

**ANALYSE DES MALADIES DE L'HYPERTENTION ET DU  
DIABETE PAR LA REGRESSION LOGISTIQUE  
ANALYSIS OF HYPERTENTION AND DIABETES DISEASES BY  
LOGISTICAL REGRESSION**

Mme Tassadit SEMAOUNE, MAA Université de Béjaïa (Algérie),

semaoune\_tassa@yahoo .fr

Pr .Taklit AKROUF, ENSSEA Alger (Algérie), taklitalitouche@gmail.com

*Reçu le:18/01/2025*

*Accepté le:21 /03/2025*

*Publié le:30/03/2025*

**Résumé:**

En termes de prévalence des maladies chroniques en Algérie, l'hypertension artérielle occupe la première place devant le diabète. Elle occupe également la première position des causes de décès par maladie chronique. Nous avons mené une étude dans la wilaya de Béjaïa sur 500 personnes âgées de plus de 25 ans sur une durée d'une année. Nous avons ensuite réalisé une analyse statistique à l'aide du logiciel SPSS en suivant la méthode de régression logistique, afin d'identifier les facteurs qui pourraient causer des complications chez les hypertendus et les diabétiques. Nous avons constaté qu'environ un diabétique sur deux est hypertendu. La tendance est significativement plus élevée chez les hommes que chez les femmes.

**Mots clés :** Régression, logistique, hypertension, diabète, Béjaïa.

**Jel Classification Codes:** C51, F1, F43

**Abstract:**

In terms of prevalence of chronic diseases in Algeria, high blood pressure ranks first ahead of diabetes. It also occupies the first position of causes of death by chronic disease. We carried out a study in the wilaya of Béjaïa on 500 people aged over 25 during one year. Then, we realized a statistical analysis using SPSS software and following logistic regression method, in order to identify the factors that could cause complications in hypertensive and diabetic patients. We noticed that about one in two

diabetics is hypertensive. The trend is significantly higher among men than women.

**Keywords:** Regression, logistic, hypertension, diabetes, Béjaïa.

**Jel Classification Codes:** C51, F1, F43

### **Introduction :**

Le diabète et l'hypertension sont des maladies dites chroniques. On utilise ce terme pour désigner des maladies qui ne peuvent pas être soignées assez rapidement. D'après la définition de l'Organisation Mondiale de la Santé (OMS), les maladies chroniques « *sont des affections de longue durée qui, en règle générale, évoluent lentement* ». Cela implique donc une longue prise en charge. (MAJNONI D'INTIGANO B., 2001)

En 2003, l'OMS propose d'adopter une autre définition des maladies chroniques : il s'agit de problèmes de santé qui nécessitent une prise en charge continue pendant des années, voire des décennies. D'après la même organisation, les maladies chroniques sont responsables de 63% des décès annuels, ce qui en fait la première cause de mortalité dans le monde.

Au cours des dernières décennies, le taux d'urbanisation est resté croissant, entraînant des transitions épidémiologiques, démographiques et nutritionnelles. En effet, de nombreux facteurs liés à l'industrialisation et à la modernisation des villes ont impliqué plusieurs changements dans le style de vie. (ESSOLI S., 2020)

L'urbanisation rapide, la transition vers des modes de vie à risques et la mondialisation des produits nocifs tels que le tabac, ont considérablement affecté la santé des individus. Selon l'OMS, ces transformations ont contribué à l'augmentation de plusieurs facteurs de risque, notamment une alimentation riche en sel et en matières grasses mais pauvre en fruits et légumes, un stress accru, et une baisse de l'activité physique, en raison d'activités professionnelles plus sédentaires, de l'usage excessif des écrans (ordinateurs, téléphones, télévision) et de la dépendance croissante aux transports motorisés.

L'exposition de la population algérienne à ces facteurs de risque augmente la morbidité et la mortalité liée aux maladies chroniques non transmissibles. (BRAHIMIA B., 2010)

Nous nous intéressons dans notre étude aux facteurs influençant la demande de soins de santé dans la wilaya de Béjaïa. Pour être plus précis, notre étude est axée sur les facteurs relatifs aux maladies chroniques. Nous nous sommes focalisés sur les deux maladies chroniques les plus fréquentes et qui deviennent de véritables problèmes de santé publique, à savoir le diabète et l'hypertension artérielle (HTA). Nous nous posons la question de recherche suivante : ***Comment effectuer une régression logistique afin d'identifier les facteurs de risque qui pourraient causer des complications chez les hypertendus et les diabétiques ?***

## **2. L'étude statistique**

Ne disposant pas d'une base de sondage sur les malades de la wilaya de Béjaïa, nous avons utilisé une méthode d'échantillonnage non probabiliste. La méthode retenue est l'Échantillonnage par quotas. Elle consiste à évaluer les caractéristiques de la population cible et à les reproduire dans l'échantillon. La limite principale de cette méthode est liée à son caractère non probabiliste. (ARDILLY P., 2004)

La taille de l'échantillon retenu pour notre étude est de 500 malades répartis comme suit : 76% d'hypertendus et 24% de diabétiques. Les patients retenus se sont tous présentés au CHU de Béjaïa en 2018, soit au service d'endocrinologie, soit au service de cardiologie. (RAPPORT, 2004)

## **3. Hypothèses de recherche**

La revue de la littérature, notamment des études déjà menées dans certains pays d'Afrique, a permis de constater que les maladies chroniques sont influencées par de nombreuses variables notamment l'âge, le sexe, le poids, etc.

Des études antérieures ont souvent indiqué qu'un âge plus avancé, le sexe féminin et une élévation de l'indice de masse corporelle constituent

d'importants prédicteurs de tension artérielle ou de diabète. De nombreux modèles de prédiction sont faits dans d'autres pays.

Cependant, leur généralisation est limitée. De plus, peu de modèles permettent d'examiner les liens entre ces maladies et l'hérédité, l'état civil du malade, l'activité physique ou le mode de vie, mais surtout aucun modèle n'a été appliqué en Algérie.

Dans le cas de notre étude qui s'intéresse aux facteurs de risque des maladies chroniques à Béjaïa, nous formulons l'hypothèse suivante : En plus de l'âge, du sexe et du poids, il existe d'autres facteurs tels que l'hérédité, l'activité physique et la situation matrimoniale, qui ont un impact sur la complication de l'hypertension et/ou du diabète. (FACTEURS, 2005)

#### **4. Définition des variables de l'étude**

Sur la base de l'hypothèse énoncée ci-dessus, nous décrivons la maladie chronique (hypertension ou diabète) comme une variable dépendante et nous utilisons sept variables indépendantes. Aussi, nous avons opté pour la méthode de régression logistique binaire.

La régression logistique binaire est un modèle utilisé pour analyser la relation entre une variable de réponse (variable dépendante) et une variable prédictive (variable indépendante). La variable de réponse est une dichotomie des données qualitatives qui prend la valeur 1 pour exprimer l'existence d'une caractéristique et 0 pour exprimer la non-existence d'une caractéristique. (THOMAS A., 2000)

#### **5. Analyse de la maladie à l'aide de la régression logistique binaire**

Pratiquement, la régression logistique permet d'estimer des coefficients de régression qui indiquent le sens (positif/négatif) et la force (0 ou 1) de l'influence d'une variable explicative sur l'appartenance à une catégorie (développer une maladie chronique ou non). (MAJNONI D'INTIGANO B., 2001)

#### **6. Facteurs de risque de complication chez les malades**

Nous allons appliquer une régression logistique de la variable "complication" chez les malades (diabétiques et hypertendus) sur différentes variables explicatives.

Pour mieux comprendre l'effet de ces variables sur "le risque de complication" ou "l'absence de risque de complication", nous effectuerons les 5 modélisations suivantes :

**a) Première modélisation**

Le modèle montre la relation entre la variable "complication" et les variables explicatives : sexe, âge, situation matrimoniale (SM), niveau d'instruction (NI), profession (PROF) et revenu (R). Le modèle possède 7 paramètres et il s'écrit sous la forme suivante :

$$y = \beta_0 + \beta_1 \text{sexe} + \beta_2 \text{âge} + \beta_3 \text{SM} + \beta_4 \text{NI} + \beta_5 \text{PROF} + \beta_6 \text{R}$$

**b) Seconde modélisation**

Le second modèle montre la relation entre la variable "complication" et les variables explicatives : âge, poids (wt), pression systolique (PAS) et taux de cholestérol (HDL). Ce modèle s'écrit comme suit :

$$y = \beta_0 + \beta_1 \text{âge} + \beta_2 \text{Wt} + \beta_3 \text{PAS} + \beta_4 \text{HDL}$$

**c) Troisième modélisation**

Le troisième modèle montre la relation entre la variable "complication" et les variables explicatives : âge, situation matrimoniale (SM) et histoire familiale du malade (hérédité). Ce modèle s'écrit comme suit :

$$y = \beta_0 + \beta_1 \text{âge} + \beta_2 \text{SM} + \beta_3 \text{hérédité}$$

**d) Quatrième modélisation**

Dans ce quatrième modèle, il n'y a que deux variables explicatives : âge et poids du malade (wt). Voici sa formule :

$$y = \beta_0 + \beta_1 \text{âge} + \beta_2 \text{wt}$$

**e) Cinquième modélisation**

Le cinquième modèle possède 7 variables : âge, situation matrimoniale (SM), hérédité, taux de cholestérol (HDL), pression

systolique (PAS), activité physique (sport) et poids (wt). Il s'écrit comme suit :

$$y = \beta_0 + \beta_1 \text{âge} + \beta_2 \text{SM} + \beta_3 \text{hérédité} + \beta_4 \text{HDL} + \beta_5 \text{PAS} + \beta_6 \text{sport} + \beta_7 \text{wt}$$

Il faut noter que seules les variables statistiquement différentes de zéro seront retenues pour la formulation des cinq modèles. (THOMAS A., 2000)

## 7. Application de la régression logistique binaire

Au vu des résultats présentés dans le tableau en annexe, les variables qui influencent positivement et fortement la complication des maladies chroniques à l'étude sont l'âge, le poids et l'hérédité. Celle qui influence négativement le risque de complication est l'activité physique. Une relation linéaire positive existe entre la variable à expliquer (complication) et les variables explicatives : âge, poids, hérédité et sport. Les autres corrélations sont significatives mais nettement plus faibles.

### 7.1. Application au modèle 1

La régression logistique binaire (méthode descendante pas à pas) fournit la classification suivante :

**Table N°1. Classification des patients (modèle 1)**

| Observations   |                    |     | Prévisions   |     |                     |
|----------------|--------------------|-----|--------------|-----|---------------------|
|                |                    |     | Complication |     | Pourcentage correct |
|                |                    |     | Oui          | Non |                     |
| <b>Étape 3</b> | Complication       | Oui | 291          | 19  | 93,9                |
|                |                    | Non | 117          | 56  | 32,4                |
|                | Pourcentage global |     |              |     | <b>71,8</b>         |

a. La valeur de césure est ,500

**Source** : obtenu à l'aide du logiciel SPSS v22.0

Les résultats de la régression montrent qu'il n'y a pas de dépendance entre les variables « profession, revenu et niveau d'instruction du malade » et la complication d'une maladie chronique.

## *Analyse des maladies de l'hypertension et du diabète par la régression logistique*

À la 3<sup>ème</sup> itération, les variables non retenues sont : le sexe et la profession. Les résultats suivants ont été obtenus :

**Table N°2. Résultats des estimations (modèle 1)**

| Variables dans l'équation     | A                  | E.S. | Wald  | ddl | Sig.        | Exp(B) | IC pour Exp(B)<br>95% |           |
|-------------------------------|--------------------|------|-------|-----|-------------|--------|-----------------------|-----------|
|                               |                    |      |       |     |             |        | Inférieur             | Supérieur |
| <b>Age</b>                    | -,207              | ,103 | 3,998 | 1   | <b>,046</b> | ,813   | ,664                  | ,996      |
| <b>Situation matrimoniale</b> |                    |      | 6,815 | 3   | ,078        |        |                       |           |
| État (1)                      | 1,259              | ,635 | 3,927 | 1   | <b>,048</b> | 3,522  | 1,014                 | 12,234    |
| État (2)                      | 1,015 <sup>-</sup> | ,826 | 1,507 | 1   | ,220        | ,363   | ,072                  | 1,832     |
| État (3)                      | ,346               | ,271 | 1,627 | 1   | ,202        | 1,413  | ,831                  | 2,404     |
| <b>Niveau d'instruction</b>   |                    |      | 8,653 | 4   | ,070        |        |                       |           |
| Niveau (1)                    | 1,347 <sup>-</sup> | ,763 | 3,119 | 1   | ,077        | ,260   | ,058                  | 1,159     |
| Niveau (2)                    | 1,104 <sup>-</sup> | ,613 | 3,250 | 1   | ,071        | ,331   | ,100                  | 1,101     |
| Niveau (3)                    | -,907              | ,616 | 2,169 | 1   | ,141        | ,404   | ,121                  | 1,350     |
| Niveau (4)                    | -,395              | ,618 | ,407  | 1   | ,523        | ,674   | ,201                  | 2,265     |
| <b>Revenu</b>                 | ,153               | ,061 | 6,230 | 1   | <b>,013</b> | 1,165  | 1,033                 | 1,313     |
| <b>Constante</b>              | ,457               | ,812 | ,316  | 1   | ,574        | 1,579  |                       |           |

a. Variables entrées à l'étape 1 : âge, sexe, situation matrimoniale, niveau d'instruction, profession, revenu

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Les facteurs affectant la maladie chronique à l'aide de la régression logistique binaire sont l'âge, la situation matrimoniale, le niveau d'instruction et le revenu. Ces résultats sont confortés par la valeur de la statistique de Khi 2 qui est significative à 0,000.

**Table N° 3. Qualité d'ajustement global et test de spécification (modèle 1)**

| Étape | -2log-<br>vraisemblance | R-deux<br>de Cox &<br>Snell | R-deux de<br>Nagelkerke | Khi-Chi-<br>deux | ddl | Sig. |
|-------|-------------------------|-----------------------------|-------------------------|------------------|-----|------|
| 3     | 578,212 <sup>b</sup>    | ,102                        | ,140                    | 51,970           | 9   | ,000 |

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Le R-deux de Nagelkerke est de 0,140. C'est par définition une version ajustée du R-deux de Cox & Snell qui est de 0,102 et donc plus proche de la réalité. Dès lors, on peut dire que les variables explicatives contribuent à expliquer 14% de la variation dans la probabilité pour une personne malade.

La méthode descendante pas à pas permet d'identifier les déterminants concourant à augmenter les probabilités d'avoir des complications. Ces déterminants sont : l'âge, la situation matrimoniale, niveau d'instruction et le revenu. On remarque sous la colonne Wald que ces variables sont toutes significatives au seuil de 5%. Nous pouvons dire qu'elles aident à prévoir la probabilité d'avoir des complications. De plus, les coefficients estimés sont inférieurs au seuil de significativité ( $\text{sig} \leq 0,05$ ). On peut donc affirmer que ces variables augment la probabilité d'avoir des complications.

Par ailleurs, le paramètre de la variable profession est de -0,044. Il est négatif mais n'est pas significatif au seuil de 5%.

## 7.2. Application au deuxième modèle

Dans ce modèle, sont incluses des variables quantitatives, qualitatives et binaires. La complication de la maladie est binaire, le poids est introduit comme variable qualitative, les autres variables sont quantitatives (l'âge, la PAS et le taux de cholestérol).

Nous avons utilisé la régression logistique binaire pour identifier les variables qui impactent les maladies chroniques de l'hypertension et du

## *Analyse des maladies de l'hypertension et du diabète par la régression logistique*

diabète. La variable à expliquer, dont les modalités sont : 0 (n'a pas de complications) et 1 (a des complications), est à la fois qualitative et binaire.

La méthode utilisée est la méthode descendante pas à pas et le critère de choix du meilleur modèle s'est fait sur la base de rapport de vraisemblance. Cela a permis d'obtenir une bonne capacité prédictive en classifiant correctement 71% de l'échantillon comme le montre ce tableau :

**Table N° 4. Classification des patients (modèle 2)**

| Observations   |                    |     | Prévisions          |     | Pourcentage correct |
|----------------|--------------------|-----|---------------------|-----|---------------------|
|                |                    |     | Complication<br>Oui | Non |                     |
| <b>Etape 2</b> | Complication       | Oui | 257                 | 53  | 82,9                |
|                |                    | Non | 87                  | 86  | 49,7                |
|                | Pourcentage global |     |                     |     | <b>71,0</b>         |

a. La valeur de césure est ,500

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Après la 1<sup>ème</sup> itération, la conclusion suivante a été obtenue : lorsque le terme constant est inclus dans le modèle, trois variables sont significatives à la 2<sup>ème</sup> itération.

**Table N° 5. Résultats des estimations (modèle 2)**

| Variables dans l'équation  |            | A     | E.S. | Wald   | ddl | Sig.        | Exp(B) | IC pour Exp(B)<br>95% |         |
|----------------------------|------------|-------|------|--------|-----|-------------|--------|-----------------------|---------|
|                            |            |       |      |        |     |             |        |                       | Inf Sup |
| <b>Étape 2<sup>a</sup></b> | <b>Age</b> | -,372 | ,084 | 19,791 | 1   | <b>,000</b> | ,689   | ,585                  | ,812    |
|                            | <b>IMC</b> | -,589 | ,071 | 69,448 | 1   | <b>,000</b> | ,555   | ,483                  | ,637    |
|                            | <b>PAS</b> | -,242 | ,057 | 18,180 | 1   | <b>,000</b> | ,785   | ,702                  | ,877    |
|                            | Constante  | 3,628 | ,485 | 55,944 | 1   | <b>,000</b> | 37,625 |                       |         |

a. Variables entrées à l'étape 1 : age, IMC, PAS, CHL.

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Les variables explicatives de la complication sont donc : l'âge, le poids et la pression artérielle des patients. La seule variable qui n'a pas été retenue est le taux de cholestérol (HDL).

Cependant, il est connu qu'avoir un haut niveau de cholestérol dans le sang, tout comme ne pas pratiquer d'activité physique évoluent dans le même sens que la prévalence d'une maladie chronique.

**Table N° 6. Qualité d'ajustement global et test de spécification (modèle 2)**

| Étape                | -2log-<br>vraisemblance | R-deux de<br>Cox & Snell | R-deux de<br>Nagelkerke | Khi-Chi-<br>deux | ddl | Sig. |
|----------------------|-------------------------|--------------------------|-------------------------|------------------|-----|------|
| Étape 2 <sup>a</sup> | 510,904 <sup>a</sup>    | ,219                     | ,300                    | 119,279          | 3   | ,000 |

a. L'estimation a été interrompue à l'itération 5 parce que les estimations des paramètres a changé de moins de ,001.

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Le log de vraisemblance est de 510,904. Dans le deuxième ensemble de résultats, les prédicteurs sont la classe d'âge, la PAS et le poids. Tous les coefficients estimés sont statistiquement significatifs, ce qui permet de confirmer les hypothèses d'influence pour chacune des variables indépendantes. Dans la sortie ci-dessus, nous pouvons voir que les variables prédictives âge, poids et PAS sont significatives parce que leur valeur-p est de 0,000. Cependant, la valeur-p pour la variable cholestérol est supérieure au niveau alpha commun de 0,05, ce qui indique qu'elle n'est pas statistiquement significative.

### **7.3. Application au troisième modèle**

Les variables explicatives introduites sont : l'âge, la situation matrimoniale et l'histoire familiale du malade.

L'utilisation de la méthode ascendante pas à pas (2 itérations) sur la base de rapport de maximum de vraisemblance, nous a permis d'obtenir une bonne capacité prédictive. Cela classe correctement 68,7% de l'échantillon.

**Table N° 7. Classification des patients (modèle 3)**

| Observations | Prévisions |
|--------------|------------|
|--------------|------------|

# Analyse des maladies de l'hypertension et du diabète par la régression logistique

|                |                    |     | Complication |     | Pourcentage correct |
|----------------|--------------------|-----|--------------|-----|---------------------|
|                |                    |     | Oui          | Non |                     |
| <b>Etape 2</b> | Complication       | Oui | 230          | 80  | 74,2                |
|                |                    | Non | 71           | 102 | 59,0                |
|                | Pourcentage global |     |              |     | <b>68,7</b>         |

a. La valeur de césure est ,500

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Le tableau suivant donne les variables explicatives retenues par le modèle, leurs coefficients de régression ainsi que le test de Wald.

Les variables explicatives introduites dans le modèle sont : l'âge, la situation matrimoniale et l'histoire familiale du malade. A la 2e itération, la variable âge n'a pas été retenue. La constante du modèle n'est pas significative et prend une valeur positive (**0,526**). Ceci pourrait s'interpréter comme l'effet positif de la constante.

**Table N° 8. Résultats des estimations (modèle 3)**

| Variables dans l'équation  |                         | A       | E.S.      | Wald   | ddl | Sig. | Exp(B) | IC 95% |       |
|----------------------------|-------------------------|---------|-----------|--------|-----|------|--------|--------|-------|
|                            |                         |         |           |        |     |      |        | Inf    | Sup   |
| <b>Étape 2<sup>a</sup></b> | <b>Situation mat</b>    | -,560   | ,161      | 12,064 | 1   | ,001 | ,571   | ,416   | ,783  |
|                            | Famille avec Diabète    |         |           | ,020   | 2   | ,990 |        |        |       |
|                            | Famille Diabète(1)      | -21,034 | 11460,784 | ,000   | 1   | ,999 | ,000   | ,000   | .     |
|                            | Famille Diabète(2)      | -,031   | ,218      | ,020   | 1   | ,887 | ,970   | ,633   | 1,485 |
|                            | <b>Famille avec HTA</b> |         |           | 35,277 | 2   | ,000 |        |        |       |
|                            | famille HTA(1)          | ,277    | ,346      | ,643   | 1   | ,423 | 1,319  | ,670   | 2,597 |
|                            | Famille HTA(2)          | 1,303   | ,230      | 32,031 | 1   | ,000 | 3,681  | 2,344  | 5,780 |
|                            | <b>Constante</b>        | ,526    | ,546      | ,929   | 1   | ,335 | 1,693  |        |       |

a. Variables entrées à l'étape 1 : âge, situation matrimoniale, Famille Diabète, famille HTA.

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Le log de vraisemblance est de 562,825. Pour le test consistant à vérifier si toutes les pentes sont égales à 000, la valeur-p est inférieure à 0,05. Significatif au niveau 0,05, il laisse entendre qu'au moins une variable indépendante exerce une influence sur la variable dépendante.

**Table N° 9. Qualité d'ajustement global et test de spécification (modèle 3)**

| Étape | -2log-<br>vraisemblanc<br>e | R-deux de<br>Cox & Snell | R-deux de<br>Nagelkerke | Khi-Chi-<br>deux | ddl | Sig. |
|-------|-----------------------------|--------------------------|-------------------------|------------------|-----|------|
| 2     | 562,825 <sup>a</sup>        | ,130                     | ,179                    | 67,358           | 4   | ,000 |

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Les prédicteurs du modèle permettent d'expliquer entre 13% et 17,9% de la variabilité dans le risque d'avoir des complications.

#### 7.4. Application au quatrième modèle

Les variables explicatives introduites sont : l'âge (âge) et le poids du malade (wt).

Nous avons utilisé la régression logistique binaire pour identifier les variables qui déterminent le risque d'avoir une complication. La variable dépendante est toujours une variable qualitative et binaire dont les modalités sont : 0 (présence de complication) et 1 (absence de complication). La méthode utilisée est la méthode descendante pas à pas (6 itérations) et le critère de choix du meilleur modèle s'est fait sur la base de rapport de vraisemblance. Cela a permis d'obtenir une bonne capacité prédictive puisqu'il classifie correctement 88,2% de l'échantillon.

**Table N° 10. Classification des patients (modèle 4)**

| Observations   |                  |     | Prévisions   |     |                        |
|----------------|------------------|-----|--------------|-----|------------------------|
|                |                  |     | Complication |     | Pourcentage<br>correct |
|                |                  |     | Oui          | Non |                        |
| <b>Etape 1</b> | Complicatio<br>n | Oui | 273          | 37  | 88,1                   |
|                |                  | Non | 20           | 153 | 88,4                   |

## *Analyse des maladies de l'hypertension et du diabète par la régression logistique*

|                    |             |
|--------------------|-------------|
| Pourcentage global | <b>88,2</b> |
|--------------------|-------------|

a. La valeur de césure est ,500

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Les résultats sont consignés dans le tableau suivant :

**Table N° 11. Résultats des estimations (modèle 4)**

| Variables dans l'équation | A      | E.S.     | Wald    | ddl | Sig.        | Exp(B)  | IC 95% |         |
|---------------------------|--------|----------|---------|-----|-------------|---------|--------|---------|
|                           |        |          |         |     |             |         | Inf    | Sup     |
| <b>Age</b>                | -1,070 | ,164     | 42,510  | 1   | <b>,000</b> | ,343    | ,249   | ,473    |
| <b>IMC</b>                |        |          | 117,415 | 4   | <b>,000</b> |         |        |         |
| IMC(1)                    | 5,282  | ,537     | 96,744  | 1   | <b>,000</b> | 196,805 | 68,693 | 563,848 |
| IMC(2)                    | -1,207 | 1,088    | 1,231   | 1   | ,267        | ,299    | ,036   | 2,522   |
| IMC(3)                    | -      | 6107,517 | ,000    | 1   | ,998        | ,000    | ,000   | .       |
| IMC(4)                    | 5,473  | ,571     | 91,847  | 1   | <b>,000</b> | 238,177 | 77,768 | 729,460 |
| <b>Constante</b>          | ,987   | ,604     | 2,669   | 1   | ,102        | 2,682   |        |         |

a. Variables entrées à l'étape 1 : âge, IMC.

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Le surpoids comme facteur de risque de l'HTA et du diabète est véhiculé dans le pays par l'urbanisation. Il est fortement influencé par des mauvaises habitudes alimentaires (comme la surconsommation du gras) et la faible pratique de l'exercice physique. Les femmes sont les plus touchées par l'excès de poids probablement pour des raisons culturelles et la prise de pilule contraceptive.

**Table N° 12. Qualité d'ajustement global et test de spécification (modèle 4)**

| Etape | -2log-vraisemblance  | R-deux de Cox & Snell | R-deux de Nagelkerke | Khi-Chi-deux | ddl | Sig  |
|-------|----------------------|-----------------------|----------------------|--------------|-----|------|
| 1     | 245,125 <sup>a</sup> | ,549                  | ,754                 | 385,058      | 5   | ,000 |

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Le log de vraisemblance est de 385,058. Le test montre que le modèle est globalement significatif (p-valeur < 0,05). Le test d'ajustement

global montre que 75,4% des variations de la variable dépendante "complication" sont expliquées par les variations des variables indépendantes.

Dans la sortie ci-dessus, nous pouvons voir que les variables prédictives âge et poids sont significatives parce que leur valeur-p est de 0,000.

Au regard des résultats de la régression, la statistique de Wald a une valeur différente de zéro et tous les trois coefficients estimés passent le seuil de significativité. Nous pouvons rejeter l'hypothèse de nullité simultanée des coefficients et donc, au moins une des variables indépendantes explique significativement le risque d'une complication chez les malades. De plus, la valeur-p vaut 0,000 et est inférieure à 0,05, ce qui montre que le modèle est globalement significatif.

## 7.5. Application au cinquième modèle

Les variables explicatives introduites pour ce modèle sont : l'âge, la situation matrimoniale, l'hérédité, le niveau de cholestérol, la pression artérielle, l'activité physique et le poids.

L'utilisation de la méthode ascendante pas à pas (4 itérations) sur la base de rapport de maximum de vraisemblance a permis d'obtenir une bonne capacité prédictive. Elle a classifié correctement 90,7% de l'échantillon (tableau ci-dessous).

**Table N°13. Classification des patients (modèle 5)**

| Observations   |                    |     | Prévisions   |     |                     |
|----------------|--------------------|-----|--------------|-----|---------------------|
|                |                    |     | Complication |     | Pourcentage correct |
|                |                    |     | Oui          | Non |                     |
| <b>Étape 4</b> | Complication       | Oui | 282          | 28  | 91,0                |
|                |                    | Non | 17           | 156 | 90,2                |
|                | Pourcentage global |     |              |     | <b>90,7</b>         |

a. La valeur de césure est 0,500

**Source :** obtenu à l'aide du logiciel SPSS v22.0

## *Analyse des maladies de l'hypertension et du diabète par la régression logistique*

Le modèle permet de d'analyser **7 déterminants** de risque de complication chez les malades. Les coefficients de régression, le test de Wald et les écarts-type sont donnés dans le tableau 14.

Nous avons observé à la 1ère itération que la valeur-p pour les variables âge, hérédité, sport et poids (0,000) sont inférieures au niveau alpha de 0,05, ce qui indique qu'elles sont statistiquement significatives. Cependant, la valeur-p (0,1) pour la variable situation matrimoniale (0,486) est supérieure au niveau alpha commun, ce qui indique qu'elle n'est pas statistiquement significative au seuil de 5%. Voici les résultats à la 4ème itération :

**Table N° 14. Résultats des estimations (modèle 5)**

| <b>Variables dans l'équation</b> | <b>A</b> | <b>E.S.</b> | <b>Wald</b> | <b>ddl</b> | <b>Sig.</b> |
|----------------------------------|----------|-------------|-------------|------------|-------------|
| <b>Age</b>                       | +1,167   | ,212        | 30,232      | 1          | <b>,000</b> |
| <b>Situation matrimoniale</b>    | ,486     | ,295        | 2,713       | 1          | ,100        |
| <b>Famille avec HTA</b>          |          |             | 11,913      | 2          | <b>,003</b> |
| famille HTA(1)                   | ,267     | ,686        | ,151        | 1          | ,697        |
| famille HTA(2)                   | 1,400    | ,413        | 11,508      | 1          | <b>,001</b> |
| <b>Famille Diabète</b>           |          |             | 8,337       | 2          | <b>,015</b> |
| Famille Diabète(1)               | -15,933  | 10562,360   | ,000        | 1          | ,999        |
| Famille Diabète(2)               | 1,383    | ,479        | 8,337       | 1          | <b>,004</b> |
| <b>Sport(1)</b>                  | -1,992   | ,650        | 9,392       | 1          | <b>,002</b> |
| <b>IMC</b>                       |          |             | 85,208      | 4          | <b>,000</b> |
| IMC(1)                           | 6,842    | ,793        | 74,424      | 1          | <b>,000</b> |
| IMC(2)                           | -1,049   | 1,164       | ,813        | 1          | ,367        |
| IMC(3)                           | -17,468  | 5511,241    | ,000        | 1          | ,997        |
| IMC(4)                           | 7,057    | ,830        | 72,239      | 1          | <b>,000</b> |
| <b>Constante</b>                 | -,831    | ,992        | ,702        | 1          | ,402        |

**Source :** obtenu à l'aide du logiciel SPSS v22.0

Au vu des résultats présentés dans les tableaux ci-haut, certaines variables influencent positivement et d'autres négativement la présence d'une complication chez les malades. Les variables qui influencent positivement la présence d'une complication sont l'âge, l'hérédité, le poids et la PAS et celle qui l'influence négativement c'est le sport.

Les variables non retenues dans ce modèle sont la situation matrimoniale et taux de cholestérol.

- **L'âge** : La probabilité d'être malade augmente avec l'âge. Plus nous avançons en âge, plus le risque de devenir malade augmente (hospitalisé ou avoir d'autres maladies chroniques). (BODSON J. S., 2005)

- **L'hérédité**: Nos résultats (dans la population) ont montré une association positive entre la présence de l'histoire familiale de la maladie et le risque d'avoir des complications. Les personnes ayant eu dans leur famille des personnes hypertendues ou diabétiques sont donc plus à risque. Un suivi plus rapproché de ces personnes s'avère nécessaire.

- **La pression artérielle (PAS)** : La pression artérielle élevée est également un facteur déterminant de risque de complication.

- **L'activité physique** : Un individu qui ne pratique pas une activité physique régulière, développe un facteur de risque d'avoir ces maladies et augmente sa probabilité d'avoir des complications par 40,01%.

- **L'IMC** : Le poids (plus important chez les femmes dans notre échantillon) est un grand facteur de risque d'hypertension et de diabète. Il est en effet de plus en plus prouvé que les personnes obèses courent un risque plus élevé d'être atteintes de ces maladies.

Ce modèle est globalement significatif. Cependant, notre échantillon compte 77,23% d'hypertendus et 22,77% de diabétiques. Il est donc nécessaire de préciser que notre modèle classe mieux les malades hypertendus que les diabétiques.

Le test global de notre modèle est présenté dans le tableau ci-dessous :

**Table N° 15. Qualité d'ajustement global et test de spécification (modèle 5)**

|        | Khi-Chi-deux | ddl | Sig. | -2log-vraisemblance  | R-deux de Cox & Snell | R-deux de Nagelkerke |
|--------|--------------|-----|------|----------------------|-----------------------|----------------------|
| Etape4 | 427,646      | 11  | ,000 | 202,536 <sup>a</sup> | ,587                  | ,806                 |

a. Une valeur khi-deux négative indique que la valeur du khi-deux a diminué depuis l'étape précédente.

**Source** : obtenu à l'aide du logiciel SPSS v22.0

Le Khi-2 du modèle vaut 427,646 sig (0,000). Significatif au niveau 0,05, il laisse entendre qu'au moins une variable indépendante exerce une influence sur la variable dépendante. Le test de spécification montre que le modèle est globalement significatif.

Cette vérification se fait en examinant le récapitulatif du modèle. Il s'agit du R-deux de Nagelkerke, celui-ci représentant la variance expliquée par le modèle. Dans le tableau précédant, le R-deux s'élève à 0,806, ce qui est jugé satisfaisant compte tenu du caractère exploratoire. Ainsi, le modèle explique 80,6% de la variance de la variable dépendante, ici la complication de la maladie. Ensuite, le pourcentage R-deux Cox & Snell permet également de vérifier la force du modèle. Ainsi, dans le même tableau, il est à 0,587, ce qui signifie que le modèle est vrai dans 58,7%.

Les deux coefficients estimés sont statistiquement significatifs pour les maladies chroniques (diabète et HTA), ce qui permet d'écarter l'hypothèse de non influence pour chacune des variables indépendantes.

## 8. Choix du modèle de l'étude et discussion des résultats

Le choix du modèle de l'étude se fera en fonction des critères de performance précédemment définis : la significativité du modèle et le test d'ajustement global.

**Table N° 16. Choix du modèle**

| Modèle  | Significativité | R-deux de Cox & Snell | R-deux Nagelkerke | de prévision |
|---------|-----------------|-----------------------|-------------------|--------------|
| Modèle1 | 0,000           | 0,102                 | 0,14              | 71,2         |
| Modèle2 | 0,000           | 0,219                 | 0,30              | 71           |
| Modèle3 | 0,000           | 0,208                 | 0,285             | 71           |
| Modèle4 | 0,000           | 0,549                 | 0,754             | 88,2         |
| Modèle5 | 0,000           | <b>0,587</b>          | <b>0,806</b>      | <b>90,7</b>  |

**Source :** réalisé par nos soins à partir des résultats de l'étude

Sur la base du maximum de vraisemblance, le cinquième modèle est le plus pertinent. Il a permis d'obtenir une bonne capacité prédictive en classifiant correctement 90,7% de l'échantillon (tableau 16).

## **9. Conclusion :**

La présente étude montre que l'inactivité physique, l'excès de poids, l'âge et les antécédents familiaux constituent d'importants facteurs de risque de complication chez les hypertendus et les diabétiques. Le risque augmente de façon linéaire avec chaque exposition additionnelle. Comme bon nombre de ces facteurs sont modifiables, les résultats pourraient avoir des implications significatives pour les politiques de santé publique et les pratiques cliniques. D'autres études sont nécessaires pour déterminer s'il est possible de retarder l'apparition de ces maladies, voire de la prévenir grâce à des interventions précoces ciblant les facteurs de risque.

L'hypertension et le diabète, en raison de leur forte prévalence, de leur gravité et de leurs conséquences invalidantes, représentent un fardeau majeur pour la santé publique dans notre pays. L'analyse des résultats a permis de tirer les conclusions suivantes :

- L'HTA et le diabète sont les maladies les plus fréquentes dans la wilaya de Béjaïa ;
- Les hommes sont légèrement plus affectés que les femmes ;
- La prévalence de l'hypertension comme le diabète augmente avec l'âge ;
- Les personnes ayant des parents hypertendus ou diabétiques présentent un risque accru de développer de l'HTA et/ou du diabète ;
- Les complications cardiovasculaires sont plus fréquemment rapportées par les patients hypertendus ;
- Le niveau d'éducation et la profession influencent la connaissance, l'adhérence au traitement et le suivi thérapeutique ;
- La présence de plusieurs facteurs de risque augmente de façon exponentielle le risque d'avoir des complications chez les patients ;
- L'inactivité physique est un facteur de risque important dans la survenue de L'HTA et du diabète.

Ces éléments doivent être pris en compte dans les stratégies de prévention de l'hypertension artérielle et du diabète.

## **10. Liste Bibliographique:**

- ARDILLY P., 2004 : *Echantillonnage et méthodes d'enquêtes* — DUNOD, France.
- MAJNONI D'INTIGANO B., 2001 : *Économie de la santé* — PUF, France.
- BODSON J. S., 2005 : *L'analyse multivariée avec SPSS* — Presses de l'université du Québec, Canada.
- BRAHIMIA B., 2010 : *Économie de la santé, évolution et tendance des systèmes de santé OCDE / Europe de l'Est / Maghreb* — Bahaeddine éditions, Algérie.
- FACTEURS, 2005 : *Mesure des facteurs de risque des maladies non transmissibles dans deux wilayas pilotes en Algérie (approche Step Wise de l'OMS)* — Algérie.
- RAPPORT, 2004 : *Rapport préliminaire mesure des facteurs de risque des maladies non transmissibles dans deux zones pilotes (approche step wise de l'OMS)* — Algérie.
- ESSOLI S., 2020 : *L'observance thérapeutique chez les patients atteints de maladies chroniques : cas du diabète et de l'hypertension artérielle* — Marrakech.
- THOMAS A., 2000 : *Econométrie des variables qualitatives* — ISBN 2100046659 — DUNOD, France.

## **11. Annexes :**

La matrice de corrélation obtenue avec SPSS :

**Table N°1. Matrice de corrélation de PEARSON (signification bilatérale)**

| Variables    |       | Compl   | sexe   | Age    | Sport             | Héré              | IMC   |
|--------------|-------|---------|--------|--------|-------------------|-------------------|-------|
| <b>Compl</b> | Corré |         |        |        |                   |                   |       |
|              | Sig   | 1       | -,029  | ,060   | -,904**<br>(,000) | ,146**<br>(,000)  | ,049  |
| <b>Age</b>   | Corré | ,256    | ,163** | 1      | -,012             |                   |       |
|              | sig   | (,000)  | (,000) | (,000) |                   | ,008              | -,052 |
| <b>Sport</b> | Corré | -,904** | ,117** |        |                   |                   |       |
|              | Sig   | (,000)  | (,004) | -,012  | 1                 | -,147**<br>(,000) | -,028 |

|             |       |        |        |       |         |      |       |
|-------------|-------|--------|--------|-------|---------|------|-------|
| <b>Héré</b> | Corré | ,146** | -,102* | ,008  | -,147** |      |       |
|             | Sig   | (,000) | (,011) |       | (,000)  | 1    | -,008 |
| <b>IMC</b>  | Corré | ,308** |        |       | -,253** |      |       |
|             | Sig   | (,000) | ,230*  | -,019 | (,000)  | ,043 | 1     |

\*\* La corrélation est significative au niveau 0.01 (bilatéral).

\* La corrélation est significative au niveau 0.05 (bilatéral).

**Source :** obtenu à l'aide du logiciel SPSS v22.0