

## L'analyse de régression logistique

**Julie Desjardins**

*Université de Montréal*

La régression logistique se définit comme étant une technique permettant d'ajuster une surface de régression à des données lorsque la variable dépendante est dichotomique. Cette technique est utilisée pour des études ayant pour but de vérifier si des variables indépendantes peuvent prédire une variable dépendante dichotomique. Contrairement à la régression multiple et l'analyse discriminante, cette technique n'exige pas une distribution normale des prédicteurs ni l'homogénéité des variances. Différents types de régression logistique existent, possédant chacun leur procédé statistique et conduisant à l'élaboration de différents modèles théoriques. Ainsi, seront abordés les types direct, séquentiel et automatisé («stepwise»). Un exemple d'utilisation de cette technique avec le logiciel SPSS sera présenté et la procédure d'analyse des résultats y sera détaillée, notamment en ce qui a trait à l'interprétation des rapports de cote.

Logistic regression analysis allows one to predict a discrete outcome such as group membership from a set of variables that may be continuous, dichotomous, or a mix. This technic is use when we want to verify if many independant variables can predict a dependant dichotomous variable. Unlike multiple regression and discriminant function analysis, the logistic regression does not require a normal distribution of predictors neither equal variance within each group (homogeneity of variances). Different types of logistic regression exist, each one having a statistic procedure and will conduct the elaboration of differents theoretical models. In this way, will be approached the direct types, sequential and stepwise. One example of using this technic with SPSS program, will be presented and the analytical procedure of the results will then be detailed, notably concerning the interpretation of the odds ratio.

La régression logistique est une technique permettant d'ajuster une surface de régression à des données lorsque la variable dépendante est dichotomique. Il s'agit en fait de connaître les facteurs associés à un phénomène en élaborant un modèle de prédiction. La popularité de cette méthode est bien connue dans les sciences de la santé et en sciences humaines, où la variable à prédire est la présence ou l'absence d'une maladie. Par exemple, il peut s'agir d'une étude sur la dépression majeure où l'on désire connaître les facteurs la prédisant le mieux, en étudiant des variables telles que l'âge, le sexe, l'estime de soi, les relations interpersonnelles etc. (Tabachnick et Fidell, 2000).

### Postulats et principes de base

La régression logistique n'exige pas que les prédicteurs soient distribués normalement, linéaires ou qu'ils possèdent une variance égale entre chaque groupe. Toutefois, cette technique s'applique uniquement à de grands échantillons. Les prédicteurs (variables indépendantes) peuvent être des variables dichotomiques ou continues (Tabachnick et Fidell, 2000).

*Équation utilisée :*

$$\ln(\hat{Y}/1-\hat{Y}) = A + \sum B_j X_{ij} \quad (1)$$

Cette équation correspond au Log naturel de la probabilité

de faire partie d'un groupe divisée par la probabilité de ne pas faire partie du groupe. En fait, elle provient de l'équation de la régression linéaire où est ajouté le log des rapports de cote (ces derniers seront définis plus loin) (Tabachnick et Fidell, 2000).

Brièvement, le fonctionnement consiste à calculer les coefficients de régression de façon itérative. Cela signifie que le programme informatique, à partir de certaines valeurs de départ pour  $b_0$  et  $b_1$ , vérifiera si les log chances estimés sont bien ajustés aux données, corrigera les coefficients, réexaminera le bon ajustement des valeurs estimées, etc., jusqu'à ce qu'aucune correction des coefficients ne puisse atteindre un meilleur résultat (Howell, 1998).

#### **Distinctions :**

Pour les études avec une variable dépendante dichotomique, d'autres procédures peuvent s'envisager, outre la régression logistique. Cette dernière est ainsi à distinguer de l'analyse discriminante, de l'analyse de fréquences multivariées et de la régression multiple, principalement puisqu'elle pallie aux limites de celles-ci. D'abord, l'analyse discriminante consiste à distinguer deux groupes ou plus sur la base d'un ensemble de variables. Il s'avère difficile de confirmer que la régression logistique surpasse complètement l'analyse discriminante. Néanmoins, cette dernière soumet les variables indépendantes à certaines conditions d'applications restrictives de normalité qui s'avèrent parfois difficiles à respecter. Or, la régression logistique n'impose pas de telles conditions aux variables indépendantes, qui peuvent aussi bien être catégorielles que continues.

Concernant l'analyse de fréquences multivariées, les prédicteurs doivent être des variables discrètes alors que dans la régression logistique, les prédicteurs peuvent être continus, dichotomiques ou discrets. Également, tel que mentionné précédemment, la distribution de ces derniers ne doit pas nécessairement être normale et l'homogénéité des variances n'est également pas requise, contrairement à la régression multiple. Ainsi, pour cette technique, ces deux aspects s'avèrent nécessaires (Tabachnick et Fidell, 2000).

Toutefois, il est à noter que la première partie du résultat obtenue par la régression logistique est analogue à celle d'une régression multiple. Ainsi, pour ces deux techniques, il s'agit d'un test visant à déterminer si le modèle (tous les prédicteurs étant réunis) prédit la variable dépendante à des niveaux supérieurs au hasard. Plus précisément, pour la régression multiple, il s'agit d'un test F, tandis que pour la régression logistique, plusieurs tests  $X^2$  sont produits (Howell, 1998).

## **Les types de régression logistique et les stratégies de modélisation**

La régression logistique peut s'effectuer de diverses façons. Elle se divise principalement en trois types, correspondant chacun à l'élaboration de modèles particuliers.

#### ***La régression logistique directe***

D'abord, la méthode directe est utilisée lorsqu'aucune hypothèse spécifique n'est formulée concernant l'ordre ou l'importance des variables prédictrices. En fait, ce type ne tient compte de la contribution que des prédicteurs supérieurs, car chacun est évalué comme s'il entrait dernièrement dans l'équation. De ce fait, si un prédicteur est fortement corrélé avec la variable dépendante, il est possible qu'il ne démontre qu'une prédiction faible en présence des autres prédicteurs. Cette méthode est faite par défaut avec SPSS (Tabachnick et Fidell, 2000).

#### ***La régression logistique séquentielle***

Selon cette méthode, le chercheur doit spécifier l'ordre d'entrée des prédicteurs dans le modèle. Afin de déterminer le rang des prédicteurs, le chercheur se base sur le cadre théorique de la problématique et des hypothèses bien définies. Il peut s'agir par exemple d'une étude se basant sur des modèles théoriques bien établis où le chercheur désire tester si d'autres variables peuvent y être ajoutées et ainsi vérifier leur contribution. Cette méthode correspond à deux stratégies de modélisation présentées ci-bas. Avec SPSS, il est possible d'effectuer des entrées séquentielles d'un ou plusieurs prédicteurs par l'utilisation successive du mode ENTER. (Tabachnick et Fidell, 2000).

#### ***Modèle explicatif***

À des fins explicatives, ce modèle permet de confirmer ou d'infirmer les hypothèses. Il est ici possible de distinguer les interactions spécifiques entre les prédicteurs et la variable à prédire de même que les variables jugées confondantes. L'analyse des données générée par une telle étude suppose l'examen minutieux d'une association entre un facteur bien identifié et la variable dépendante. Certaines hypothèses secondaires, portant par exemple sur des interactions spécifiques ou sur le rôle modifiant de certains tiers facteurs avec la variable dépendante, peuvent se greffer à l'hypothèse centrale. Tous les autres facteurs, jugés potentiellement confondants, pourront être retenus s'ils sont jugés porteurs de tels effets. Le test statistique ici effectué correspond à son rôle plus conventionnel d'outil de décision, c'est-à-dire le rejet ou non de l'hypothèse (Bernard, P., 2003).

### *Modèle prédictif*

Ce modèle conduit à l'élaboration d'instruments de prédiction de l'événement. Sur la base d'un ensemble de facteurs, il devient alors possible de mesurer la probabilité que l'événement se produise. Ainsi, à partir d'un bassin de variables constitué, une sélection des facteurs les plus pertinents et les plus discriminants doit être effectuée. Ces facteurs serviront à l'élaboration de l'instrument de prédiction (Bernard, P., 2003).

### *Modèle descriptif*

Les études descriptives permettent de mesurer l'importance d'un phénomène et d'en tracer le profil suivant un certain nombre de variables. Il s'agit ainsi d'explorer les données et de vérifier les associations possibles conduisant à la formulation d'hypothèses. De ce fait, ce modèle tend à suggérer des hypothèses plutôt que de les confirmer étant donné le caractère exploratoire de l'étude (Bernard, P., 2003).

### **La régression logistique Automatisée («Stepwise»)**

Cette technique permet de générer des hypothèses. En effet, elle est principalement utilisée lorsque le champ de recherche est moins exploré et que la connaissance des variables prédictives possibles s'avère limitée. Ainsi, elle ne se base pas sur des hypothèses, c'est-à-dire que les prédictifs ne sont pas entrés selon un ordre prédéterminé. Cette technique fonctionne de façon à conserver les meilleurs prédictifs de l'ensemble des variables incluses dans le modèle. Cette sélection des prédictifs est déterminée uniquement par des critères statistiques. En d'autres mots, chaque variable indépendante est entrée dans le modèle une à la fois, selon le niveau de signification statistique. À ce sujet, un seuil critère de 0,10 est ici toléré afin que des variables moins importantes ne soient pas totalement exclues suite à l'importante association de prédictifs avec la variable dépendante. En fait, il s'agit ici de conserver autant d'informations que possible. Sans quoi, cette technique ne tient pas compte des prédictifs ayant une contribution plus faible étant donné la présence de meilleurs prédictifs. Comme il ne s'agit pas de conserver uniquement les meilleurs prédictifs, mais bien de comprendre l'ensemble du phénomène et tous les facteurs explicatifs, il est justifié de tolérer un seuil significatif moins sévère. Cette technique renvoie au modèle descriptif. Avec SPSS, il s'agit de sélectionner la méthode Descendante ou Ascendante pas à pas (Forward ou Backward contionnal) (Tabachnick et Fidell, 2000).

### **Limites de cette technique**

La régression logistique nécessite des échantillons de grande taille pour atteindre un bon niveau de stabilité. Un

nombre minimal de 50 participants par variable est suggéré. Également, les prédictifs choisis doivent être spécifiques à un groupe ou l'autre, puisqu'il s'agit d'une variable dichotomique. En d'autres mots, les catégories doivent être mutuellement exclusives et exhaustives, car un prédictif ne peut pas appartenir aux deux groupes à la fois. Cette technique s'avère très sensible à la multicollinéarité entre les prédictifs, celle-ci se vérifiant à l'aide d'une matrice de corrélation. Ainsi, il est nécessaire d'examiner les corrélations entre les prédictifs avant de procéder à l'élaboration du modèle. Lorsque certains prédictifs sont fortement corrélés entre eux, il est préférable d'en éliminer quelques-uns puisqu'il s'agit probablement de variables redondantes. Dans le même ordre d'idées, la régression logistique présume que les réponses des différents cas sont indépendants les uns des autres. Il est alors prétendu que chaque réponse provient de cas différents, c'est-à-dire non reliés. Ainsi, si les variables résultantes sont formées par la période de temps pendant laquelle les mesures sont prises (avant et après traitement) ou si les variables proviennent d'un groupe par couplage (chaque sujet du groupe expérimental est jumelé à un sujet du groupe contrôle), la régression logistique devient inappropriée compte tenu des erreurs probables de corrélations (Tabachnick et Fidell, 2000).

### **Utilisation avec SPSS**

#### *Mise en situation*

L'exemple suivant se base sur une étude exploratoire portant sur la médication psychotrope chez les adolescents. La régression logistique est alors utilisée puisque l'objectif de l'étude consiste à préciser dans quelle mesure les variables sont associées au recours à la médication psychotrope et peuvent prédire ce phénomène. Plusieurs questionnaires ont été utilisés afin de dresser un portrait des jeunes sous médication. Des analyses descriptives et bivariées ont été effectuées à partir des scores de ces questionnaires de manière à cibler les prédictifs potentiels. Quelques variables, dont les symptômes, le sexe et l'âge se sont avérées significatives dans les tests de moyenne et seront alors incluses dans la régression. Pour ces variables, le seuil significatif choisi se situe entre 0,05 et 0,10 puisque l'exemple a trait à une étude exploratoire. L'exemple est effectué selon la méthode automatisée («stepwise»).

#### *Éléments importants à considérer*

D'abord, tel qu'illustré dans la mise en situation, il s'avère nécessaire de vérifier préalablement que les prédictifs peuvent influencer sur la variable à prédire à l'aide de tests statistiques. Les variables significatives sont alors retenues et incluses dans le modèle. Ensuite, il est préférable

d'effectuer une matrice de corrélation afin de vérifier la multicollinéarité des variables. Ainsi, si deux variables corrélaient fortement entre elles, elles ne peuvent toutes les deux être incluses dans le modèle. Il faudra alors procéder à une sélection selon la pertinence en regard des objectifs de l'étude.

Par exemple, dans la présente mise en situation, une variable, les «troubles intériorisés», a dû être écartée du modèle puisqu'elle incluait une seconde variable du modèle, soit «l'anxiété-dépression». Il en est ainsi puisque la technique exige que les variables doivent faire partie de catégories distinctes. Il est recommandé de procéder à l'élaboration de quelques modèles à partir de l'ensemble des variables à l'étude avant d'effectuer le modèle uniquement avec les variables significatives. Ainsi, il est préférable d'examiner quel modèle peut être généré à partir de l'ensemble des variables. Il s'agit de vérifier quelles variables sont significatives dans le modèle alors qu'elles sont en présence de toutes les autres variables. Bien entendu, c'est davantage le modèle avec les variables préalablement choisies (puisque significatives dans les tests de moyenne) qui sera présenté dans l'article. Par ailleurs, il est à noter que dans la régression logistique, la signification des tests statistiques est fournie pour chacune des variables indépendantes. Ces tests de signification sont souvent basés sur les statistiques de Wald, lesquelles possèdent une distribution chi carré avec 1 degré de liberté. Lorsque la valeur absolue du coefficient de régression est large, l'erreur standard estimée est également très large, ce qui a pour conséquence de ne pas rejeter l'hypothèse nulle, alors que cela aurait dû être le cas (erreur de type 2). Comme alternative, la signification d'un rapport de cote peut être examiné en utilisant un intervalle de confiance. Si ce dernier est plus grand que 1, l'hypothèse nulle est retenue, autrement elle doit être rejetée. Également, tel que c'est le cas avec la régression linéaire, la contribution d'une variable individuelle dans la régression logistique s'avère difficile à déterminer, puisque l'apport de chacune des variables dépend de la présence des autres variables dans le modèle. Cela devient particulièrement problématique lorsque les variables indépendantes sont fortement corrélées, d'où l'importance d'effectuer une matrice de corrélation au préalable. (Morgan, Vaske, Gliner & Harmon, 2003).

### Logiciel SPSS

Il est à noter que la variable à prédire, ici l'appartenance au groupe médicamenté, doit être représentée par 1. Tandis que l'autre groupe, ici non médicamenté, doit être identifié par 0.

La syntaxe requise se présente comme suit :

### LOGISTIC REGRESSION VAR=groupe

```

/METHOD=ENTER sexe age
/METHOD=FSTEP(COND) at_anxd at_ppen at_patt
at_cdel at_inte at_ext
/CONTRAST (sexe)=Indicator
/CRITERIA PIN(.05) POUT(.10) ITERATE(20) CUT(.5)

```

La variable «groupe» représente la variable dépendante. Pour les variables indépendantes, il est possible de composer différents blocs de variables pour celles provenant de la même catégorie telles que des données socio-démographiques, des symptômes, etc. Il est à noter que de regrouper par bloc permet de traiter ceux-ci en deux étapes, ce qui s'avère avantageux lorsqu'un bloc de variables est potentiellement plus faible que l'autre. De cette façon, les blocs seront considérés comme indépendants, tout en faisant partie du même modèle. La procédure est simple, il s'agit de séparer les groupes de variables par une barre oblique (/). Ensuite, il s'avère nécessaire de préciser quelles variables sont catégorielles en utilisant le terme Contrast. Il faut également indiquer si la variable catégorielle est Première ou Dernière (ce qui dépend des hypothèses de recherche et des variables à comparer; celle que l'on indique comme étant Première est celle que l'on attend comme étant un bon prédicteur, tandis qu'indiquer Dernier se rapporte à une variable considérée comme un faible prédicteur). Pour ce faire, on précise Indicator à la variable choisie. En terminant, la méthode doit être déterminée. Pour le premier bloc, ici incluant le sexe et l'âge, il s'agit de la méthode Enter. Pour le deuxième groupe de variables est indiquée la méthode Forward conditional (en français; Ascendante pas à pas (Wald)). (voir exemple de listing SPSS en annexe 1).

### Interprétation

Pour chacun des blocs de variables, on s'attarde aux items suivants :

#### *Vérifier la force d'association du modèle*

Cette vérification se fait en examinant le récapitulatif du modèle (model summary). Il s'agit du R<sup>2</sup> de Nagelkerke, celui-ci représentant la variance expliquée par le modèle. Le domaine d'étude et les théories sous-jacentes doivent être utilisées pour juger cette variance. Dans l'exemple suivant, le R<sup>2</sup> s'élève à 0,244, ce qui est jugé satisfaisant compte tenu du caractère exploratoire et nouveau de cette étude. Ainsi, le modèle explique 24,4% de la variance de la variable dépendante, ici la prescription de médicaments. Ensuite, le pourcentage total permet également de vérifier la force du modèle. Ainsi, dans le tableau de classification, vis-à-vis le «pourcentage correct» (percentage correct) et le «pourcentage global» (overall percentage), il est indiqué 70,3%, ce qui signifie que le modèle est vrai dans 70,3% des

cas. En d'autres mots, si un adolescent présente les caractéristiques énumérées dans le modèle, il fera partie du groupe médicamenteux dans 70,3% des cas. Ainsi, le modèle classe correctement les sujets dans 70,3% des cas.

#### *Signification des prédicteurs et interprétations des rapports de cote*

Par la suite, il s'agit d'observer quelles variables ont été incluses dans l'équation et d'examiner ensuite lesquelles sont significatives dans la case Sig. Lorsque tel est le cas, c'est-à-dire que les coefficients de Wald sont significatifs, on procède à l'interprétation des rapports de cote (ou «odds ratio») qui se situent dans la case Exp(B). Il est à noter que les rapports de cote correspondent au nombre de fois d'appartenance à un groupe lorsque la valeur du prédicteur augmente de 1. Plus précisément, un rapport de cote plus grand que 1 indique une augmentation des chances de faire partie du groupe médicamenteux, tandis qu'un rapport de cote de moins de 1 diminue les probabilités d'appartenance à ce groupe.

Ainsi, un adolescent présentant des problèmes de la pensée est 1,105 fois plus susceptible de se retrouver dans le groupe médicamenteux que les jeunes ne présentant pas ce profil. Également, la variable «âge» étant significative, cela indique que d'être plus jeune (puisque le rapport de cote se trouve sous zéro, soit 0,71) augmente les chances de faire partie du groupe d'adolescents médicamenteux. À titre explicatif, si la valeur du rapport de cote s'élevait à 3 pour les problèmes d'agitation, cela traduirait qu'un adolescent présentant de tels symptômes a 3 fois de chances d'être médicamenteux. De la même façon, un rapport de cote de 15 pour des symptômes dépressifs indiquerait que la présence de ces symptômes augmente jusqu'à 10 fois la chance d'être médicamenteux, il s'agirait alors d'un excellent prédicteur de la prise de médication. Il est à noter que le rapport de cote peut également être vu comme un indice de la force de la relation. Toutefois, il ne s'agit aucunement d'un lien causal (Morgan, 2003).

Voici brièvement comment rapporter les résultats dans un article :

*Le modèle obtenu explique 24% de la variance de la prescription de médicaments. Celui-ci démontre que les adolescents plus jeunes ( $B = -0,349$ ,  $p < 0,10$ ), présentant des problèmes de la pensée ( $B = 0,1$ ,  $p < 0,05$ ) sont plus susceptibles de faire partie du groupe médicamenteux, les rapports de cote étant respectivement de 0,71 et de 1,105. De plus, il est démontré que d'être un garçon est davantage associé à la médication que d'être une fille ( $B = 0,897$ ,  $p < 0,10$ ), le rapport de cote étant de 2,451. Par ailleurs, on note qu'un jeune présentant moins de comportements délinquants, (rapport de cote de 0,957) a plus de chances de se retrouver dans le groupe médicamenteux ( $B = -0,04$ ,  $p < 0,05$ ). Notons que le seuil de signification choisi s'élève ici à 0,10*

*compte tenu du caractère exploratoire de l'étude.*

En conclusion, cette présentation sommaire de la régression logistique rappelle qu'elle constitue une excellente technique lorsqu'il s'agit de déterminer des prédicteurs d'un phénomène. Bien qu'elle compte certains postulats et qu'elle exige une interprétation rigoureuse, elle s'applique dans une multitude de recherches. La création de modèles à partir de celle-ci exige une réflexion sur la problématique de même qu'une analyse minutieuse des résultats afin de divulguer une explication juste et détaillée du phénomène à l'étude.

#### **References**

- Bernard, P.-M. (2003). *La régression logistique*. Département de médecine sociale et préventive. Université Laval.
- Howell, C.D. (1998). *Méthodes statistiques en sciences humaines*. Québec : Édition Deboek..
- Morgan, G.A., Vaske, J.J., Gliner, J.A. & Harmon, R. (2003). Logistic Regression and Discriminant Analysis : Use and Interpretation. *Journal of the American of Child and Adolescent Psychiatry*, 42 (8), 994-997.
- Tabachnick, B.G. & Fidell, L.S. (2000). *Using Multivariate Statistics*, Fourth Edition. United States of America: Allyn and Bacon.

*Received May 1, 2005*

*Accepted September 25, 2005*

## Annexe 1. Listing SPSS produit par une regression logistique

**Logistic Regression****Case Processing Summary**

Unweighted Cases <sup>a</sup>		N	Percent
Selected Cases	Included in Analysis	101	98,1
	Missing Cases	2	1,9
	Total	103	100,0
Unselected Cases		0	,0
Total		103	100,0

a. If weight is in effect, see classification table for the total number of cases.

**Dependent Variable Encoding**

Original Value	Internal Value
médicamentés	0
1,00	1

**Categorical Variables Codings**

		Frequency	Paramete (1)
Sexe:	Masculin	60	1,000
	Féminin	41	,000

**Block 0: Beginning Block****Classification Table<sup>b</sup>**

Observed		Predicted			
		groupe des sujets		Percentage Correct	
		médicamentés	1,00		
Step 0	groupe des sujets	médicamentés	57	0	100,0
		1,00	44	0	,0
Overall Percentage					56,4

a. Constant is included in the model.

b. The cut value is ,500

**Variables in the Equation**

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-,259	,201	1,664	1	,197	,772

**Variables not in the Equation**

Step	Variables	Score	df	Sig.
0	SEXE(1)	2,490	1	,115
	AGE	2,731	1	,098
Overall Statistics		5,163	2	,076

**Block 1: Method = Enter****Omnibus Tests of Model Coefficients**

Step	Chi-square	df	Sig.
Step 1 Step	5,278	2	,071
Block	5,278	2	,071
Model	5,278	2	,071

**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	133,060	,051	,068

**Classification Table<sup>a</sup>**

Observed		Predicted			
		groupe des sujets		Percentage Correct	
		médicamentés	1,00		
Step 1	groupe des sujets	médicamentés	41	16	71,9
		1,00	25	19	43,2
Overall Percentage					59,4

a. The cut value is ,500

## Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1	SEXE(1)	,667	,424	2,471	1	,116	1,949
	AGE	-,264	,161	2,681	1	,102	,768
	Constant	3,330	2,448	1,851	1	,174	27,951

a. Variable(s) entered on step 1: SEXE, AGE.

## Block 2: Method = Forward Stepwise (Conditional)

## Omnibus Tests of Model Coefficients

Step		Chi-square	df	Sig.
Step 1	Step	10,679	1	,001
	Block	10,679	1	,001
	Model	15,957	3	,001
Step 2	Step	4,289	1	,038
	Block	14,968	2	,001
	Model	20,246	4	,000

## Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	122,380	,146	,196
2	118,091	,182	,244

Classification Table<sup>a</sup>

Observed			Predicted		
			groupe des sujets		Percentage Correct
			médicamentés	1,00	
Step 1	groupe des sujets	médicamentés	47	10	82,5
		1,00	22	22	50,0
		Overall Percentage			68,3
Step 2	groupe des sujets	médicamentés	44	13	77,2
		1,00	17	27	61,4
		Overall Percentage			70,3

a. The cut value is ,500

## Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1	SEXE(1)	,973	,464	4,403	1	,036	2,647
	AGE	-,361	,177	4,165	1	,041	,697
	AT_PPEN	,087	,028	9,278	1	,002	1,091
	Constant	-,793	2,890	,075	1	,784	,453
Step 2	SEXE(1)	,897	,473	3,600	1	,058	2,451
	AGE	-,349	,180	3,737	1	,053	,706
	AT_PPEN	,100	,030	11,233	1	,001	1,105
	AT_CDEL	-,044	,022	4,011	1	,045	,957
	Constant	1,366	3,129	,191	1	,662	3,922

a. Variable(s) entered on step 1: AT\_PPEN.

b. Variable(s) entered on step 2: AT\_CDEL.

Model if Term Removed<sup>a</sup>

Variable	Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of the Change
Step 1 AT_PPEN	-66,556	10,732	1	,001
Step 2 AT_PPEN	-65,682	13,272	1	,000
AT_CDEL	-61,202	4,313	1	,038

a. Based on conditional parameter estimates

## Variables not in the Equation

Step	Variables	Score	df	Sig.
Step 1	AT_ANXD	,067	1	,796
	AT_PATT	1,533	1	,216
	AT_CDEL	4,168	1	,041
	Overall Statistics	7,315	3	,063
Step 2	AT_ANXD	,097	1	,756
	AT_PATT	3,362	1	,067
	Overall Statistics	3,363	2	,186