

Etude de Cas - Modélisation de l'efficacité du système de fermeture
artérielle en fonction des conditions de la chirurgie.

Pierre Lague - Mattéo Chopin - Victor Haguët

Mars 2022

Table des matières

1	Introduction	2
1.1	Déroulement du projet	2
2	Présentation du sujet	2
2.1	Contexte	2
2.2	Description du sujet	2
2.3	Présentation des travaux de recherche	2
2.4	Méthode d'étude	3
2.5	Les populations étudiées	3
2.6	Déroulé type d'une étude	3
2.7	Analyse des données récoltées	3
2.8	Conclusion de l'étude	3
3	Etude des données et création d'un modèle multivarié	4
3.1	Préparation des données avant les tests d'indépendance	4
3.2	Analyse univariée	4
3.2.1	Tests d'indépendance avec la variable de groupe / type de fermeture (VCS)	4
3.2.2	Tests d'indépendance à la variable Succès Technique (Réussite)	7
3.3	Analyse multivariée : régression logistique	8
4	Annexes	12

1 Introduction

Dans le cadre de certaines opérations médicales, il peut être nécessaire de faire appel à des dispositifs de fermeture artérielle. Deux types de dispositifs principaux permettent ce genre d'opérations. Les sutures directes et les dispositifs à plug.

Dans le cadre de cette étude, nous allons chercher à comparer le dispositif PROGLIDE (à suture) et le dispositif FEMOSEAL (à plug) à l'aide d'une trentaine de variables explicatives afin de vérifier l'hypothèse que les systèmes à suture seraient moins efficaces que les systèmes à plug.

A quoi la réussite d'une opération de fermeture artérielle est lié dans une opération ? Trouve-t-elle un lien avec la méthode de fermeture ?

1.1 Déroulement du projet

A partir de la base de données des systèmes, l'étude abordera les sujets suivants :

- Une étude brève de la bibliographie autour de ce sujet
- Une modélisation multivariée à partir de variables explicatives préalablement choisies
- Une analyse des résultats obtenus et explication des différents choix statistiques (tests, modèles, ...)

2 Présentation du sujet

2.1 Contexte

La rédaction de ce document rentre dans le cadre de l'UE STA 1604 - Etude de cas proposée en L3 Statistique à l'université Bretagne sud. L'UE est encadrée par Mme Béatrice GUYOMAR'CH, CHU de Nantes.

2.2 Description du sujet

Les dispositifs de fermeture artérielle sont différenciables en deux types de catégories. Les dispositifs de sutures directes et les dispositifs à plug.

Les dispositifs de sutures directes consistent à aller directement réaliser une suture sur l'artère pour la refermer. Les dispositifs à plug consistent à prendre en "sandwich" une artère afin d'exercer une pression pour favoriser sa fermeture.

Des deux types de dispositifs, il semble que ce soit ceux utilisant des plugs qui sont de plus en plus privilégiés.

L'étude porte sur deux dispositifs de fermeture artérielle. Le premier est le système PROGLIDE qui est un dispositif à suture artérielle. Le second est le système FEMOSEAL qui est un dispositif à plug.

2.3 Présentation des travaux de recherche

Des études randomisées ont été menées afin de comparer l'efficacité des systèmes de fermeture artérielle FEMOSEAL et PROGLIDE après traitement endovasculaire de l'artériopathie oblitérante des membres inférieurs.

Buts des études

Les buts de ces études se déclinent en buts principaux et secondaires. Le but principal étant :

- comparer objectivement l'efficacité des deux systèmes de fermeture.

Les buts secondaires étant :

- comparer la sécurité de l'utilisation d'un DFV à plug *versus* celle d'un DFV à suture.
- comparer les deux DFV en terme de sécurité d'une reprise précoce de la déambulation à H+5 après opération.
- comparer les deux DFV en terme de qualité de vie du patient à H+5 et à 1 mois.
- comparer sur le long terme l'efficacité économique de la prise en charge avec les DFV FEMOSEAL et les DFV PROGLIDE

Critères d'évaluation

La plupart de ces objectifs seront complétés en respectant des critères d'évaluation précis (réintervention chirurgicale, saignement au point de ponction, faux anévrisme, hématome au point de ponction empêchant la reprise de la marche, lésion nerveuse, infection).

2.4 Méthode d'étude

Les recherches effectuées présentent les caractéristiques suivantes : étude prospective (commence dans le présent et se poursuit dans le temps), étude multicentrique (les volontaires sont issus de différents centres hospitaliers), étude contrôlée (on donne le médicament ou le traitement à tester aux personnes du groupe "de traitement" tandis que le groupe témoin reçoit un placebo), étude randomisée (étude expérimentale dans laquelle un traitement (ou une intervention) est comparé à un autre traitement, à une absence de traitement ou à un placebo), étude ouverte (les participants à l'étude, les thérapeutes et le personnel impliqué savent quelle thérapie les participants à l'étude reçoivent).

2.5 Les populations étudiées

La sélection des patients repose sur des critères d'inclusion ou d'exclusion (patient majeur sans tutelle, patient présentant un besoin pour l'AOMI, patient affilié à un régime de sécu., obligations légales etc., allergies).

2.6 Déroulé type d'une étude

Chaque étude a dû établir un planning à tenir, des conditions d'évaluation, des techniques d'analyse spécifiques, des règles d'arrêt de participation de l'étude ainsi que des critères d'arrêt d'une partie ou de la totalité de la recherche.

2.7 Analyse des données récoltées

Une politique de récolte et de traitement des données est mise en place pour respecter la confidentialité du patient (identification du patient, secret médical).

Pour l'analyse des données, le critère principal de succès de l'intervention est analysé comme une variable qualitative binaire (Succès respectivement OUI ou NON). La valeur est mise à NON dès la première complication majeure. La comparaison des taux de succès sera testé avec un test du Chi2 entre les deux groupes de DFV.

Le taux de survenue de complications jusqu'à 1 mois après l'intervention sont analysés comme des critères binaires.

Le temps avant la reprise de la déambulation, le score de qualité de vie seront analysés comme des variables quantitatives. Des tests de comparaison de score moyen seront utilisés (t-test).

Enfin des modèles logistiques univariés et multivariés vont permettre d'obtenir les coefficients de régression et des critères de qualité de modèle pour savoir comment nos variables explicatives décrivent notre variable cible (succès de l'intervention).

2.8 Conclusion de l'étude

Pour une étude sur 233 patients (3 patients ont été enlevés de l'analyse pour diverses causes), 230 ont été randomisés. Parmi ces 230 patients, 113 ont été pris en charge avec un DFV FEMOSEAL et 117 ont été pris en charge avec un DFV PROGLIDE. Respectivement 106 et 107 ont été stable à 1 mois après l'intervention. Ces chiffres affirment que l'utilisation de FEMOSEAL n'est pas forcément la meilleure et qu'elle ne permet pas tout le temps une meilleure intervention.

3 Etude des données et création d'un modèle multivarié

Afin de répondre à la problématique, une préparation des données sera d'abord nécessaire. Ensuite, chacune des variables explicatives sera soumise à un test afin de savoir si celles-ci sont dépendantes à la variable aux variables VCS et succès. Enfin, une régression logistique pourra être effectuée.

3.1 Préparation des données avant les tests d'indépendance

Tout d'abord, il a été choisi de faire un formatage des données afin de faciliter la compréhension des variables. Pour les variables qualitatives, chacune d'entre elles ont été attribuées à un format spécifique. Par exemple, pour la variable sexe, le format correspondant était le suivant :

- Si sexe = 1, alors sexe = "homme"
- Si sexe = 2, alors sexe = "femme"

Ainsi, celles-ci seront plus facilement utilisables et interprétables dans la suite de cette étude.

Les variables quantitatives ont été transformées en qualitatives puis rajoutées dans notre jeu de données. Cette étape a pour but de pouvoir utiliser les données de nos variables quantitatives dans la régression logistique. Ainsi, nos nouvelles variables qualitatives seront constituées des modalités suivantes.

- "Moins de Q1" sont les individus ayant une valeur inférieure au premier quartile de sa distribution
- "Entre Q1 et Me" sont les individus ayant une valeur comprise entre le premier quartile et la médiane
- "Entre Me et Q3" sont les individus ayant une valeur comprise entre la médiane et le troisième quartile
- "Plus de Q3" sont les individus ayant une valeur supérieure au troisième quartile de sa distribution

Si les quartiles et la médiane sont égaux, on répartit la variable en seulement 2 modalités séparées par la médiane.

3.2 Analyse univariée

Maintenant que les variables explicatives ont été formatées, il est alors possible de tester la dépendance des variables explicatives avec la variable VCS qui correspond aux deux dispositifs de suture artérielle puis à la variable réussite correspondant aux succès de chaque opération.

3.2.1 Tests d'indépendance avec la variable de groupe / type de fermeture (VCS)

Variabes qualitatives

Pour les variables qualitatives, nous avons choisi de tester leur indépendance avec la variable de groupe (VCS) avec le test de khi-2. Ce test est utilisé pour tester l'indépendance de variables qualitatives et n'est utilisable que si le jeu de données étudié comporte un certain nombre de données. Ayant un jeu de données suffisamment grand, nous avons donc pu faire appel à ce test pour nos variables qualitatives.

Le test de khi-2 propose l'hypothèse H_0 "La variable étudiée n'est pas indépendante avec la variable VCS" avec un niveau de confiance à 5%. Cette hypothèse sera rejetée pour une $p - value < 0.05$.

Obs.	Variable	Prob	Indep
1	Age	0.5032	Indépendante
2	Antivitamine K	0.7570	Indépendante
3	Antécédent AVC/AIT	0.3187	Indépendante
4	Cardiopathie Ischémique	0.3265	Indépendante
5	Coté Ponction	0.0983	Indépendante
6	Diabète	0.6548	Indépendante
7	Durée de l'hospitalisation	1.0000	Indépendante
8	Durée de l'intervention (min)	0.0661	Indépendante
9	Dyslipidémie	0.5646	Indépendante
10	IMC	0.8497	Indépendante
11	Inhibiteurs ACE	1.0000	Indépendante
12	Insuffisance Cardiaque	0.0611	Indépendante
13	Lésion traitée par ballon	1.0000	Indépendante
14	Lésion traitée par stent	0.8419	Indépendante
15	Médicament Antiplaquettaire	0.8240	Indépendante
16	Nombre de localisations	0.4982	Indépendante
17	Rutherford	0.3986	Indépendante
18	Score ASA	0.4005	Indépendante
19	Sexe	0.4172	Indépendante
20	Statine	0.5352	Indépendante
21	Succès Technique	<.0001	Non indépendante
22	Tabac	0.6138	Indépendante
23	Tension Artérielle	0.2956	Indépendante
24	Type d'Anesthésie	0.7419	Indépendante

FIGURE 1 – Test de Khi-2 pour les variables quantitatives en fonction de VCS

Ainsi, toutes les variables qualitatives semblent être indépendantes avec la variable VCS sauf la variable succès qui semble partager une dépendance avec la variable VCS.

Variabes quantitatives

Pour savoir si il y a indépendance entre le type de fermeture (VCS) et les variables quantitatives, il faut d'abord choisir le test de référence. Celui ci sera paramétrique si la distribution est normale, et non paramétrique sinon. On peut formuler au préalable des hypothèses sur les histogrammes :

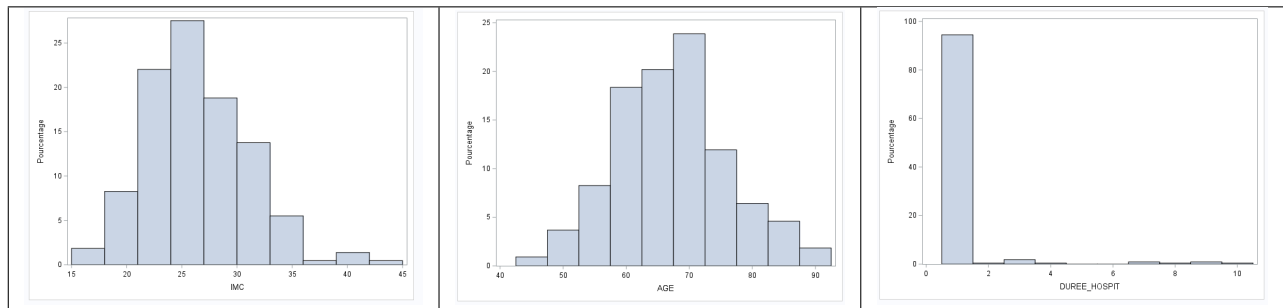


TABLE 1 – Histogrammes des variables quantitatives (IMC, AGE, DUREE_HOSPIT)

On remarque que les variables Age et IMC tendent à suivre une loi gaussienne. On peut aussi faire la même hypothèse sur la durée de l'intervention :

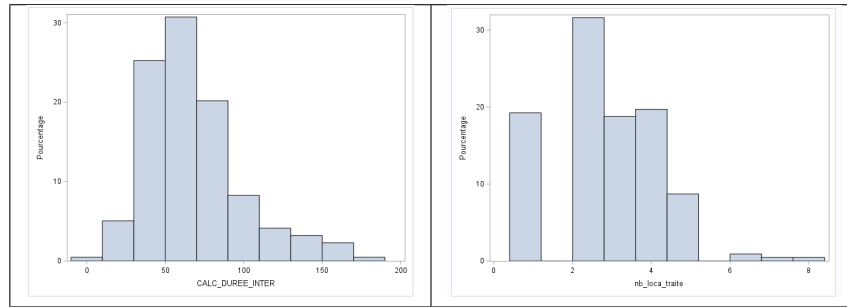


TABLE 2 – Histogrammes des variables quantitatives (CALC_DUREE_INTER, LOCA_TRAITEE)

Pour les variables quantitatives, un premier test de normalité a été réalisé, le test de Kolmogorov-Smirnov. Ce test se base sur l'hypothèse H_0 "La variable suit une loi normale" avec un niveau de confiance à 5%. Cette hypothèse sera rejetée pour une $p - value < 0.05$.

Nous obtenons les résultats suivants :

Obs.	VarName	pValue	Dist
1	IMC	0.0100	Non normale
2	AGE	0.0248	Non normale
3	DUREE_HOSPIT	0.0100	Non normale
4	CALC_DUREE_INTER	0.0100	Non normale
5	nb_loca_traite	0.0100	Non normale

FIGURE 2 – Tests de Kolmogorov-Smirnov pour les variables quantitatives

Il en ressort qu'aucune variable quantitative ne suit une distribution normale ce qui contredit les hypothèses sur les histogrammes.

Ainsi, pour pouvoir tester leur indépendance avec la variable VCS, nous avons choisi de faire appel au test non paramétrique de Wilcoxon-Mann-Whitney, qui compare les médianes pour chaque type de fermeture artérielle.

D'après le test de Wilcoxon, si la p-value obtenue est supérieure à 0.05, alors l'hypothèse H_0 serait repoussée avec H_0 : Il existe une dépendance entre la variable étudiée et la variable de groupe.

Nous avons obtenu ceci :

Obs.	Variable	Prob1	Indep
1	IMC	0.3396	Indépendante
2	Age	0.0718	Indépendante
3	Durée de l'hospitalisation	0.4805	Indépendante
4	Durée de l'intervention (min)	0.1295	Indépendante
5	Nombre de localisations	0.3230	Indépendante

FIGURE 3 – Test de Wilcoxon pour les variables quantitatives en fonction de VCS

Interprétation

D'après ces différents tests, seule la variable de succès (Succès Technique) semble partager une dépendance avec la variable groupe comme le montre les résultats à la table 11.

Ainsi, toutes les variables du jeu de données exceptées la variable Succès Technique sont indépendantes à la variable VCS.

3.2.2 Tests d'indépendance à la variable Succès Technique (Réussite)

Il peut alors être intéressant de tester l'indépendance de chaque variable à la variable Réussite. Pour cela, les mêmes tests seront réutilisés.

Variables qualitatives

Une nouvelle fois, nous avons fait appel au Test de khi-2.

Obs.	Variable	Prob	Indep
1	Age	0.0092	Non indépendante
2	Antivitamine K	0.1949	Indépendante
3	Antécédent AVC AIT	0.5312	Indépendante
4	Cardiopathie Ischémique	0.6311	Indépendante
5	Côté Ponction	0.0134	Non indépendante
6	Diabète	0.5041	Indépendante
7	Durée de l'hospitalisation	0.8295	Indépendante
8	Durée de l'intervention (min)	0.0194	Non indépendante
9	Dyslipidémie	0.1414	Indépendante
10	Groupe	<.0001	Non indépendante
11	IMC	0.5516	Indépendante
12	Inhibiteurs ACE	0.5325	Indépendante
13	Insuffisance Cardiaque	0.1535	Indépendante
14	Lésion traitée par ballon	0.1860	Indépendante
15	Lésion traitée par stent	0.8327	Indépendante
16	Médicament Antiplaquettaire	0.7802	Indépendante
17	Nombre de localisations	0.1783	Indépendante
18	Rutherford	0.3497	Indépendante
19	Score ASA	0.3263	Indépendante
20	Sexe	0.3521	Indépendante
21	Statine	0.5821	Indépendante
22	Tabac	0.6534	Indépendante
23	Tension Artérielle	0.7963	Indépendante
24	Type d'Anesthésie	0.9287	Indépendante

FIGURE 4 – Test de Khi-2 pour les variables quantitatives en fonction de VCS

Plusieurs variables partagent une dépendances avec la variable succès. Il s'agit des variables suivantes :

- Age
- Côté de la Ponction
- Durée de l'intervention
- VCS (groupes)

Variables quantitatives

De même, pour les variables quantitatives, nous avons réutilisé le test de Wilcoxon puisque ces variables ne suivent pas une distribution normale. Nous avons obtenu les résultats suivants :

Obs.	Variable	Prob1	Indep
1	IMC	0.3084	Indépendante
2	Age	0.1472	Indépendante
3	Durée de l'hospitalisation	0.3844	Indépendante
4	Durée de l'intervention (min)	0.2791	Indépendante
5	Nombre de localisations	0.1428	Indépendante

FIGURE 5 – Test de Wilcoxon pour les variables quantitatives en fonction de succès

Interprétation

D'après ces différents tests, quatre variables partagent une dépendance avec la variable succès comme le montrent les résultats à la table 12 .

3.3 Analyse multivariée : régression logistique

Il est maintenant possible de créer un modèle à partir des différentes variables explicatives. Ce modèle ne sera créé qu'à partir des variables qualitatives ce qui permettra de réaliser une régression logistique.

Tout d'abord, une phase de sélection des variables à l'aide d'une régression simple doit avoir lieu sur chaque variable explicative qualitative. On retiendra une variable explicative avec un niveau de confiance de 20%.

Au final, quatre variables seront retenues dans les modèles de régression logistique simples. Il s'agit des variables DYSPLI, CI, ballon, nb_loca avec les résultats suivants.

	P-value	Rapport de cote	Intervalle de confiance
DYSPLI	0.1426	0.648	[0.363 :1.157]
CI	0.1614	2.125	[0.740 :6.099]
ballon	0.1868	0.687	[0.394 :1.200]
nb_loca	0.1792	0.683	[0.392 :1.191]
VCS	< 0.0001	5.800	[3.113 :10.807]

TABLE 3 – Variables sélectionnées pour la régression logistique

Ces dernières sont alors utilisées pour créer un modèle multinomial. Il en ressort cette table :

Analyse des effets Type 3			
Effet	DDL	Khi-2 de Wald	Pr > khi-2
DYSPLI	1	1.8758	0.1708
CI	1	0.8171	0.3660
ballon	1	1.6040	0.2053
nb_loca	1	0.2893	0.5907
VCS	1	29.0198	<.0001

FIGURE 6 – p-value des variables retenues

La p-value de la variable nb_loca est supérieure à 20% et cela implique qu'il faut la retirer du modèle. Il faut alors refaire le modèle sans cette dernière.

Analyse des effets Type 3			
Effet	DDL	Khi-2 de Wald	Pr > khi-2
DYSPLI	1	2.1280	0.1446
CI	1	0.8098	0.3682
ballon	1	2.3918	0.1220
VCS	1	29.2481	<.0001

FIGURE 7 – P-value des valeurs retenues

La p-value de la variable CI est également trop grande. Cette variable sera donc retirée du modèle.

	P-value	Estimation du point	Intervalle de confiance
DYSPLI	0.1592	0.632	[0.334 :1.197]
ballon	0.1360	0.628	[0.341 :1.157]
VCS	< 0.0001	5.901	[3.143 :11.079]

TABLE 4 – p-value, estimation du point et intervalle de confiance des variables sélectionnés

Pour ce modèle, toutes les p-value sont inférieures à 20%. Par conséquent, les trois variables qui vont servir à établir le modèle sont *DYSPLI*, *ballon*, *VCS*.

Afin d'améliorer le modèle, on va essayer de trouver une interaction entre deux variables qui rendrait le modèle plus efficace.

Ici, les interactions ont des p-values suivantes :

Intéractions	p-value
<i>DYSPLI * ballon</i>	0.0285
<i>DYSPLI * VCS</i>	0.0154
<i>ballon * VCS</i>	0.3233

TABLE 5 – p-value des interactions

Au final, les interactions *DYSPLI * ballon* et *DYSPLI * VCS* sont intéressantes à rajouter au modèle. Le modèle final est donc Reussite = *DYSPLI* *ballon* *VCS* *DYSPLI * ballon* *DYSPLI * VCS*.

On obtient les valeurs suivantes pour chaque variable.

Analyse des valeurs estimées du maximum de vraisemblance							
Paramètre			DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept			1	0.6782	0.1695	16.0123	<.0001
DYSPLI	NON		1	-0.3546	0.1695	4.3789	0.0364
ballon	NON		1	-0.2050	0.1642	1.5596	0.2117
VCS	FEMOSEAL		1	0.8610	0.1705	25.5078	<.0001
DYSPLI*ballon	NON	NON	1	0.3703	0.1642	5.0893	0.0241
DYSPLI*VCS	NON	FEMOSEAL	1	-0.4226	0.1705	6.1462	0.0132

FIGURE 8 – Logarithme du rapport de côte et P-value de chaque variable

Sachant que la colonne estimation correspond au logarithme des rapport de côte, il est possible d'en déduire les informations suivantes :

- *DYSPLI* : Il y a 0.7014 fois plus de risque que l'intervention soit un échec si le patient est sujet à la displidémie.
- *ballon* : Il y a 0.8146 fois plus de risque que l'intervention soit un échec si la lésion est traitée par ballon.
- Il y a 2.3655 fois plus de risque que l'intervention soit un échec si on utilise la méthode FEMOSEAL plutôt que la méthode PROGLIDE.
- *DYSPLI * ballon* : Il y a 1.4481 fois plus de risque que l'opération est soit un échec si le patient est atteint de displidémie ou/et si la lésion est traitée par ballon.
- *DYSPLI * VCS* : Il y a 0.6553 fois plus de risque que l'opération est soit un échec si le patient est atteint de displidémie ou/et si on utilise la méthode FEMOSEAL plutôt que la méthode PROGLIDE.

Il est possible de vérifier la qualité d'un modèle à l'aide du test de Hosmer et de Lemeshow. Pour avoir un test de bonne qualité, il faut avoir une p-value de plus de 5%, ce qui est le cas :

Test d'adéquation de Hosmer et de Lemeshow		
khi-2	DDL	Pr > khi-2
2.2227	6	0.8981

FIGURE 9 – Test de Hosmer et Lemeshow

De même, pour qu'un modèle soit de bonne qualité, il faut que l'aire sous la courbe ROC soit supérieur à 0.7 :

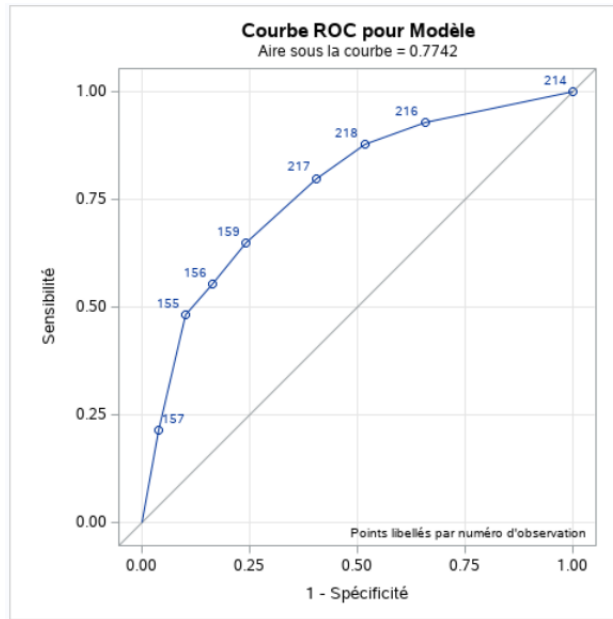


FIGURE 10 – Courbe ROC du modèle

Ici, l'aire sous la courbe ROC est de 0.77 ce qui prouve que le modèle de bonne qualité.

4 Annexes

Variable	Modalité	FEMOSEAL	PROGLIDE	P-value	Indépendance
Age	0-60 ans	46% (23)	54% (27)	0.5032	Indépendante
	61-65 ans	44% (23)	56% (30)	0.5032	
	66-71 ans	56% (28)	44% (22)	0.5032	
	72 et +	54% (35)	46% (30)	0.5032	
	Variable quantitative	68 [61,73]	65 [61,72]	0.0718	
Antivitamine K	NON	50% (104)	50% (103)	0.7570	Indépendante
	OUI	46% (5)	54% (6)	0.7570	
Antécédent AVC AIT	NON	48% (92)	52% (97)	0.3187	Indépendante
	OUI	58% (17)	42% (12)	0.3187	
Cardiopathie Ischémique	NON	52% (72)	48% (65)	0.3265	Indépendante
	OUI	46% (37)	54% (44)	0.3265	
Coté Ponction	Droite	58% (53)	42% (39)	0.0983	Indépendante
	Droite&Gauche	36% (11)	64% (19)	0.0983	
	Gauche	46% (45)	54% (51)	0.0983	
Diabète	NON	52% (70)	48% (65)	0.6548	Indépendante
	Type1	60% (3)	40% (2)	0.6548	
	Type2	46% (36)	54% (42)	0.6548	
Durée de l'hospitalisation	Moins d'un jour	50% (103)	50% (103)	1.0000	Indépendante
	Plus d'un jour	50% (6)	50% (6)	1.0000	
	Variable quantitative	1 [1,1]	1 [1,1]	0.4805	
Durée de l'intervention (min)	Entre 45 et 60	36% (18)	64% (31)	0.0661	Indépendante
	Entre 60 et 79	54% (30)	46% (26)	0.0661	
	Moins de 45	62% (33)	38% (20)	0.0661	
	Plus de 80	46% (28)	54% (32)	0.0661	
	Variable quantitative	60 [40,80]	60 [48,80]	0.1295	
Dyslipidémie	NON	48% (34)	52% (38)	0.5646	Indépendante
	OUI	52% (75)	48% (71)	0.5646	
IMC	Entre 23 et 26	54% (33)	46% (28)	0.8497	Indépendante
	Entre 26 et 30	48% (27)	52% (29)	0.8497	
	Moins de 22	46% (25)	54% (29)	0.8497	
	Plus de 30	52% (24)	48% (23)	0.8497	
	Variable quantitative	25.6 [23.4,29.1]	25.4 [22.8,29.6]	0.3396	
Inhibiteurs ACE	NON	50% (66)	50% (66)	1.0000	Indépendante
	OUI	50% (43)	50% (43)	1.0000	
Insuffisance Cardiaque	NON	52% (105)	48% (98)	0.0611	Indépendante
	OUI	26% (4)	74% (11)	0.0611	
Lésion traitée par ballon	NON	50% (57)	50% (57)	1.0000	Indépendante
	OUI	50% (52)	50% (52)	1.0000	
Lésion traitée par stent	NON	52% (15)	48% (14)	0.8419	Indépendante
	OUI	50% (94)	50% (95)	0.8419	
Médicament Antiplaquettaire	Asp&Clo	50% (26)	50% (27)	0.8240	Indépendante
	Aspirine	48% (53)	52% (56)	0.8240	
	Clopidogrel	54% (30)	46% (26)	0.8240	
Nombre de localisations	<=2	48% (53)	52% (58)	0.4982	Indépendante
	>2	52% (56)	48% (51)	0.4982	
	Variable quantitative	3 [2,4]	2 [2,4]	0.3230	
Rutherford	Asymptotique	100% (1)	0% (0)	0.3986	Indépendante
	Boitement mineur	0% (0)	100% (1)	0.3986	
	Boitement moyen	46% (27)	54% (32)	0.3986	
	Boitement sévère	56% (57)	44% (45)	0.3986	
	Douleur au repos	48% (16)	52% (18)	0.3986	
Score ASA	Perte de tissu mineur	38% (8)	62% (13)	0.3986	Indépendante
	Invalide	28% (3)	72% (8)	0.4005	
	Maladie légère	52% (44)	48% (40)	0.4005	
	Maladie sévère	52% (58)	48% (55)	0.4005	
Sexe	Sain	40% (4)	60% (6)	0.4005	Indépendante
	FEMME	56% (27)	44% (22)	0.4172	
Statine	HOMME	48% (82)	52% (87)	0.4172	Indépendante
	NON	54% (30)	46% (26)	0.5352	
Succès Technique	OUI	48% (79)	52% (83)	0.5352	Non indépendante
	NON	24% (19)	76% (60)	<.0001	
Tabac	NON	64% (90)	36% (49)	<.0001	Indépendante
	NON	48% (60)	52% (67)	0.6138	

FIGURE 11 – Dépendance de chaque variable en fonction de la variable groupe

Variable	modalités	OUI	NON	P-value	Indépendance	
Age	Variable quantitative	68 [60,73]	65 [62,71]	0.1472	Indépendante	
	0-60 ans	70% (35)	30% (15)	0.0092	Non indépendante	
	61-65 ans	46% (24)	54% (29)	0.0092		
	66-71 ans	64% (32)	36% (18)	0.0092		
Antivitamine K	NON	64% (134)	36% (73)	0.1949	Indépendante	
	OUI	46% (5)	54% (6)	0.1949		
Antécédent AVC AIT	NON	62% (119)	38% (70)	0.5312	Indépendante	
	OUI	68% (20)	32% (9)	0.5312		
Cardiopathie Ischémique	NON	64% (89)	36% (48)	0.6311	Indépendante	
	OUI	62% (50)	38% (31)	0.6311		
Coté Ponction	Droite	66% (61)	34% (31)	0.0134	Non indépendante	
	Droite&Gauche	40% (12)	60% (18)	0.0134		
	Gauche	68% (66)	32% (30)	0.0134		
Diabète	NON	66% (88)	34% (47)	0.5041	Indépendante	
	Type1	40% (2)	60% (3)	0.5041		
	Type2	62% (49)	38% (29)	0.5041		
Durée de l'hospitalisation	Moins d'un jour	64% (131)	36% (75)	0.8295	Indépendante	
	Plus d'un jour	66% (8)	34% (4)	0.8295		
	Variable quantitative	1 [1,1]	1 [1,1]	0.3844		
Durée de l'intervention (min)	Variable quantitative	60 [41,80]	59 [47,88]	0.2791	Indépendante	
	Entre 45 et 60	48% (24)	52% (25)	0.0194		Non indépendante
	Entre 60 et 79	76% (42)	26% (14)	0.0194		
	Moins de 45	72% (38)	28% (15)	0.0194		
	Plus de 80	58% (35)	42% (25)	0.0194		
Dyslipidémie	NON	56% (41)	44% (31)	0.1414	Indépendante	
	OUI	68% (98)	32% (48)	0.1414		
Groupe	FEMOSEAL	82% (90)	18% (19)	<.0001	Non indépendante	
	PROGLIDE	44% (49)	56% (60)	<.0001		
IMC	Entre 23 et 26	60% (37)	40% (24)	0.5516	Indépendante	
	Entre 26 et 30	70% (39)	30% (17)	0.5516		
	Moins de 22	66% (36)	34% (18)	0.5516		
	Plus de 30	58% (27)	42% (20)	0.5516		
	Variable quantitative	25.6 [22.8,29]	25.4 [23.4,30.4]	0.3084		
Inhibiteurs ACE	NON	62% (82)	38% (50)	0.5325	Indépendante	
	OUI	66% (57)	34% (29)	0.5325		
Insuffisance Cardiaque	NON	66% (132)	34% (71)	0.1535	Indépendante	
	OUI	46% (7)	54% (8)	0.1535		
Lésion traitée par ballon	NON	60% (68)	40% (46)	0.1860	Indépendante	
	OUI	68% (71)	32% (33)	0.1860		
Lésion traitée par stent	NON	66% (19)	34% (10)	0.8327	Indépendante	
	OUI	64% (120)	36% (69)	0.8327		
Médicament Antiplaquettaire	Asp&Clo	66% (35)	34% (18)	0.7802	Indépendante	
	Aspirine	62% (67)	38% (42)	0.7802		
	Clopidogrel	66% (37)	34% (19)	0.7802		
Nombre de localisations	<=2	60% (66)	40% (45)	0.1783	Indépendante	
	>2	68% (73)	32% (34)	0.1783		
	Variable quantitative	3 [2,4]	2 [2,4]	0.1428		
Rutherford	Asymptotique	0% (0)	100% (1)	0.3497	Indépendante	
	Boitement mineur	100% (1)	0% (0)	0.3497		
	Boitement moyen	60% (35)	40% (24)	0.3497		
	Boitement sévère	62% (64)	38% (38)	0.3497		
	Douleur au repos	64% (22)	36% (12)	0.3497		
Score ASA	Perte de tissu mineur	80% (17)	20% (4)	0.3497	Indépendante	
	Invalide	54% (6)	46% (5)	0.3263		

FIGURE 12 – Dépendance de chaque variable en fonction de la variable réussite

* Création d'une bibliothèque SAS et importation des données;

```
libname dir '/home/u48674897/EDC/DATA2';
options fmtsearch = (dir.formats);
options nofmterr;
```

* Importation de la table;

```
filename step '/home/u48674897/EDC/DATA2/step_sujet1.xlsx';
```

```
proc import datafile = step
  out = dir.step
  dbms = xlsx
  replace;
getnames = yes;
sheet = 'data';
run;
```

* Création des formats de la table;

```
proc format lib = dir.formats;
value SEXE          1 = "HOMME"          2 = "FEMME";
value OUI_NON       0 = "NON"           1 = "OUI";
value TABAC         0 = "NON"           1 = "oui <= 40 PA"      2 = "oui >= 40 PA";
value DIABETE       0 = "NON"           1 = "Type1"            2 = "Type2";
value ASAJ          1 = "Sain"          2 = "Maladie légère"  3 = "Maladie sévère"
                    4 = "Invalide";
value RUTH_J        0 = "Asymptotique"  1 = "Boitement mineur" 2 = "Boitement moyen"
                    3 = "Boitement sévère" 4 = "Douleur au repos" 5 = "Perte de tissu mineur";
value AAPJ          0 = "Non"           1 = "Aspirine"        2 = "Clopidogrel"
                    3 = "Asp&Clo";
value COTEPONCTION  1 = "Gauche"        2 = "Droite"           3 = "Droite&Gauche";
value ANESTHESIE    1 = "Générale"      2 = "Locale Régionale" 3 = "Locale&Sédation";
value GROUPS        1 = "FEMOSEAL"      2 = "PROGLIDE";
value AGE           0 = "0-60 ans"      1 = "61-65 ans"       2 = "66-71 ans"
                    3 = "72 et +";
value DUREE_H       0 = "Moins d'un jour" 1 = "Plus d'un jour";
value IMC           0 = "Moins de 22"    1 = "Entre 23 et 26"  2 = "Entre 26 et 30"
                    3 = "Plus de 30";
value DUREE_INTER   0 = "Moins de 45"    1 = "Entre 45 et 60"  2 = "Entre 60 et 79"
                    3 = "Plus de 80";
value nb_loca       0 = "<=2"           1 = ">2";
run;
```

* Restructuration de la table d'origine et application des formats;

```
data dir.step;
set dir.step;
```

```
DDN2 = input('01/'!!substr(DDN, 4, 10),ddmmyy10.);
AGE = round((DATE_RANDOMISATION-DDN2)/365,1);
IMC = input(CALC_IMCJ0, 4.);
```

```
if SEXE = "HOMME" then SEX = 1; else SEX = 2;
if HTA = "OUI" then TENSIONART = 1; else TENSIONART = 0;
if DYSLIPIDEMIE = "OUI" then DYSPLI = 1; else DYSPLI = 0;
if MALADIE_CEREBROVASC = "OUI" then AVC = 1; else AVC = 0;
if CARDIO_ISCHEMIQUE = "OUI" then CARDIO_IHD = 1; else CARDIO_IHD = 0;
if IC = "OUI" then CI = 1; else CI = 0;
if AVK_J0 = "OUI" then AVK = 1; else AVK = 0;
if IEC_J0 = "OUI" then IEC = 1; else IEC = 0;
if STATINE_J0 = "OUI" then STATINE = 1; else STATINE = 0;
if stent_traite = "OUI" then stent = 1; else stent = 0;
if ballon_traite = "OUI" then ballon = 1; else ballon = 0;
if Succes = "OUI" then Reussite = 1; else Reussite = 0;
if Group = "Femoseal" then VCS = 1; else VCS = 2;
if TABAC = "NON" then FUMEUR = 0; else if TABAC = "oui<= 40 PA" then FUMEUR = 1;
else FUMEUR = 2;
if DIABETE = "non" then DIAB = 0; else if DIABETE = "type 1" then DIAB = 1;
else DIAB = 2;
if AAPJ0 = "NON" then AAP = 0; else if AAPJ0 = "aspirine" then AAP = 1;
else if AAPJ0 = "clopidogrel" then AAP = 2;
else AAP = 3;
if COTE_PONCTION = "GAUCHE" then COTE_PONCT = 1; else if COTE_PONCTION = "DROIT" then COTE_PONCT = 2;
else COTE_PONCT = 3;
if ANESTHESIE = "générale" then ANESTH = 1; else if ANESTHESIE = "loco régionale" then ANESTH = 2;
else ANESTH = 3;
if AGE < 61 then AGE2 = 0; else if AGE < 66 and AGE >= 61 then AGE2 = 1;
else if AGE < 72 and AGE >= 66 then AGE2 = 2;
else if AGE >= 72 then AGE2 = 3;
else DUREE_HOSPIT <= 1 then DUREE_H = 0; else DUREE_H = 1;
```

```

if IMC < 23 then IMC2 = 0; else if IMC >= 23 and IMC < 26 then IMC2 = 1;
else if IMC >= 26 and IMC < 30 then IMC2 = 2;
else if IMC >= 30 then IMC2 = 3;

if CALC_DUREE_INTER < 45 then DUREE_INTER = 0; else if CALC_DUREE_INTER >= 45 and CALC_DUREE_INTER < 60 then DUREE_INTER = 1;
else if CALC_DUREE_INTER >= 60 and CALC_DUREE_INTER < 80 then DUREE_INTER = 2;
else if CALC_DUREE_INTER >= 80 then DUREE_INTER = 3;

if nb_loca_traite <= 2 then nb_loca = 0; else nb_loca = 1;

format DDN2 ddmyy10. RUTH_J0 RUTH_J.
SEX SEXE. TENSIONART OUIINON.
DYSPLI OUIINON. AVC OUIINON.
CARDIO_IHD OUIINON. CI OUIINON.
AVK OUIINON. IEC OUIINON.
STATINE OUIINON. stent OUIINON.
ballon OUIINON. Reussite OUIINON.
VCS GROUPS. DIAB DIABET.
ASAJ0 ASAJ. AAP AAPJ.
COTE_PONCT COTE_PONCTION. FUMEUR TABAC.
ANESTH ANESTHESIE. AGE2 AGE.
DUREE_H DUREE_H. IMC2 IMC.
DUREE_INTER DUREE_INTER. nb_loca nb_loca.;

drop DDN TABAC SEXE HTA DYSLIPIDEMIE MALADIE_CEREBROVASC
CARDIO_ISCHEMIQUE IC AVK_J0 IEC_J0 STATINE_J0 stent_traite
ballon_traite Succes Group TABAC DIABETE AAPJ0 COTE_PONCTION
ANESTHESIE CALC_IMCJ0;

run;

%let variables_qualitatives = RUTH_J0 SEX TENSIONART DYSPLI AVC CARDIO_IHD CI AVK IEC
STATINE stent ballon DIAB ASAJ0 AAP COTE_PONCT
FUMEUR ANESTH AGE2 DUREE_H IMC2 DUREE_INTER nb_loca;
%let variables_quantitatives = IMC AGE DUREE_HOSPIT CALC_DUREE_INTER nb_loca_traite;

* Transformation des variables quantitatives en qualitatives en s'aidant des quantiles;

proc means data=dir.step q1 median q3 p1 p5 p95 p99;
var &variables_quantitatives;
run;

proc contents data=dir.step;
run;

/* *----- Analyse Univariée sur le Groupe -----* */
/* *----- VARIABLES QUALITATIVES -----* */

* Calcul des répartitions des variables qualitatives et tests d'indépendance du Khi-2;

ods output Chisq = dir.chi2;
ods output CrossTabFreqs = dir.frequencies (keep = RowPercent Frequency VCS Reussite &variables_qualitatives);
proc freq data=dir.step;
table (&variables_qualitatives Reussite) * VCS / chisq nocum;
run;

* Préparation de la table des fréquences (Variable, Modalité, RowPercent, Frequency);

data dir.frequencies;
set dir.frequencies;
informat Variable $30.;
Modalite = strip(compress(vvalue(SEX) || vvalue(TENSIONART) || vvalue(RUTH_J0)
|| vvalue(DYSPLI) || vvalue(AVC)
|| vvalue(CARDIO_IHD) || vvalue(CI)
|| vvalue(AVK) || vvalue(IEC)
|| vvalue(STATINE) || vvalue(stent)
|| vvalue(ballon) || vvalue(Reussite)
|| vvalue(DIAB) || vvalue(ASAJ0)
|| vvalue(AAP) || vvalue(COTE_PONCT)
|| vvalue(FUMEUR) || vvalue(ANESTH)
|| vvalue(AGE2) || vvalue(DUREE_H)
|| vvalue(IMC2) || vvalue(DUREE_INTER)
|| vvalue(nb_loca), "."));

if _N_ < 22 then Variable = "Rutherford";
else if _N_ < 31 and _N_ >= 22 then Variable = "Sexe";
else if _N_ < 40 and _N_ >= 31 then Variable = "Tension Artérielle";
else if _N_ < 49 and _N_ >= 40 then Variable = "Dyslipidémie";
else if _N_ < 58 and _N_ >= 49 then Variable = "Antécédent AVC|AIT";
else if _N_ < 67 and _N_ >= 58 then Variable = "Cardiopathie Ischémique";
else if _N_ < 76 and _N_ >= 67 then Variable = "Insuffisance Cardiaque";
else if _N_ < 85 and _N_ >= 76 then Variable = "Antivitamine K";

```



```

else if _N_ < 94 and _N_ >= 85 then Variable = "Inhibiteurs ACE";
else if _N_ < 103 and _N_ >= 94 then Variable = "Statine";
else if _N_ < 112 and _N_ >= 103 then Variable = "Lésion traitée par stent";
else if _N_ < 121 and _N_ >= 112 then Variable = "Lésion traitée par ballon";
else if _N_ < 130 and _N_ >= 121 then Variable = "Diabète";
else if _N_ < 148 and _N_ >= 130 then Variable = "Score ASA";
else if _N_ < 160 and _N_ >= 148 then Variable = "Médicament Antiplaquettaire";
else if _N_ < 172 and _N_ >= 160 then Variable = "Coté Ponction";
else if _N_ < 184 and _N_ >= 172 then Variable = "Tabac";
else if _N_ < 196 and _N_ >= 184 then Variable = "Type d'Anesthésie";
else if _N_ < 211 and _N_ >= 196 then Variable = "Age";
else if _N_ < 220 and _N_ >= 211 then Variable = "Durée de l'hospitalisation";
else if _N_ < 235 and _N_ >= 220 then Variable = "IMC";
else if _N_ < 250 and _N_ >= 235 then Variable = "Durée de l'intervention (min)";
else if _N_ < 259 and _N_ >= 250 then Variable = "Nombre de localisations";
else if _N_ >= 259 then Variable = "Succès Technique";

```

```

drop RUTH_J0 SEX TENSIONART DYSPLI AVC CARDIO_IHD CI AVK IEC STATINE stent ballon
      Reussite DIAB ASAJ0 AAP COTE_PONCT FUMEUR ANESTH AGE2 DUREE_H IMC2 DUREE_INTER nb_loca;

```

```
if VCS NE . and RowPercent NE . then output;
```

```
run;
```

```
* Tests d'indépendance du Khi-2;
```

```

data dir.chi2;
set dir.chi2;
informat Variable $30.;
if Prob < 0.05
if (Table) = "Table RUTH_J0 * VCS" then Indep = "Non indépendante"; else Indep = "Indépendante";
else if (Table) = "Table SEX * VCS" then Variable = "Rutherford";
else if (Table) = "Table TENSIONART * VCS" then Variable = "Sexe";
else if (Table) = "Table DYSPLI * VCS" then Variable = "Tension Artérielle";
else if (Table) = "Table AVC * VCS" then Variable = "Dyslipidémie";
else if (Table) = "Table CARDIO_IHD * VCS" then Variable = "Antécédent AVC|AIT";
else if (Table) = "Table CI * VCS" then Variable = "Cardiopathie Ischémique";
else if (Table) = "Table AVK * VCS" then Variable = "Insuffisance Cardiaque";
else if (Table) = "Table IEC * VCS" then Variable = "Antivitamine K";
else if (Table) = "Table STATINE * VCS" then Variable = "Inhibiteurs ACE";
else if (Table) = "Table stent * VCS" then Variable = "Statine";
else if (Table) = "Table ballon * VCS" then Variable = "Lésion traitée par stent";
else if (Table) = "Table Reussite * VCS" then Variable = "Lésion traitée par ballon";
else if (Table) = "Table DIAB * VCS" then Variable = "Succès Technique";
else if (Table) = "Table ASAJ0 * VCS" then Variable = "Diabète";
else if (Table) = "Table AAP * VCS" then Variable = "Score ASA";
else if (Table) = "Table COTE_PONCT * VCS" then Variable = "Médicament Antiplaquettaire";
else if (Table) = "Table FUMEUR * VCS" then Variable = "Coté Ponction";
else if (Table) = "Table ANESTH * VCS" then Variable = "Tabac";
else if (Table) = "Table AGE2 * VCS" then Variable = "Type d'Anesthésie";
else if (Table) = "Table DUREE_H * VCS" then Variable = "Age";
else if (Table) = "Table IMC2 * VCS" then Variable = "Durée de l'hospitalisation";
else if (Table) = "Table DUREE_INTER * VCS" then Variable = "IMC";
else if (Table) = "Table nb_loca * VCS" then Variable = "Durée de l'intervention (min)";
else if (Table) = "Table nb_loca * VCS" then Variable = "Nombre de localisations";
if Statistic="Chi-Square" then output;

```

```
* SOURCE D'ERREUR DANS LE CODE Khi-2 OU Chi-Square (langue ou version de SAS);
```

```
keep Variable Prob Indep;
```

```
run;
```

```
* Préparation du tableau pour les variables qualitatives ;
```

```

proc sort data = dir.frequencies ;
by Variable Modalite;
run;

```

```

proc transpose data = dir.frequencies out = dir.frequencies2;
VAR Frequency;
ID VCS;
BY Variable Modalite;
run;

```

```

proc transpose data = dir.frequencies out = dir.frequencies3;
VAR RowPercent;
ID VCS;
BY Variable Modalite;
run;

```

```

proc sort data=dir.chi2;
by Variable;
run;

```

```

data dir.freqs;
merge dir.frequencies2 (rename=(FEMOSEAL=FEMFREQ PROGLIDE=PROGFREQ))
      dir.frequencies3
      dir.chi2;
by Variable;
VAR1 = catt(round(FEMOSEAL,2), "%", " (", FEMFREQ, ")");
VAR2 = catt(round(PROGLIDE,2), "%", " (", PROGFREQ, ")");
keep Variable Modalite VAR1 VAR2 prob indep;
format VAR1 $20. VAR2 $20. indep $17.;
run;

/* *-----VARIABLES QUANTITATIVES -----* */

%let variables_quantitatives = IMC AGE DUREE_HOSPIT CALC_DUREE_INTER nb_loca_traite;

ods output TestsForNormality=dir.quali_tests;
proc univariate data=dir.step normal;
var &variables_quantitatives;
run;

* Histogrammes des variables quantitatives ;

%macro hist;
%do i = 1 %to %sysvalf(%sysfunc(countw(&variables_quantitatives)));
  %let var = %scan(&variables_quantitatives,&i);
  PROC SGPLOT data=dir.step ;
    histogram &var;
  run;
%end;
%mend;

%hist;

* Tests de normalité de Kolmogorov-Smirnov ;

data dir.quali_tests;
set dir.quali_tests;
informat Variable $37.;
if Prob < 0.05 then Dist="Non normale"; else Dist="Normale";
if Test="Kolmogorov-Smirnov" then output;
keep VarName Pvalue Dist;
run;

ods output WilcoxonTest=dir.quali_tests2;
proc npar1way data=dir.step wilcoxon;
class VCS;
var &variables_quantitatives;
run;

* Tests non paramétriques de Wilcoxon-Mann-Whitney
Vu que les variables ne suivent pas le loi normale,
nous allons nous concentrer sur l'étude de leurs médianes
avec les tests de Wilcoxon et Mann-Whitney ;

data dir.quali_tests2;
set dir.quali_tests2;
informat Variable $37.;
if Prob1 < 0.05 then Indep="Non indépendante"; else Indep="Indépendante";
if Variable = "IMC" then Variable = "IMC";
else if Variable = "AGE" then Variable = "Age";
else if Variable = "CALC_DUREE_INTER" then Variable = "Durée de l'intervention (min)";
else if Variable = "DUREE_HOSPIT" then Variable = "Durée de l'hospitalisation";
else if Variable = "nb_loca_traite" then Variable = "Nombre de localisations";
keep Variable Prob1 Indep;
run;

* Calcul des répartitions des variables quantitatives ;

proc means noprint data=dir.step;
class VCS;
var &variables_quantitatives;
output out=dir.quantiles
q1(IMC)= q1(Age)= q1(CALC_DUREE_INTER)= q1(DUREE_HOSPIT)= q1(nb_loca_traite)=
median(IMC)= median(Age)= median(CALC_DUREE_INTER)= median(DUREE_HOSPIT)= median(nb_loca_traite)=
q3(IMC)= q3(Age)= q3(CALC_DUREE_INTER)= q3(DUREE_HOSPIT)= q3(nb_loca_traite)=/autoname;
run;

* Préparation du tableau pour les variables quantitatives ;

data dir.quantiles;
set dir.quantiles;

```

```
drop _TYPE_ _FREQ_;
if VCS NE . then output;
run;
```

```
proc transpose data = dir.quantiles out = dir.quantiles2;
  ID VCS;
run;
```

```
data dir.quantiles2;
set dir.quantiles2;
informat Var $37. Variable $37.;
```

```
if substr(_NAME_,length(_NAME_)-2,3) = "_Q1" then
do;
  Var = substr(_NAME_,1,length(_NAME_)-3);
  Stat = "Q1";
end;
else if substr(_NAME_,length(_NAME_)-2,3) = "_Q3" then
do;
  Var = substr(_NAME_,1,length(_NAME_)-3);
  Stat = "Q3";
end;
else if substr(_NAME_,length(_NAME_)-6,7) = "_Median" then
do;
  Var = substr(_NAME_,1,length(_NAME_)-7);
  Stat = "Median";
end;
```

```
if Var = "IMC" then Variable = "IMC";
else if Var = "AGE" then Variable = "Age";
else if Var = "CALC_DUREE_INTER" then Variable = "Durée de l'intervention (min)";
else if Var = "DUREE_HOSPIT" then Variable = "Durée de l'hospitalisation";
else if Var = "nb_loca_traite" then Variable = "Nombre de localisations";
keep Variable Stat FEMOSEAL PROGLIDE;
run;
```

```
proc sort data=dir.quantiles2;
by Variable;
run;
```

```
proc transpose data = dir.quantiles2 out = dir.quantiles3;
  VAR FEMOSEAL PROGLIDE;
  BY Variable;
  ID Stat;
run;
```

```
data dir.quantiles3;
set dir.quantiles3;
AGREGAT = catt(Me, " [, Q1, ", Q3, " ]");
format AGREGAT $20.;
drop Me Q1 Q3;
run;
```

```
proc sort data=dir.quantiles3;
by Variable;
run;
```

```
proc transpose data = dir.quantiles3 out = dir.quantiles4;
  VAR AGREGAT;
  BY Variable;
  ID _NAME_;
run;
```

```
proc sort data=dir.quali_tests2;
by Variable;
run;
```

```
data dir.meds;
merge dir.quantiles4 (rename=(FEMOSEAL=VAR1 PROGLIDE=VAR2))
  dir.quali_tests2;
by Variable;
keep Variable VAR1 VAR2 Prob1 Indep;
format VAR1 $15. VAR2 $15. Indep $17.;
run;
```

```
* Fusion des tables de toutes les variables ;
```

```
data dir.fusion;
set dir.freqs dir.meds (rename=(Prob1=Prob));
if Modalite NE "" then Modalite = Modalite; else Modalite = "Variable quantitative";
run;
```

```

/* *----- TABLEAU -----* */
proc report data = dir.fusion headline headskip spacing=2 nowd;
columns Variable Modalite VAR1 VAR2 prob Indep;
define Variable / group;
define VAR1 / width=25 center "FEMOSEAL";
define VAR2 / width=25 center "PROGLIDE";
define prob / width=25 "P-value";
define Indep / group width=25 center "Indépendance";
run;

/* *----- Analyse Univariée sur le Succès -----* */

/* *----- VARIABLES QUALITATIVES -----* */

* Calcul des répartitions des variables qualitatives et tests d'indépendance du Khi-2;

ods output Chisq = dir.chi2;
ods output CrossTabFreqs = dir.frequencies (keep = RowPercent Frequency VCS Reussite &variables_qualitatives);
proc freq data=dir.step;
table (&variables_qualitatives VCS) * Reussite / chisq nocum;
run;

* Préparation de la table des fréquences (Variable, Modalité, RowPercent, Frequency);

data dir.frequencies;
set dir.frequencies;
informat Variable $30.;
Modalite = strip(compress(vvalue(SEX) || vvalue(TENSIONART) || vvalue(RUTH_J0)
|| vvalue(DYSPLI) || vvalue(AVC)
|| vvalue(CARDIO_IHD) || vvalue(CI)
|| vvalue(AVK) || vvalue(IEC)
|| vvalue(STATINE) || vvalue(stent)
|| vvalue(ballon) || vvalue(VCS)
|| vvalue(DIAB) || vvalue(ASAJ0)
|| vvalue(AAP) || vvalue(COTE_PONCT)
|| vvalue(FUMEUR) || vvalue(ANESTH)
|| vvalue(AGE2) || vvalue(DUREE_H)
|| vvalue(IMC2) || vvalue(DUREE_INTER)
|| vvalue(nb_loca), "."));

if _N_ < 22 then Variable = "Rutherford";
else if _N_ < 31 and _N_ >= 22 then Variable = "Sexe";
else if _N_ < 40 and _N_ >= 31 then Variable = "Tension Artérielle";
else if _N_ < 49 and _N_ >= 40 then Variable = "Dyslipidémie";
else if _N_ < 58 and _N_ >= 49 then Variable = "Antécédent AVC|AIT";
else if _N_ < 67 and _N_ >= 58 then Variable = "Cardiopathie Ischémique";
else if _N_ < 76 and _N_ >= 67 then Variable = "Insuffisance Cardiaque";
else if _N_ < 85 and _N_ >= 76 then Variable = "Antivitamine K";
else if _N_ < 94 and _N_ >= 85 then Variable = "Inhibiteurs ACE";
else if _N_ < 103 and _N_ >= 94 then Variable = "Statine";
else if _N_ < 112 and _N_ >= 103 then Variable = "Lésion traitée par stent";
else if _N_ < 121 and _N_ >= 112 then Variable = "Lésion traitée par ballon";
else if _N_ < 130 and _N_ >= 121 then Variable = "Diabète";
else if _N_ < 148 and _N_ >= 130 then Variable = "Score ASA";
else if _N_ < 160 and _N_ >= 148 then Variable = "Médicament Antiplaquettaire";
else if _N_ < 172 and _N_ >= 160 then Variable = "Coté Ponction";
else if _N_ < 184 and _N_ >= 172 then Variable = "Tabac";
else if _N_ < 196 and _N_ >= 184 then Variable = "Type d'Anesthésie";
else if _N_ < 211 and _N_ >= 196 then Variable = "Age";
else if _N_ < 220 and _N_ >= 211 then Variable = "Durée de l'hospitalisation";
else if _N_ < 235 and _N_ >= 220 then Variable = "IMC";
else if _N_ < 250 and _N_ >= 235 then Variable = "Durée de l'intervention (min)";
else if _N_ < 259 and _N_ >= 250 then Variable = "Nombre de localisations";
else if _N_ >= 259 then Variable = "Groupe";

drop RUTH_J0 SEX TENSIONART DYSPLI
AVC CARDIO_IHD CI AVK IEC STATINE stent ballon
VCS DIAB ASAJ0 AAP COTE_PONCT FUMEUR ANESTH
AGE2 DUREE_H IMC2 DUREE_INTER nb_loca;

if Reussite NE . and RowPercent NE . then output;

run;

* Tests d'indépendance du Khi-2;

data dir.chi2;
set dir.chi2;
informat Variable $30.;
if Prob < 0.05 then Indep="Non indépendante"; else Indep="Indépendante";
if (Table) = "Table RUTH_J0 * Reussite" then Variable = "Rutherford";
else if (Table) = "Table SEX * Reussite" then Variable = "Sexe";

```

```

else if (Table) = "Table TENSIONART * Reussite" then Variable = "Tension Artérielle";
else if (Table) = "Table DYSPLI * Reussite" then Variable = "Dyslipidémie";
else if (Table) = "Table AVC * Reussite" then Variable = "Antécédent AVC|AIT";
else if (Table) = "Table CARDIO_IHD * Reussite" then Variable = "Cardiopathie Ischémique";
else if (Table) = "Table CI * Reussite" then Variable = "Insuffisance Cardiaque";
else if (Table) = "Table AVK * Reussite" then Variable = "Antivitamine K";
else if (Table) = "Table IEC * Reussite" then Variable = "Inhibiteurs ACE";
else if (Table) = "Table STATINE * Reussite" then Variable = "Statine";
else if (Table) = "Table stent * Reussite" then Variable = "Lésion traitée par stent";
else if (Table) = "Table ballon * Reussite" then Variable = "Lésion traitée par ballon";
else if (Table) = "Table VCS * Reussite" then Variable = "Groupe";
else if (Table) = "Table DIAB * Reussite" then Variable = "Diabète";
else if (Table) = "Table ASAJ0 * Reussite" then Variable = "Score ASA";
else if (Table) = "Table AAP * Reussite" then Variable = "Médicament Antiplaquettaire";
else if (Table) = "Table COTE_PONCT * Reussite" then Variable = "Coté Ponction";
else if (Table) = "Table FUMEUR * Reussite" then Variable = "Tabac";
else if (Table) = "Table ANESTH * Reussite" then Variable = "Type d'Anesthésie";
else if (Table) = "Table AGE2 * Reussite" then Variable = "Age";
else if (Table) = "Table DUREE_H * Reussite" then Variable = "Durée de l'hospitalisation";
else if (Table) = "Table IMC2 * Reussite" then Variable = "IMC";
else if (Table) = "Table DUREE_INTER * Reussite" then Variable = "Durée de l'intervention (min)";
else if (Table) = "Table nb_loca * Reussite" then Variable = "Nombre de localisations";
if Statistic="Chi-Square" then output;

```

* SOURCE D'ERREUR DANS LE CODE Khi-2 OU Chi-Square (langue ou version de SAS);

```

keep Variable Prob Indep;
run;

```

* Préparation du tableau pour les variables qualitatives ;

```

proc sort data = dir.frequencies ;
by Variable Modalite;
run;

```

```

proc transpose data = dir.frequencies out = dir.frequencies2;
VAR Frequency;
ID Reussite;
BY Variable Modalite;
run;

```

```

proc transpose data = dir.frequencies out = dir.frequencies3;
VAR RowPercent;
ID Reussite;
BY Variable Modalite;
run;

```

```

proc sort data=dir.chi2;
by Variable;
run;

```

```

data dir.freqs;
merge dir.frequencies2 (rename=(OUI=OUIFREQ NON=NONFREQ))
      dir.frequencies3
      dir.chi2;
by Variable;
VAR1 = catt(round(OUI,2), "%", " (", OUIFREQ, ")");
VAR2 = catt(round(NON,2), "%", " (", NONFREQ, ")");
keep Variable Modalite VAR1 VAR2 prob indep;
format VAR1 $20. VAR2 $20. indep $17.;
run;

```

/* *-----VARIABLES QUANTITATIVES -----* */

```

ods output WilcoxonTest=dir.quali_tests2;
proc npar1way data=dir.step wilcoxon;
class Reussite;
var &variables_quantitatives;
run;

```

* Tests non paramétriques de Wilcoxon-Mann-Whitney
Vu que les variables ne suivent pas la loi normale,
nous allons nous concentrer sur l'étude de leurs médianes
avec les tests de Wilcoxon et Mann-Whitney ;

```

data dir.quali_tests2;
set dir.quali_tests2;
informat Variable $37.;
if Prob1 < 0.05 then Indep="Non indépendante"; else Indep="Indépendante";
if Variable = "IMC" then Variable = "IMC";
else if Variable = "AGE" then Variable = "Age";
else if Variable = "CALC_DUREE_INTER" then Variable = "Durée de l'intervention (min)";

```

```

else if Variable = "DUREE_HOSPIT"      then Variable = "Durée de l'hospitalisation";
else if Variable = "nb_loca_traite"     then Variable = "Nombre de localisations";
keep Variable Prob1 Indep;
run;

```

* Calcul des répartitions des variables quantitatives ;

```

proc means noprint data=dir.step;
class Reussite;
var &variables_quantitatives;
output out=dir.quantiles
q1(IMC)= q1(Age)= q1(CALC_DUREE_INTER)= q1(DUREE_HOSPIT)= q1(nb_loca_traite)=
median(IMC)= median(Age)= median(CALC_DUREE_INTER)= median(DUREE_HOSPIT)= median(nb_loca_traite)=
q3(IMC)= q3(Age)= q3(CALC_DUREE_INTER)= q3(DUREE_HOSPIT)= q3(nb_loca_traite)=/autoname;
run;

```

* Préparation du tableau pour les variables quantitatives ;

```

data dir.quantiles;
set dir.quantiles;
drop _TYPE_ _FREQ_;
if Reussite NE . then output;
run;

```

```

proc transpose data = dir.quantiles out = dir.quantiles2;
  ID Reussite;
run;

```

```

data dir.quantiles2;
set dir.quantiles2;
informat Var $37. Variable $37.;

```

```

if      substr(_NAME_,length(_NAME_)-2,3) = "_Q1"      then
do;
  Var = substr(_NAME_,1,length(_NAME_)-3);
  Stat = "Q1";
end;
else if substr(_NAME_,length(_NAME_)-2,3) = "_Q3"      then
do;
  Var = substr(_NAME_,1,length(_NAME_)-3);
  Stat = "Q3";
end;
else if substr(_NAME_,length(_NAME_)-6,7) = "_Median" then
do;
  Var = substr(_NAME_,1,length(_NAME_)-7);
  Stat = "Median";
end;

```

```

if      Var = "IMC"      then Variable = "IMC";
else if Var = "AGE"     then Variable = "Age";
else if Var = "CALC_DUREE_INTER" then Variable = "Durée de l'intervention (min)";
else if Var = "DUREE_HOSPIT" then Variable = "Durée de l'hospitalisation";
else if Var = "nb_loca_traite" then Variable = "Nombre de localisations";
keep Variable Stat OUI NON;
run;

```

```

proc sort data=dir.quantiles2;
by Variable;
run;

```

```

proc transpose data = dir.quantiles2 out = dir.quantiles3;
  VAR OUI NON;
  BY Variable;
  ID Stat;
run;

```

```

data dir.quantiles3;
set dir.quantiles3;
AGREGAT = catt(Me, " [", Q1, ",", Q3, "]");
format AGREGAT $20.;
drop Me Q1 Q3;
run;

```

```

proc sort data=dir.quantiles3;
by Variable;
run;

```

```

proc transpose data = dir.quantiles3 out = dir.quantiles4;
  VAR AGREGAT;
  BY Variable;
  ID _NAME_;
run;

```

```

proc sort data=dir.quali_tests2;
by Variable;
run;

data dir.meds;
merge dir.quantiles4 (rename=(OUI=VAR1 NON=VAR2))
      dir.quali_tests2;
by Variable;
keep Variable VAR1 VAR2 Prob1 Indep;
format VAR1 $15. VAR2 $15. Indep $17.;
run;

* Fusion des tables de toutes les variables ;

data dir.fusion;
set dir.freqs dir.meds (rename=(Prob1=Prob));
if Modalite NE "" then Modalite = Modalite; else Modalite = "Variable quantitative";
run;

/* *----- TABLEAU -----* */

proc report data = dir.fusion headline headskip spacing=2 nowd;
columns Variable Modalite VAR1 VAR2 prob Indep;
define Variable / group;
define VAR1 / width=25 center "OUI";
define VAR2 / width=25 center "NON";
define prob / width=25 "P-value";
define Indep / group width=25 center "Indépendance";
run;

* Modèle de régression logistique simple pour chaque variable explicative ;

%macro logistic;
%let variables = RUTH_J0 SEX TENSIONART DYSPLI AVC CARDIO_IHD CI AVK IEC VCS
                STATINE stent ballon DIAB ASAJ0 AAP FUMEUR ANESTH DUREE_H IMC2 nb_loca;

%do i = 1 %to %sysfunc(countw(&variables));
%let var = %scan(&variables,&i);

title "&var";
ODS EXCLUDE ResponseProfile FitStatistics Association;
PROC LOGISTIC data=dir.step descending ;
class &var;
model Reussite = &var; /*permet d'avoir le test global*/
run;

%end;
%mend logistic;

%logistic;

* p-value < 20% retenues ;
%let retenues = DYSPLI CI ballon nb_loca VCS;

* Modèle de régression logistique multinomial ;

%macro multivar;
title "&retenues";
PROC LOGISTIC data=dir.step descending ;
class &retenues;
model Reussite = &retenues; /*permet d'avoir le test global*/
run;
%mend;

%multivar;

* On retire nb_loca ;
%let retenues = DYSPLI CI ballon VCS;
%multivar;

* On retire CI ;
%let retenues = DYSPLI ballon VCS;
%multivar;

* On s'arrête si significativité à 20% ;

* On teste les potentiels croisements ;

%macro multivarX;
%do i = 1 %to %sysfunc(countw(&retenues))-1;
%let var1 = %scan(&retenues,&i);

```

```
%do j = %sysevalf(&i+1) %to %sysfunc(countw(&retenues));
  %let var2 = %scan(&retenues,&j);
  title "&retenues &var1.*&var2";
  PROC LOGISTIC data=dir.step descending ;
  class &retenues;
  model Reussite = &retenues &var1.*&var2; /*permet d'avoir le test global*/
  run;
%end;
%end;
%mend;

%multivarX;

* On rajoute les croisements significatifs DYSPLI*ballon et DYSPLI*VCS;

title "DYSPLI ballon VCS DYSPLI*ballon DYSPLI*VCS";
ods graphics on;
PROC LOGISTIC data=dir.step descending plots(only)=roc(id=obs);
class DYSPLI ballon VCS;
model Reussite = DYSPLI ballon VCS DYSPLI*ballon DYSPLI*VCS /lackfit;
roc DYSPLI ballon VCS DYSPLI*ballon DYSPLI*VCS;
run;
```