

Détection des patients pour la prévention primaire et secondaire des maladies vasculaires

Ismat Chaib Draa, Sébastien Amiot, Julien Taillard, Abdelahad Chraibi, David Delerue, Adrien Hertault, Bruno Lernout

► **To cite this version:**

Ismat Chaib Draa, Sébastien Amiot, Julien Taillard, Abdelahad Chraibi, David Delerue, et al.. Détection des patients pour la prévention primaire et secondaire des maladies vasculaires. 10ème conférence Francophone en Gestion et Ingénierie des Systèmes Hospitaliers, GISEH2020, Oct 2020, Valenciennes, France. hal-03188041

HAL Id: hal-03188041

<https://hal-uphf.archives-ouvertes.fr/hal-03188041>

Submitted on 1 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Détection des patients pour la prévention primaire et secondaire des maladies vasculaires

Chaib Draa Ismat¹, Amiot Sebastien², Julien Taillard¹, Abdelahad Chraibi¹, David Delerue¹, Hertault Adrien² ; Bruno Lernout²

¹ Alicante, Philippe Girard, Seclin, ismat.chaib-draa@alicante.fr

² Centre Hospitalier de Valenciennes, 57 Avenue Desandrouins, 59300 Valenciennes, lernout-b@ch-valenciennes.fr

Résumé. Cet article présente les premières briques du travail réalisé dans le cadre du projet intitulé « Passeport Vasculaire », ainsi que les premiers résultats préliminaires. Ce travail est réalisé conjointement par le service de chirurgie vasculaire du Centre Hospitalier de Valenciennes (CHV) et l'entreprise Alicante. L'objectif est de détecter parmi l'ensemble des patients admis ou pré-admis au CHV le risque de pathologies vasculaires. Pour mener à terme ce projet, l'expertise des chirurgiens du pôle vasculaire du CHV est mise à disposition, ainsi qu'une combinaison de plusieurs techniques d'algorithmique et d'intelligence artificielle (IA). Ces techniques d'IA ont été exploitées pour tirer profit de l'importante masse de données hétérogènes et complémentaires disponibles au CHV.

Mots clés : prévention des maladies vasculaires, intelligence artificielle, classification

Contexte et motivation

Le système vasculaire également appelé système cardiovasculaire est un réseau interconnecté constitué de vaisseaux sanguins (veines et artères) relié au cœur et aux poumons afin d'apporter l'oxygène et les nutriments à tous les organes et tissus du corps. Le système vasculaire fait partie des systèmes vitaux de l'organisme. Pour cette raison, les maladies touchant le système vasculaire impactent de manière significative toutes les facettes de la santé. Les maladies cardiovasculaires demeurent la principale cause de mortalité dans le monde chez les adultes d'âge moyen en 2019. Parmi ces maladies, nous avons les cardiopathies coronariennes, les maladies cérébro-vasculaires, les artériopathies périphériques, les cardiopathies rhumatismales et les thromboses veineuses profondes et embolies pulmonaires.

Toujours selon l'OMS, les maladies cardio-vasculaires sont la première cause de mortalité dans le monde. Chaque année, 17.7 millions de décès sont imputables aux maladies cardio-vasculaires, soit 31% de la mortalité mondiale totale. Parmi ces décès, 7.4 millions sont causés par des cardiopathies coronarienne et 6.7 millions sont dus à un AVC (chiffres 2015). Les maladies de l'appareil circulatoire dans le territoire du HAINAUT CAMBRESIS représentent 2563 décès en moyenne annuelle (26% des décès) avec :

- Une surmortalité en regard de la France hexagonale de +33% (+31% pour les hommes et +36% pour les femmes)
- 678 décès évitables par an avec la même mortalité qu'en France hexagonale (313 pour les hommes et 365 pour les femmes)
- Chez les moins de 65 ans la surmortalité est plus marquée avec près de +78% chez les femmes et près de +60% chez les hommes.

Parmi les maladies cardio-vasculaires les plus répandues, nous avons les infarctus et les accidents vasculaires cérébraux qui sont généralement des événements aigus et sont principalement dus au blocage d'une artère empêchant le sang de parvenir au cœur ou au cerveau. Les infarctus et les AVC sont généralement causés par la présence de plusieurs facteurs de risque associés comme le tabagisme, une mauvaise alimentation et l'obésité, la sédentarité et la consommation excessive de l'alcool,

l'hypertension, le diabète et l'hyperlipidémie. La gravité des maladies vasculaire ainsi que la mortalité qui en résulte souligne l'importance de la mobilisation du monde médical autour de cette problématique et la nécessité de mettre en œuvre des processus de détection des profils à risque pour la prévention.

1 Objectif du projet

1.1 Problématique

Les patients souffrant de maladies cardiovasculaires ou exposés à un risque élevé de ces maladies nécessitent une détection précoce pour aboutir à une prise en charge efficace.

Partant de ce postulat, le service de chirurgie vasculaire du CHV a décidé de mettre en place un processus de détection des patients à risque vasculaire car le dépistage systématique n'a, en effet pas trouvé son efficacité et demeure très couteux en temps et en logistique pour de maigres résultats. Ce projet vise à faire bénéficier tous les patients pré-admis ou admis au CHV d'actions de prévention des maladies vasculaires. Ces actions seront ciblées, personnalisées et pourront intervenir quel que soit le service d'admission du patient à l'hôpital. Ces actions de prévention seront basées sur les données rétrospectives des patients ainsi que leurs parcours de soins réalisés au CHV. Ces données sont de nature hétérogène et comportent des données structurées du type PMSI (E.Faure, 2015), des mesures biologiques et également des données non structurées comme les courriers médicaux. Les données exploitées dans le cadre de ce projet sont détaillées dans la section 2.1. Certains travaux se sont intéressés à la détection des facteurs de risques vasculaire en exploitant des algorithmes d'IA (Rajmohan, 2014). D'autres travaux comme (Motwani, 2016) et (Sun, 2008) se sont également intéressés à la prédiction des maladies vasculaires. La différence majeure entre notre travail et les travaux cités réside dans le nombre de critères que nous prenons en compte et les différentes sources de données utilisées. En effet, les travaux cités prennent en moyenne 5 à 6 critères de risque pour la modélisation de leurs algorithmes d'IA lorsque nous en considérons plus de 30. Ajoutons également le fait que ces travaux ont été réalisés sur des données structurées et l'exploration des données rétrospectives et celles des courriers n'a pas été considérée.

1.2 Aperçu du passeport vasculaire

La figure 1 ci-dessous représente l'architecture simplifiée du mode de fonctionnement de la détection des profils à risque vasculaire.

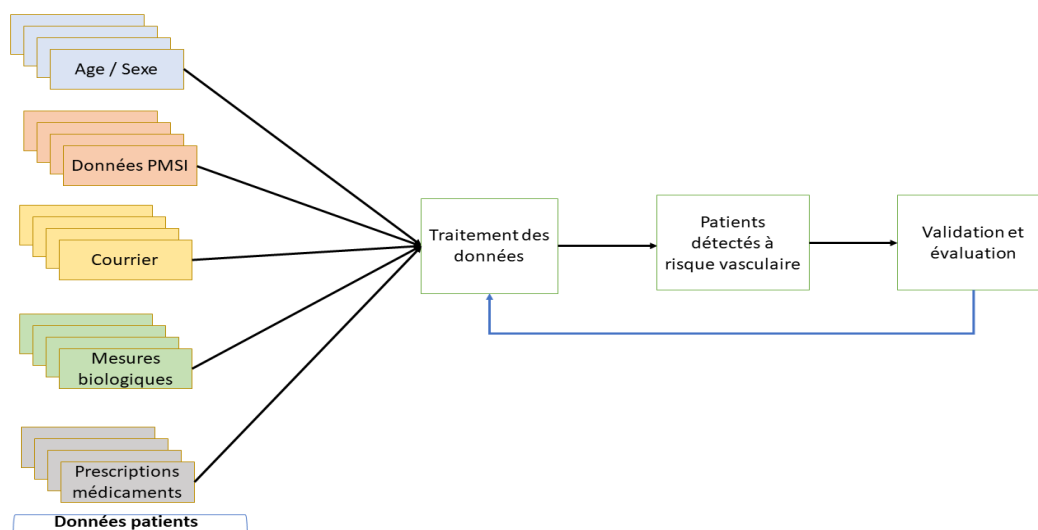


Figure 1 Architecture simplifiée du passeport vasculaire

Dans la figure ci-dessus, nous avons en entrée les données exploitées dans le cadre de la prévention primaire et secondaire des maladies vasculaires. Ces données rétrospectives sont désidentifiées et comportent :

- L'âge et le sexe du patient
- Les données PMSI (section 2.1.1)
- Les courriers médicaux
- Les mesures biologiques
- Les prescriptions médicamenteuses

Ces données en entrée ont été sélectionnées sur la base des règles d'experts exposées par les chirurgiens ainsi qu'une multitude d'algorithmes d'intelligence artificielle et de traitement automatique du langage naturel (TALN). Ces données alimenteront nos différents classifieurs pour aboutir à des listes de patients éligibles à la prévention primaire ou secondaire. Ces listes seront par la suite consultées, évaluées et validées par les chirurgiens vasculaires du CHV. Après évaluation, les algorithmes de détection seront corrigés et réévalués avant d'être déployés.

2 Matériels et méthodologies

La section 2 présente en détail les données utilisées ainsi que les différentes techniques d'IA pour tirer profit de l'importante masse de données disponible au CHV.

2.1 Les données utilisées

Comme mentionné précédemment, dans le cadre de ce travail, différentes sources de données disponibles au CHV ont été exploitées. Ces données peuvent être structurées comme les données PMSI (Programme de Médicalisation des Systèmes d'Information), les données biologiques et les prescriptions médicamenteuses. Nous avons également des données textuelles non-structurées à travers les courriers médicaux rédigés à la fin de chaque séjour.

2.1.1 Les données PMSI

Le Programme de Médicalisation des Systèmes d'Information (PMSI) est un dispositif faisant partie de la réforme du système de santé français dont le but est de réduire les inégalités de ressources entre les établissements de santé. Dans le but de mesurer l'activité et les ressources des établissements, il est nécessaire de disposer d'informations quantifiées et standardisées. Dans le cadre de ce travail, nous nous intéressons principalement aux principes du PMSI MCO (Médecine Chirurgie Obstétrique).

Les données PMSI que nous exploitons sont : le motif de la prise en charge médicale, la durée du séjour, l'unité médicale responsable de la prise en charge, le diagnostic principal de la prise en charge ainsi que les diagnostics associés et reliés. Les actes chirurgicaux effectués au cours du séjour sont également pris en considération.

2.1.2 Les courriers

A la fin de chaque séjour médical, un courrier est rédigé par les médecins responsables du séjour mentionnant le déroulement du séjour du patient. Les courriers représentent une mine d'information considérable mais souvent inexploités notamment car ceux-ci demeurent non-standardisées. Nous pouvons retrouver dans ces courriers diverses informations comme les antécédents médicaux du patient, son mode de vie, les prescriptions médicamenteuses, etc. Les courriers sont traités via des techniques de traitement automatique du langage naturel (TALN) pour en extraire des informations sur le patient qui ne sont pas disponibles dans les données PMSI. Les informations mentionnées dans le courrier renforcent également les autres sources d'informations. Le traitement de courriers représente l'étape la

plus importante de ce projet et nous permet d'avoir tous les critères de risque concernant le patient ainsi que les examens médicaux réalisés et leurs résultats. Les courriers nous offrent donc une vision exhaustive sur le séjour du patient.

2.1.3 Les mesures biologiques

Les mesures biologiques représentent les résultats des différentes analyses de sang réalisées au sein de l'hôpital au cours d'un séjour pour un patient donné. Ces mesures indiquent les différentes constantes du patient comme par exemple la glycémie à jeun, les taux de cholestérol, l'acide urique, etc. Les mesures biologiques sont accessibles sous une forme structurée et numérique. Par exemple : Vitamine B12 ;412.0 ; ng/l.

2.2 Prescriptions médicamenteuses

Les médicaments prescrits lors d'un séjour pour un patient sont également utilisés et sont disponibles sous forme de données structurées. La prise de certains médicaments peut également être marqueurs de certains facteurs de risque.

2.2 Algorithmes

2.2.1 Constitution du dataset

Pour mener à terme l'identification des patients éligibles à une détection des pathologies vasculaires, nous collectons les données susmentionnées. Nous nous intéressons en premier lieu aux patients ayant déjà effectué un séjour dans le service de chirurgie vasculaire. Pour ce faire, nous réalisons des requêtes via l'entrepôt de données du CHV pour accéder à l'ensemble des séjours des patients ayant dans leurs parcours des séjours à caractère vasculaire.

Cette recherche se fait sur la base de plusieurs critères :

- La présence du champ unité médicale (UM) "chirurgie vasculaire"
- La présence de termes évocateurs dans les courriers comme "service de chirurgie vasculaire", "revascularisation", les noms des chirurgiens du service vasculaire, etc.
- Traitement du langage naturel pour l'extraction du champ lexical relatif au service de chirurgie vasculaire
- La présence d'un diagnostic principal propre aux pathologies vasculaires comme "I251" pour une cardiopathie artérioscléreuse.
- La présence d'actes de chirurgie vasculaire comme le code "DDMA025" pour une revascularisation coronaire.

Chaque patient ayant effectué un séjour en chirurgie vasculaire est sélectionné pour que l'ensemble de ses séjours au sein du CHV soient étudiés. Cette étude servira à enrichir et valider notre base de connaissance ainsi que les critères de détection des futurs profils à risque. L'étude de ces patients est réalisée via des algorithmes d'IA tels que les arbres de décision et le traitement automatique du langage naturel. Avec les arbres de décision, nous nous intéressons principalement aux données PMSI pour retrouver les diagnostics les plus récurrents pour les actes chirurgicaux vasculaires ainsi que les médicaments prescrits. Le traitement du courrier est réalisé pour appuyer la recherche de diagnostics et l'enrichir. Ce processus est schématisé dans la figure 2 suivante.

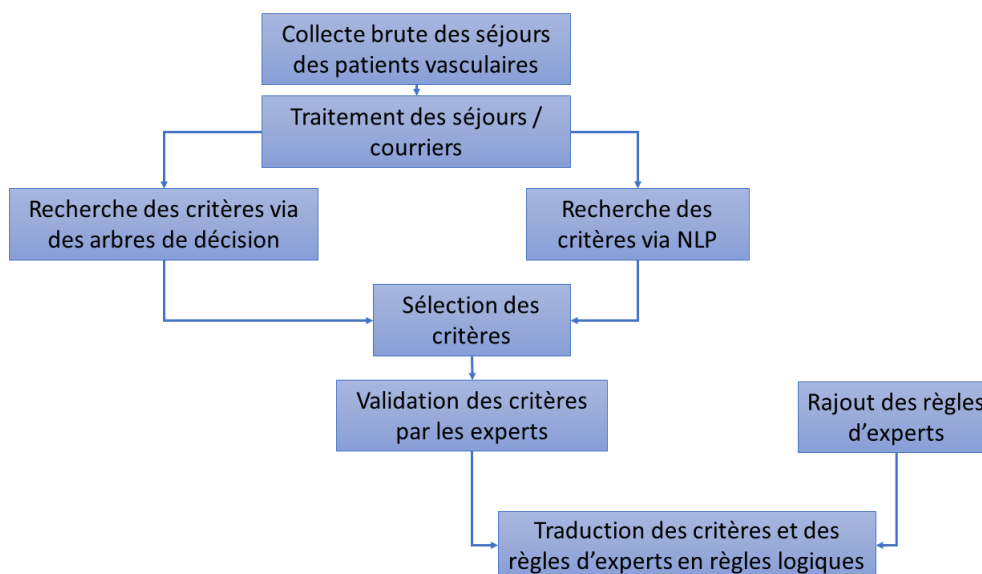


Figure 2 Processus de recherche des critères de détection des profils à risque

2.2.2 Traitement des données PMSI avec des arbres de décision

Les arbres de décision (AD) sont une catégorie d'arbres utilisée dans l'exploration de données et en informatique décisionnelle. Ils emploient une représentation hiérarchique de la structure des données sous forme des séquences de décisions en vue de la prédiction d'un résultat ou d'une classe. Chaque observation, qui doit être attribuée à une classe, est décrite par un ensemble de variables qui sont testées dans les nœuds de l'arbre. Les tests s'effectuent dans les nœuds internes et les décisions sont prises dans les nœuds feuille. Les arbres de décision ont été utilisés dans le cadre de ce travail pour consolider les règles exposées par les chirurgiens vasculaires et pour découvrir davantage de facteurs de risque des pathologies vasculaires. Pour l'exploitation des arbres de décisions, nous avons exclusivement les diagnostics et les actes disponibles dans le PMSI de chaque séjour. À chaque acte chirurgical est associé un diagnostic principal ainsi que plusieurs diagnostics associés et reliés. Pour l'ensemble des patients sélectionnés, nous avons regroupé tous les actes chirurgicaux similaires pour déterminer quels étaient les différents diagnostics qui y ont mené. Ainsi la représentation des données en entrée de nos arbres de décision peut être représentée comme suit :

Acte	Age	Sexe	Diagnostic principal	Diagnostics reliés	Diagnostics associés
DDMA025	45	Femme	Z95	I20	E78

Tableau 1 Exemple des entrées d'un arbre de décision vasculaire

Dans le tableau ci-dessus, pour un acte de revascularisation « DDMA025 », le diagnostic principal « Z95 » représente un pontage aorto-coronarien, le diagnostic relié « I20 » représente une angine de poitrine instable et le diagnostic associé « E78 » est synonyme d'hyperlipidémie. Ainsi ce tableau représente une observation de six variables.

Pour davantage de lisibilité, nous avons remplacé les codes diagnostics par leurs noms. La figure suivante représente un extrait d'un arbre de décision réalisé pour l'acte de revascularisation « DDMA025 ».

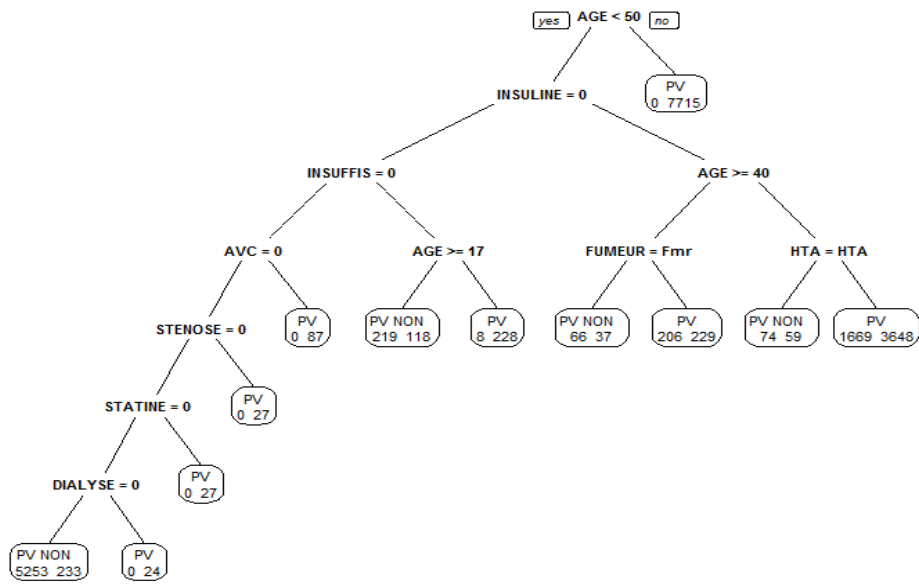


Figure 3 Aperçu d'arbre de décision pour l'acte "DDMA025"

Cet arbre de décision nous montre de manière explicite quels sont les diagnostics les plus récurrents ainsi que la tranche d'âge la plus fréquente chez les patients ayant subi une revascularisation. Nous retrouvons par exemple l'hypertension artérielle (HTA), le tabagisme, l'insuffisance cardiaque, un antécédent d'AVC, etc. Ces différents diagnostics sont sauvegardés et seront validés par la suite par les chirurgiens.

2.2.3 Traitement du courrier

Les différents courriers de séjour des patients vasculaires sont traités pour extraire des termes pertinents pour la détection des futurs profils à risque comme illustré ci-dessous.

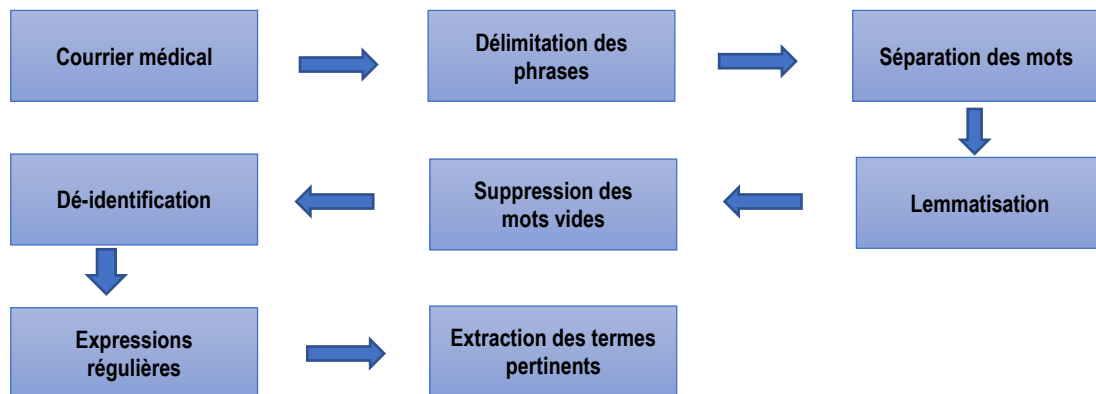


Figure 4 TALN pour les courriers médicaux

Les courriers associés à chaque séjour et à chaque patient sont traités en commençant par une délimitation des phrases et une séparation des mots. Ensuite, l'ensemble des mots composants le courrier sont réduits à leurs formes canoniques à travers la lemmatisation. Les mots vides comme par exemple : le, la, de, chez, sont supprimés. Les noms propres, les adresses, les dates de naissance ainsi que les numéros de téléphones sont codifiés pour ne pas être identifiables. Ce processus consiste en la dé-identification du courrier. Cette phase est suivie par l'application de multiples expressions régulières pour extraire les antécédents médicaux des patients, les motifs d'hospitalisation, les différents

appareillages utilisés durant le séjour, le mode de vie du patient ainsi que les prescriptions médicamenteuses. En addition à ce processus, certains critères de risque sont recherchés directement à partir du courrier comme le tabagisme, l'insuffisance rénale, l'hypertension artérielle, etc. Ces critères sont considérés comme des facteurs de risques indéniables dans la littérature des maladies vasculaires (Ford, 2004). Nous présentons dans la section résultats (Section 3) les termes médicaux les plus récurrents qui résultent du traitement des courriers.

2.2.4 Validation des critères de risque et règles experts

Les résultats des arbres de décision ainsi que le traitement des courriers ont permis d'obtenir un nombre conséquent de diagnostics et de termes évocateurs aux pathologies vasculaires. La validation de ces résultats doit impérativement se faire par les experts du domaine. Pour ce faire, les différents concepts et termes médicaux résultants de notre phase d'apprentissage ont été hiérarchisés selon leurs natures comme illustré ci-dessous.

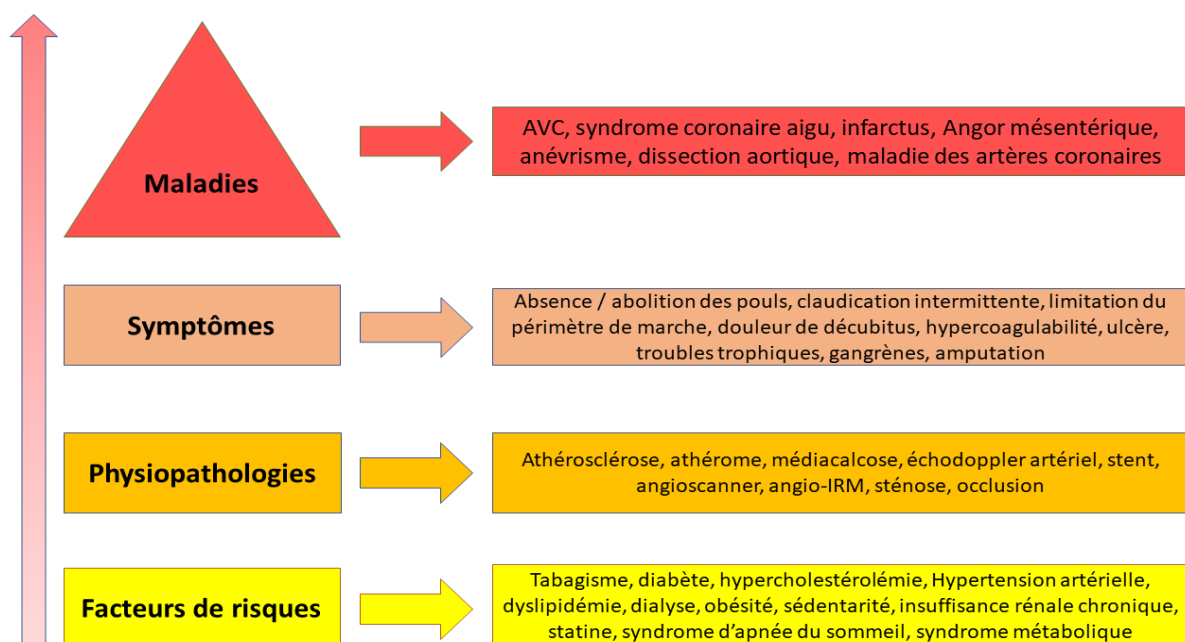


Figure 5 Hiérarchisation des critères de risque

Les quatre niveaux sélectionnés sont les facteurs de risque, les physiopathologies, les symptômes et les maladies. A ce stade du projet, chaque niveau représente un degré de gravité auquel est associé un score de 1 pour les facteurs de risque, 2 pour les physiopathologies, 3 pour les symptômes et 4 pour les maladies. Les scores associés à chaque critère de risque permettent de calculer un score final pour chaque patient. Sachant que ces facteurs de risques se potentialisent (Fédération française de cardiologie, 2017), un patient avec un léger degré de diabète et une légère dyslipidémie sera plus à risque qu'un patient avec un diabète très avancé. Partant de ce postulat, le score final de chaque patient représente la somme des scores des différents critères corrélés à l'âge et au sexe. En effet, les femmes âgées de plus de 65 ans ont un risque plus élevé que les hommes du même âge. Inversement, avant 65 ans, le risque des pathologies vasculaires est plus élevé chez les hommes. D'autre part, le dépistage des maladies vasculaires se veut systématique dès que le patient est âgé de plus de 70 ans peu importe son sexe et ses antécédents médicaux. Concernant la dichotomie entre la prévention primaire ou secondaire, cette dernière se fait en fonction de la présence ou l'absence des maladies (niveau 4). Les patients ayant déjà eu des maladies vasculaires en dehors du CHV rentrent dans le cadre de la prévention secondaire. Le reste des patients détectés seront éligibles à la prévention primaire. Dans la section suivante, nous présentons les résultats préliminaires de ce projet.

3 Résultats

Dans cette section, nous présentons les premiers résultats préliminaires obtenus. Ces résultats ont été présentés aux chirurgiens sous forme de listes qui ont été pré-validées et discutées. Sur l'ensemble des patients ayant effectué un séjour au CHV ces deux dernières années, 8250 profils à risque vasculaires ont été détectés. 8250 patients sur 264 000. Parmi ces 8250 patients, 4280 ont été détectés uniquement sur la base de l'âge supérieur à 70 ans. 3970 patients sont éligibles à un dépistage sur la base des différentes règles implémentées. Ces patients sont répartis selon quatre classes différentes. Cette classification a été réalisée en fonction du score de chaque patient. Le nombre de patients ayant eu au moins un séjour vasculaire à l'hôpital ces deux dernières années est de 2580. Rappelons que l'ensemble de nos études rétrospectives pour la détection des critères de risques ont été basées sur ces patients vasculaires. Nous présentons en premier lieu les résultats des statistiques descriptives sur les critères de risque (Figure 5) retrouvés chez les patients vasculaires en comparaison avec ces critères chez des patients non-vasculaires. Nous avons pris un échantillon de 2500 patients pour chaque population, vasculaire (VA) et non-vasculaire (NVA)

Les figures 6,7,8 et 9 représentent respectivement la répartition des facteurs de risque, physiopathologies, symptômes et maladies chez les deux population VA et N-VA.

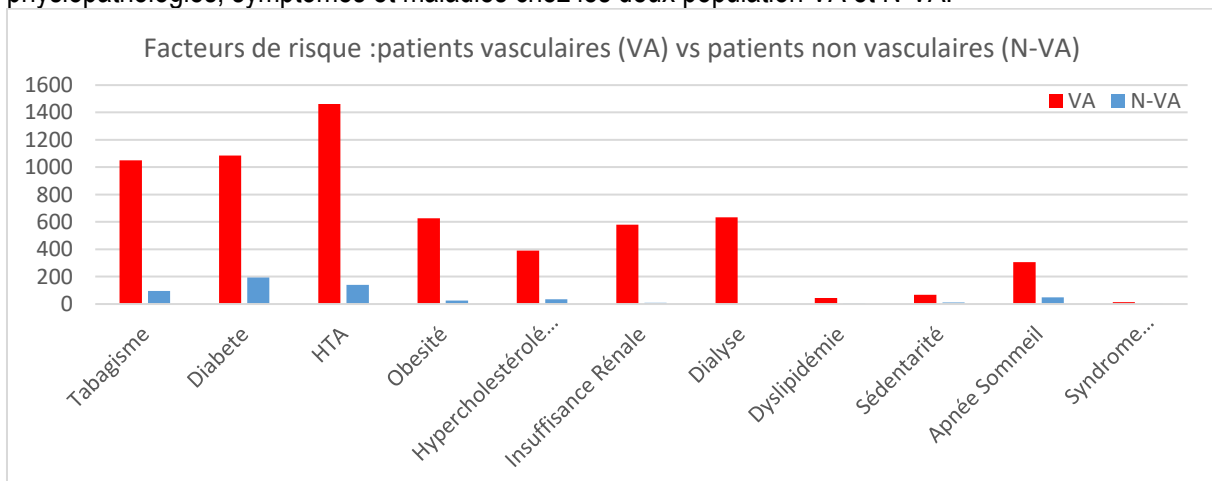


Figure 6 Occurrences des facteurs de risque chez des patients vasculaires et des patients non vasculaires

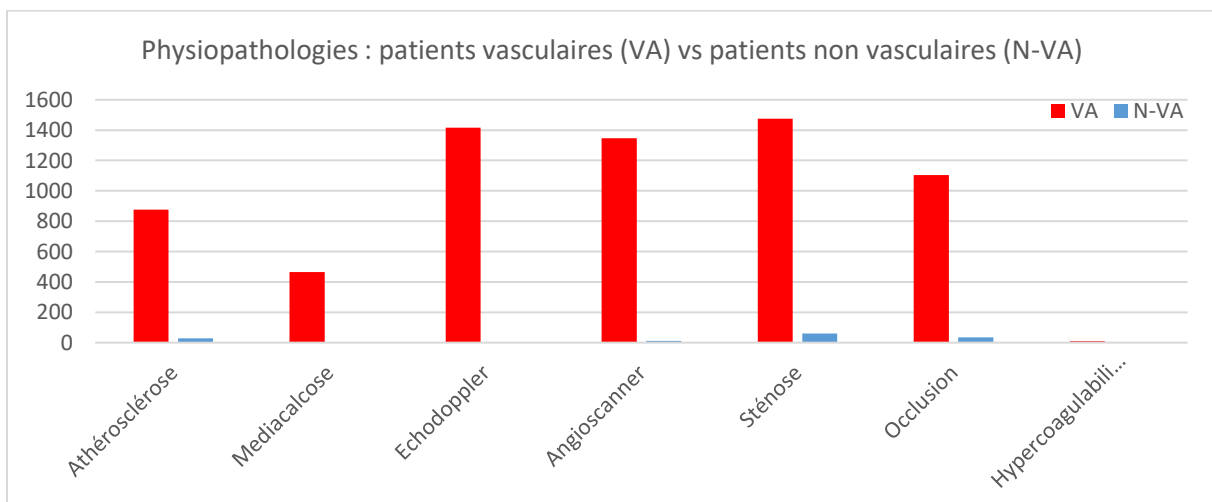


Figure 7 Occurrences des physiopathologies chez des patients vasculaires et des patients non vasculaires

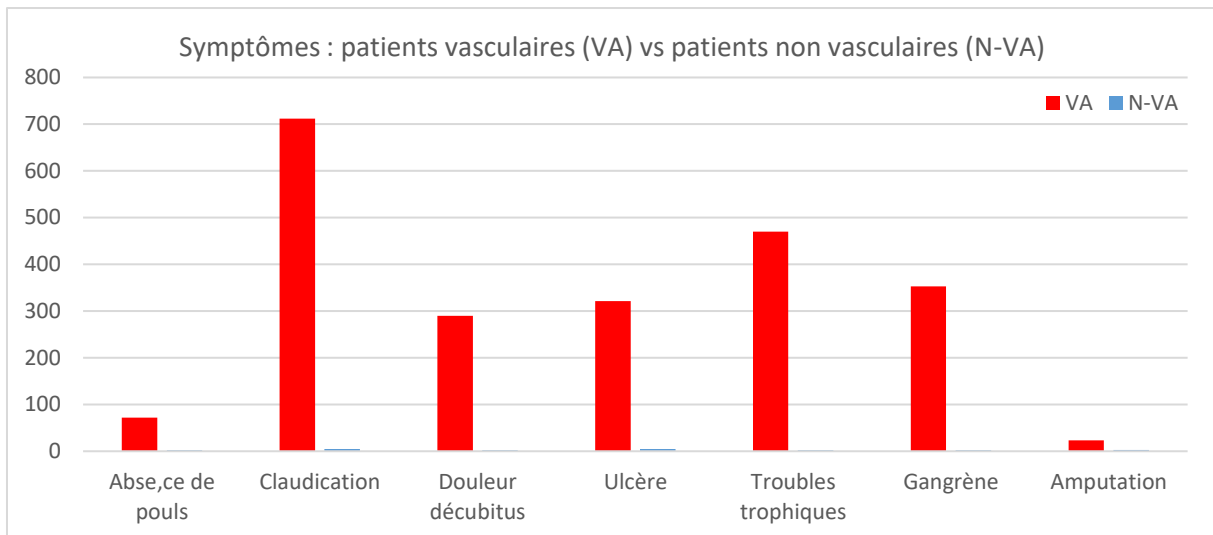


Figure 8 Occurrences des symptômes chez des patients vasculaires et des patients non vasculaires

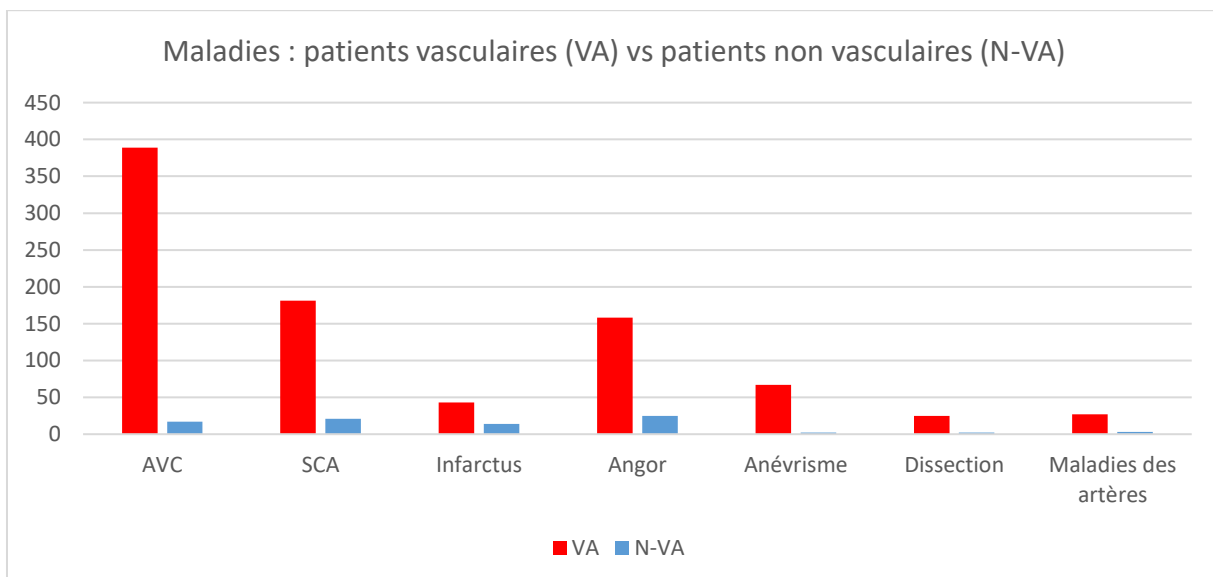


Figure 9 Occurrences des maladies chez des patients vasculaires et des patients non vasculaires

Les histogrammes ci-dessus confirment la spécificité des critères de risque présentés dans la Figure 5 aux pathologies vasculaires. Dans la figure 6, nous remarquons que l'hypertension artérielle, le diabète, le tabagisme, l'insuffisance rénale, la dialyse, l'obésité et l'hypercholestérolémie sont les facteurs de risque les plus communs à la population vasculaire. En comparaison avec la population non-vasculaire, ces facteurs de risques sont moins présents. Le même constat est réalisé pour les physiopathologies, les symptômes et les maladies. Les physiopathologies les plus communes sont la présence d'un échodoppler, un angioscanner, la sténose, l'occlusion et l'athérosclérose. Les symptômes les plus courants chez la population vasculaire sont la claudication, les troubles trophiques, la gangrène et les douleurs du décubitus. Enfin, pour les maladies, les AVC, les syndromes coronariens aigus et les angines de poitrine (angor) sont les plus récurrents. Nous avons également réalisé des combinaisons de facteurs de risques pour estimer l'impact de certaines agrégations sur les pathologies vasculaires. La figure 10 ci-dessous illustre les occurrences du tabac avec les autres facteurs de risque.

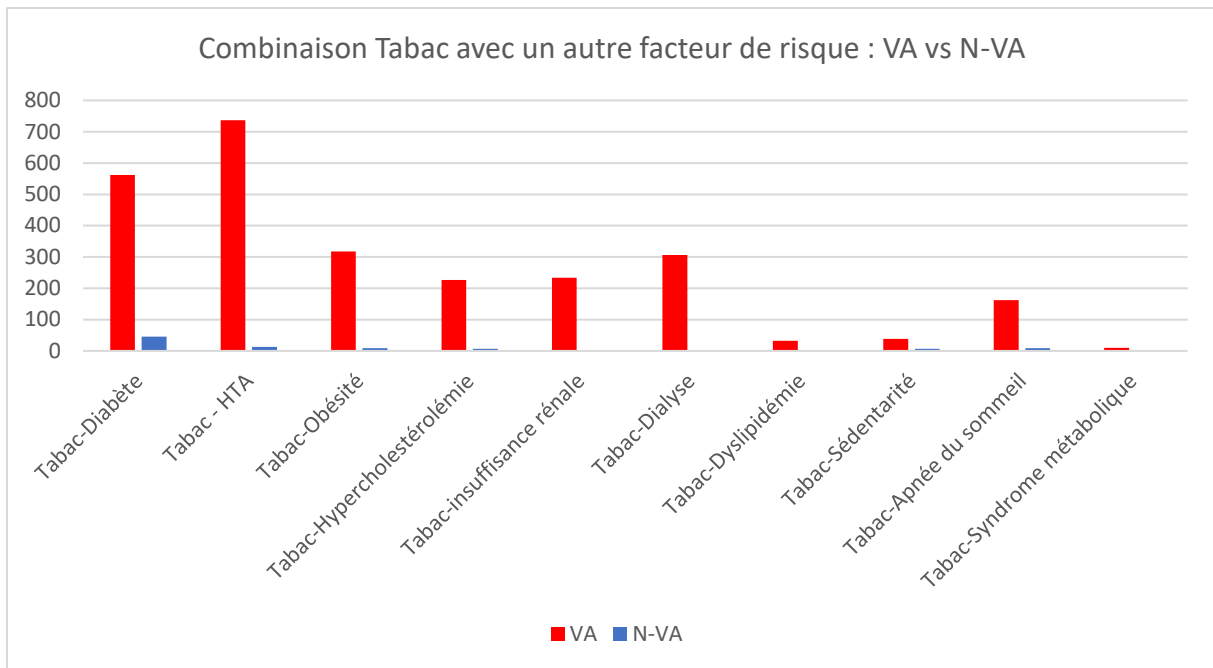


Figure 10 Occurrences de la combinaison Tabac - autre facteur chez deux populations VA - NVA

Rappelons que la présence de chaque critère de risque associe au patient un score. A ce stade, nous considérons que les facteurs de risques associent un score de 1, les physiopathologies, un score de 2, les symptômes, un score de 3 et les maladies, un score de 4. La somme de ces scores est corrélée à l'âge et au sexe de chaque patient. L'ensemble des patients est ensuite classé dans quatre classes distinctes allant du score le moins élevé pour la classe 1 au score le plus élevé pour la classe 4. Les différentes classes représentent la gravité d'un cas et/ou la probabilité qu'un patient soit éligible à un dépistage d'une pathologie vasculaire. La décision finale revient bien évidemment aux chirurgiens vasculaires de l'hôpital. La figure 11 représente la répartition des patients dans les différentes classes pour un ensemble de 3970 patients. Notons que la classe 1 comporte le plus grand nombre de patients. 6% des patients n'ont pas pu être classifiés par manque de données pertinentes.

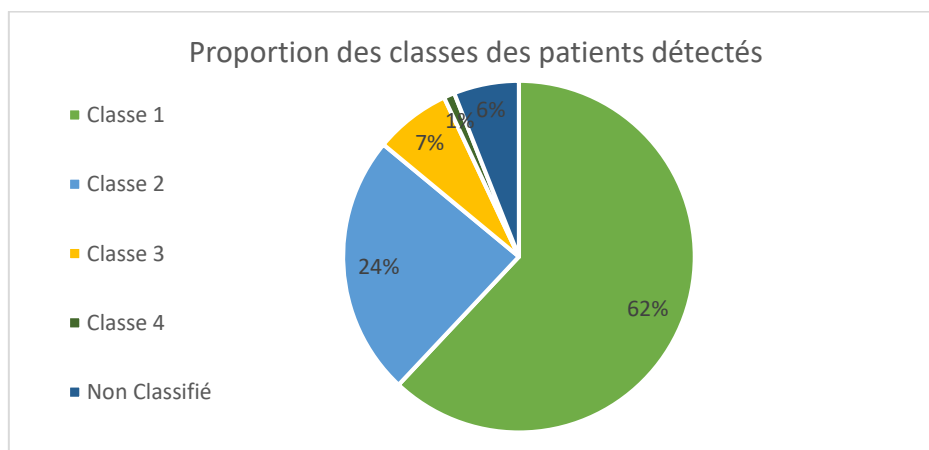


Figure 11 Répartition des patients selon les classes de risque vasculaire

Un échantillon d'une centaine de patients détectés a été présenté aux chirurgiens vasculaires et pré-validés compte tenu des critères de risque de chaque patient. Cet échantillon a été sélectionné de manière à être représentatif pour chaque classe de risque. Rappelons que le choix entre la prévention primaire ou secondaire a été réalisé sur la base de la présence ou l'absence de maladies.

4 Conclusion et perspectives

Pour conclure, le travail présenté dans cet article présente les premières briques réalisées dans le cadre du projet passeport vasculaire. Ce travail a consisté en l'analyse des données rétrospectives des patients ayant subi des chirurgies vasculaires au CHV pour détecter un ensemble de critères de risque et les hiérarchiser avec les experts de l'hôpital. Ces critères d'entrée sont issus de l'analyse des données PMSI ainsi que les courriers disponibles dans l'entrepôt de données de l'hôpital. Cette analyse a été réalisée avec des techniques de TALN et des arbres de décision. Sur la base de ces critères de risque, les règles d'experts ont été traduites en règles logiques pour détecter 8250 patients éligibles à un dépistage de pathologies vasculaires. Un échantillon des patients détectés a été évalué et pré-validé par les médecins de l'hôpital. A chaque patient, un score de gravité a été associé, sur la base de ce score, le patient a été classifié. Ces quatre classes ont été sélectionnées intuitivement.

La suite de ce travail consistera à améliorer la scorification des facteurs de risque et trouver des corrélations entre chaque facteur de risque et chaque physiopathologie, symptôme ou maladie sous-jacente. Nous travaillons également sur l'étude de l'importance de chaque critère de risque et son impact précis sur l'apparition de pathologies vasculaires. D'autres techniques ont été exploitées également sur les mesures biologiques ainsi que le séquençement des unités médicales mais n'ont pas été présentées dans cet article.

Références

- E.Faure. (2015). Le programme de médicalisation des systèmes d'information (PMSI). *caducee.net*.
- Fédération française de cardiologie. (2017). <https://www.fedecardio.org/Je-m-informe/Reduire-le-risque-cardio-vasculaire/les-facteurs-de-risque-cardio-vasculaires>. *fede cardio*.
- Ford, E. G. (2004). The distribution of 10-year risk for coronary heart disease among US adults: findings from the National Health and Nutrition Examination Survey III. . *Journal of the American College of Cardiology*, 43(10), 1791-1796., 10.
- Motwani, M. D.-M. (2016). Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *European heart journal*, 38(7)., pp. 500-507.
- OMS, 2. (2017). Maladies cardiovasculaires. Dans OMS, [https://www.who.int/fr/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/fr/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)).
- Rajmohan, K. P. (2014). Prediction and Diagnosis of Cardio Vascular Disease--A Critical Survey. In *2014 World Congress on Computing and Communication Technologies* (pp. 246-251). *IEEE*.
- Sun, Y. V. (2008). Application of machine learning algorithms to predict coronary artery calcification with a sibship-based design. *Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society*,, pp. 350-360.