable-diseases#tab=tab_1

2. Ministry of Health (Morocco) , *Enquête Nationale sur les Facteurs de Risque Communs des Maladie Non Transmissibles, STEPS, 2017-2018*, 12-13 (Rabat,2019)

3. World Health Organization, *Global Report on Diabetes* (Genève, 2016)

4. Ministry of Health (Morocco), *Enquête Nationale sur La Population et la Santé Familiale (ENPSF-2018)*, (Rabat, 2019)

5. M. Chahboune, N. El Khoudri, M. Chebabe, A. Benmoussa, N. Tachfouti, Revue Marocaine de Santé Publique, **6**, (2019).

6. E. B. Sellam, A. Bour, Médecine des maladies Métaboliques, **10**, 63-69 (2016).

7. N. Tachfouti, M. Berraho, B. Nadia et C. Nejjari, Revue Marocaine de Santé Publique, **1**, (2014).

8. Statistics Canada, *Survey Methods and Practices*, 151-175 (Ottawa, 2010)

9. Ministry of Health (Morocco), *Projections des populations cibles des programmes de santé*, (Rabat, 2017)

10. A. Kamal, *Sondage et Tests Statistiques*, 4-44 (Mohammadia, Imprimerie de Fédala, 1998).

11. W. E. Deming, Revue de statistique appliquée, **12**, 11-55 (1964)

12. HCP (Morocco), *Projections de la population et des ménages 2014-2050*, (Rabat, 2017)

13. T. Ashis, H. Md Zobayer, Scientific Reports, **10**, 1-7 (2020).

14. W. C. Cockerham, B. W. Hamby et G. R. Oates, American journal of preventive medicine, **52**, S5-S12 (2017).

15. A. Field, *Discovering Statistics Using SPSS*, 218-268 (London, SAGE Publications Ltd, 2005)

16. G. Bourbaonnais, *Econométrie*, 295-301 (Paris, Dunod, 2005)

17. D. R. Cox et E. J. Snell, *Analysis of binary data*, (Routledge, CRC press, 1989)

18. N. J. Nagelkerke, Biometrika, **78**, 691-692 (1991)

19. A. Wald, Transactions of the American Mathematical society, **54**, 426-482 (1943)

20. R. R. Kalyani, S. H. Golden et W. T. Cefalu, Diabetes Care, **40**, 440-443 (2017)

21. S. Wang, W. Ma, Z. Yuan, S. M. Wang, X. Yi, H. Jia, F.Xue, BMJ open, **6**, (2016)

22. C. Sacerdote et al., International journal of epidemiology, **41**, 1162-1173 (2012)

23. L. N. Borrell, F. J. Dallo, K. White, American journal of public health, **96**, 1637-1642 (2006)

24. X. Shang et al., PLoS One, **8**, e66479 (2013)

25. S. M. Whitaker et al., American journal of men's health, **8**, 349-356 (2014)

26. . K. O. Hajian-Tilaki, B. Heidari, Journal of Public Health, **32**, 202-209 (2010)

27. J. C. Phelan, B. G. Link, P. Tehranifa, Journal of Health and Social Behavior, **51**, S28-S40 (2010)

28. B. D. Carter et al, New England Journal of Medicine, **372**, 631-640 (2015)

29. G. Edwin AM, K. M. Gillespie, Diabetologia, **44**, 3-15, (2001)

30. A. Aregbesola et al, Annals of Clinical Biochemistry, **54**, 113-120 (2017)

31. N. Sobers-Grannum et al., PloS One, **10**, e0126799 (2015)

32. J. Li et al., Frontiers in Endocrinology, **10**, 147 (2019)