

Original Article

Aide à la décision par les réseaux de neurones artificiels : cas de l'infarctus du myocarde au service des urgences cardiologiques à Oran

M. RAIHAH¹, M.F. MESLI¹, A. ABDELOUAHAB¹, H.M. ALI LAHMAR²

¹ Service de Biostatistique, faculté de médecine d'Oran.

² Service de cardiologie, Etablissement Hospitalier Universitaire d'Oran.

Résumé

Introduction : l'infarctus du myocarde (IDM) constitue une urgence cardiologique dont le diagnostic au service d'urgence demeure un problème clinique. Dans cette étude, un réseau neuronal artificiel (RNA) a été utilisé dans la prédiction de l'IDM chez les patients qui consultent aux urgences pour douleur thoracique.

Patients et méthodes : les données des patients ont été aléatoirement réparties en séries de formation et de test. Le modèle de réseau neuronal a été construit sur la base des données de formation et testés sur les observations du groupe test. Les performances diagnostiques du modèle prédictif ont été analysées grâce à la courbe ROC (Receiver Operating Curve). **Résultats :** l'enquête a porté sur 540 patients consultant pour douleur thoracique d'origine non traumatique. Le diagnostic de l'IDM a été retenu chez 118 patients (21,8 %). En utilisant les observations du groupe test, le modèle prédictif neuronal avait une surface sous les courbes ROC de 97,2 %, une sensibilité de 81,1 % et une spécificité de 96%. **Conclusion :** cette étude a montré que les RNA ont de bonnes capacités prédictives. Ils peuvent aider les médecins au service des urgences à prendre des décisions en produisant rapidement des données pertinentes.

Mots clés: Réseau de neurones artificiels. Infarctus du myocarde. Douleur thoracique. Diagnostic. Prédiction.

Received date : 04/03/19

Accepted date :

1. Introduction

Malgré les progrès réalisés dans le développement de nouveaux biomarqueurs cardiaques [1, 2], de techniques d'imagerie [3], et de méthode de stratification du risque [4], les médecins ne sont toujours pas en mesure d'identifier avec précision tous les patients souffrant d'un infarctus du myocarde (IDM) aux urgences cardiologiques. Bien que l'impression clinique du médecin soit un indicateur très sensible de l'infarctus, entre 4 % et 11,8 % des patients se présentant à l'urgence avec un infarctus ne seront pas diagnostiqués et seront renvoyés de l'hôpital sans traitement [5, 6]. De même, plus de 80% des patients admis dans des unités de soins coronariens

consécutivement à des indications suggérant un IDM seront renvoyés sans pouvoir confirmer un diagnostic d'infarctus [7, 8]. Comme la majorité des décès d'origine cardiaque surviennent en phase pré-hospitalière [9-11], une détection précoce des premiers symptômes d'une cardiopathie ischémique permettrait très probablement d'assurer une meilleure prise en charge et en conséquence adéquate, d'améliorer la qualité de vie des patients et de réduire le coût pour la société. Plusieurs études épidémiologiques ont en effet montré l'impact potentiel en termes de mortalité et de morbidité cardiaque d'une diminution de la durée entre l'apparition des symptômes et la prise en charge effective des patients [11, 12].

Il faudrait en pratique réduire cette durée à moins de deux heures contre quatre en moyenne actuellement [11, 13-15]

Or, le problème n'est pas simple ; même les meilleurs experts se trompent parfois. En outre, l'ischémie cardiaque peut être instable : présente à domicile du patient, elle peut disparaître à l'arrivée au service des urgences, pour réapparaître quelques heures après quand le patient est revenu à domicile ; d'où l'intérêt des solutions de prévision des infarctus.

Un certain nombre d'approches ont été proposées pour améliorer la pertinence du diagnostic de l'IDM par les cliniciens [16-21]. À ce jour, aucune de ces approches n'a été largement adoptée. Il a été suggéré que les médecins n'utilisent une méthode qui vise à améliorer la précision du diagnostic, que si elle est facile à utiliser et qu'elle améliore considérablement et constamment leurs performances [22, 23].

Ces dernières années, l'utilisation des réseaux de neurones artificiels (RNA) s'est développée dans de nombreuses disciplines notamment en médecine. Ils sont surtout appliqués pour résoudre des problèmes de classification et de prédiction [24]. Dans le cadre du traitement des données, les RNA constituent une méthode d'approximation de systèmes complexes, grâce à leur capacité d'être un approximateur universel, ont prouvé leur capacité à extraire de données d'expérimentation des modèles performants, sans avoir à effectuer d'hypothèse sur la forme générale de ces derniers [25]. Les RNA sont également applicables dans toutes les situations où il existe une relation non linéaire entre une variable prédictive et une variable prédite [26].

Le diagnostic clinique et la prévision des maladies se font à partir de l'analyse de symptômes et de facteurs de risque. Or, l'interaction entre ces éléments est souvent complexe. Un mauvais diagnostic peut conduire au décès de l'individu. Il est donc nécessaire de pouvoir évaluer le risque avec

le plus de précision possible. Les RNA sont de plus en plus utilisés dans cette discipline comme outil d'aide à la décision [26].

Comme une association linéaire entre les variables associées à l'IDM ne peut pas être considérée comme admise, cette étude a pour objectif d'élaborer, à l'aide du réseau de neurones artificiels, un modèle de prédiction d'IDM à partir des données disponibles lors de la présentation des patients aux urgences cardiologiques de l'établissement hospitalier universitaire (EHU) d'Oran.

2. Patients et méthodes

2.1 Population d'étude

Il s'agit d'une étude réalisée au service des urgences cardiologiques de l'EHU d'Oran, Algérie, entre janvier 2015 et décembre 2015. Les patients, âgés de plus de 18 ans et consultant aux urgences pour des douleurs thoraciques ont été inclus dans l'étude. Nous avons exclu les patients présentant une douleur thoracique d'origine traumatique. Au total, 540 patients ont été inclus dans l'étude.

2.2 Recueil des données

L'enquête a été réalisée au moyen d'un questionnaire qui a porté sur l'identité des patients, leurs antécédents personnels d'HTA, de diabète, de dyslipidémie et leurs antécédents coronariens personnels et familiaux. Les habitudes tabagiques et la durée des symptômes ont été également rapportées. Le délai entre l'apparition des symptômes et la consultation a été définie par trois classes : moins de 6 heures, de 6 à 12 heures et supérieure à 12 heures. Les niveaux d'exposition au tabac ont été classés en trois catégories : non-fumeurs, anciens fumeurs ceux qui ont cessé de fumer pendant plus d'un an et fumeurs actuels ceux qui ont fumé au cours des 12 mois précédents.

Les patients ont été examinés avec prise de la pression artérielle (systolique et diastolique) et de la fréquence cardiaque. Aussi, un électrocardiogramme (ECG) a été réalisé pour chaque patient à la recherche de sus-décalage du segment ST, de sous-décalage du segment ST et d'onde Q de nécrose.

Tous les patients avec un diagnostic présumé d'IDM ont été admis dans une unité de soins intensifs de cardiologie et suivis jusqu'à la sortie d'hôpital. Le diagnostic d'IDM était basé sur la présence d'au moins deux des critères suivants [27] : douleur angineuse intense prolongée supérieure à 30 minutes, modification du segment ST et/ou des ondes T évocatrices d'ischémie transmurale et élévation de la troponine.

2.3 Analyse prédictive

Le RNA utilisé dans cette étude pour la prédiction de l'IDM était le perceptron multicouche à rétropropagation [28]. L'algorithme de rétropropagation calcule l'erreur quadratique moyenne qui mesure l'erreur entre la sortie fournie par le réseau et la sortie désirée. Disposant d'un ensemble d'exemples qui sont des couples (entrées, sorties désirées), le réseau est d'abord initialisé, c'est à dire que les poids synaptiques sont donnés de façon aléatoire pour permettre au réseau de commencer son apprentissage. A chaque étape, un exemple est présenté en entrée et le réseau calcule une sortie. Ensuite l'erreur est calculée en comparant la sortie calculée à la sortie attendue. Cette erreur est ensuite rétropropagée dans le réseau, donnant lieu à une modification de chaque poids, ayant participé à l'erreur constatée. Ce processus est répété en présentant successivement chaque exemple. L'apprentissage consiste à minimiser l'erreur quadratique moyenne commise sur l'ensemble des exemples.

Pour réduire la taille de l'échantillon des données et déterminer l'ensemble optimal de variables d'entrée, des analyses de sensibilité des réseaux qualifiés ont été exécutées principalement pour prioriser les variables dans l'ensemble des données [29]. En détail, cette méthode examine chaque variable d'entrée disponible en employant un réseau neuronal. Pour chaque variable retirée de la liste d'entrée, l'erreur quadratique moyenne de sortie est calculée. Seules les variables qui ont comme conséquence une détérioration de la performance du modèle en étant abandonnées sont maintenues dans l'architecture finale du

réseau. Plus l'erreur est faible, meilleure est la performance du réseau.

Étant donné que les modèles de réseaux de neurones se construisent par apprentissage à partir d'un certain nombre d'observations, on a réparti aléatoirement les données des patients en séries de formation et du test. On a utilisé 70 % des observations (378 patients) pour la formation (apprentissage) et 30 % pour le test (162 patients), et ceci afin de tester la capacité prédictive réelle du réseau. Afin d'éviter le surentraînement du réseau, une base de validation de 108 patients prélevée sur la base d'apprentissage a permis d'interrompre le processus de modification des poids du réseau de neurones lorsque l'erreur de validation quittait son minimum.

L'adéquation du modèle prédictif a été vérifiée par le test du Chi² de Hosmer-Lemeshow [30]. Un modèle est bien calibré lorsque les probabilités prédites du modèle ne sont statistiquement pas significativement différentes ($p > 0,05$) des fréquences observées. Dans cette étude, les patients ont été divisés en décile de probabilité prédite par les modèles, et un test de Chi² à 8 degrés de liberté entre les effectifs attendus par la probabilité prédite et le nombre observé d'effectif dans chaque décile a été réalisé. Les performances diagnostiques du modèle prédictif ont été analysées grâce à la courbe ROC (Receiver Operating Curve). Dans l'analyse ROC, la performance diagnostique (prédiction de l'IDM) est rapportée en termes de deux index à savoir la fraction vrai-positifs (sensibilité) et la fraction faux-positifs (1-spécificité). La surface sous la courbe (AUC) a été calculée par la méthode de Hanley et McNeil [31]. Aussi, la sensibilité, la spécificité et les valeurs prédictives positives et négatives ont été calculées.

2.4 Statistiques

Les variables qualitatives étaient exprimées en pourcentage et les variables quantitatives en moyenne. Le test du Chi² a été utilisé pour comparer les pourcentages et le test t de

Student pour comparer les moyennes pour un seuil de significativité à 5 %. Toutes les analyses ont été effectuées à l'aide du logiciel SAS JMP Pro version 10.

3 Résultats

3.1 Description de la population d'étude

L'enquête a porté sur 540 patients âgés entre 24 et 92 ans. Leurs caractéristiques démographiques et cliniques sont présentées au tableau 1. Le sex-ratio était de 1,2 (54,4 % d'hommes) et la moyenne d'âge de 57,1 ans (écart-type [ET] : 12,3). Sur l'ensemble des patients, 22,6 % étaient diabétiques, 29,4 % étaient hypertendus et 17,4 % avaient une dyslipidémie. Trente-neuf patients (7,2 %) avaient des antécédents coronariens personnels et 18,3 % des patients avaient des antécédents coronariens familiaux. Les anciens fumeurs présentaient 11,9 % de notre population d'étude, tandis que 12,8 % des patients fumaient encore. La proportion des patients qui avaient un délai < 6 heures entre le début des symptômes et la consultation était de 36,9 %. Les pressions systoliques et diastoliques moyennes étaient de 129 mmHg (ET : 19,7) et de 75,1 mmHg (ET : 12,7) respectivement. La fréquence cardiaque moyenne était de 75,5 bpm (ET : 13,4).

Concernant les données du tracé ECG, 31,3 % des patients avaient un sus-décalage du segment ST, 15,4 % des patients avaient un sous décalage du segment ST et 20,9 % des patients avaient une onde Q.

Tableau 1. Caractéristiques des patients inclus dans l'étude (n = 540)

Comme présenté dans le tableau 2, le diagnostic de l'IDM a été retenu chez 21,8 % des patients. L'IDM était significativement associé au sexe masculin ($p = 0,008$), au diabète ($p = 0,02$), à l'HTA ($p = 0,001$), à la dyslipidémie ($p < 10^{-3}$) et aux antécédents coronariens personnels ($p = 0,009$). Par contre, aucune différence n'était retrouvée entre les antécédents coronariens familiaux la survenue de l'IDM. Aussi, l'âge moyen, la PAS moyenne et la PAD moyenne ne différaient pas significativement entre les groupes IDM et sans IDM. Enfin, la survenue de l'IDM était significativement associée au sus-décalage du segment ST ($p < 10^{-3}$), au sous décalage du segment ST ($p = 0,023$) et à l'apparition d'une onde Q sur le tracé ECG ($p < 10^{-3}$).

Caractéristiques	Effectif (%)	Moyenne (ET)
Age (années)		57,1 (12,3)
Sexe masculin	294 (54,4)	
Diabète	122 (22,6)	
HTA	159 (29,4)	
Dyslipidémie	94 (17,4)	
Antécédents coronariens personnels	39 (7,2)	
Antécédents coronariens familiaux	99 (18,3)	
Anciens fumeurs	64 (11,9)	
Fumeurs actuels	69 (12,8)	
Symptômes < 6 heures	199 (36,9)	
Symptômes entre 6 et 12 heures	31 (24,3)	
PAS (mmHg)		29 (19,7)
PAD (mmHg)		5,1 (12,7)
Fréquence cardiaque (bpm)		5,5 (13,4)
Sus décalage ST	169 (31,3)	
Sous décalage ST	83 (15,4)	
Onde Q	113 (20,9)	

Tableau 2. Comparaison entre les patients ayant un IDM et les patients non IDM

Variables	IDM n = 118	Pas d'IDM n = 422	p
Moyenne d'âge (ET) (années)	58,9 (12,3)	56,6 (12,3)	0,069
Sexe masculine n (%)	77 (65,3)	217 (51,4)	0,008
Diabète n (%)	36 (30,5)	86 (20,4)	0,020
HTA n (%)	49 (41,5)	110 (26,1)	0,001
Dyslipidémie n (%)	38 (32,2)	56 (13,3)	$< 10^{-3}$
Antécédents coronariens personnels n (%)	15 (12,7)	24 (5,7)	0,009
Antécédents coronariens familiaux n (%)	21 (17,8)	78 (18,5)	0,865
Anciens fumeurs n (%)	22 (18,6)	42 (10,0)	0,011
Fumeurs actuels n (%)	26 (22,0)	43 (10,2)	$< 10^{-3}$
Symptômes < 6 heures n (%)	59 (50,0)	140 (33,2)	0,001
Symptômes entre 6 et 12 heures n (%)	29 (24,6)	102 (24,2)	0,964
Moyenne de la PAS (ET) (mmHg)	129,3 (24,6)	129 (18,2)	0,881
Moyenne de la PAD (ET) (mmHg)	76,2 (14,8)	74,8 (12)	0,317
Moyenne de la fréquence cardiaque (ET) (bpm)	83,2 (19,3)	73,3 (10,3)	$< 10^{-3}$
Sus décalage ST n (%)	71 (60,2)	98 (23,2)	$< 10^{-3}$
Sous décalage ST n (%)	26 (22,0)	57 (13,5)	0,023
Onde Q n (%)	71 (60,2)	42 (10,0)	$< 10^{-3}$

2. Analyse prédictive par réseau de neurones artificiels

Le réseau utilisé dans cette étude était composé de trois couches : une couche d'entrée, une couche cachée et une couche de sortie. L'architecture de ce réseau est illustrée sur la figure 1.

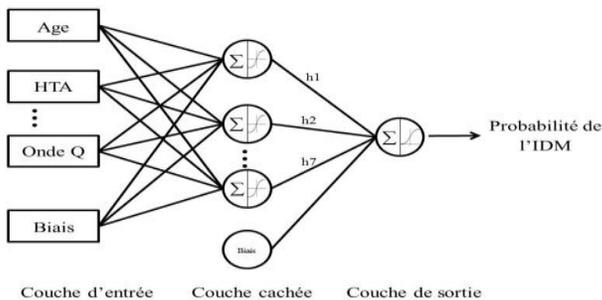


Figure 1 : Architecture du réseau de neurones pour le diagnostic de l'IDM.

La sensibilité des variables utilisées dans le modèle neuronal est résumée au tableau 3. Les erreurs ont été comparées à l'erreur de base de 0,266. On constate que six variables (sexe, diabète, antécédents coronariens familiaux, durée des symptômes, PAS et PAD) dégradent le modèle. De ce fait, la couche d'entrée du réseau est composée de neuf neurones d'entrée (âge, HTA, dyslipidémie, antécédents coronariens personnels, statut tabagique, fréquence cardiaque, sus-décalage du ST, sous-décalage du ST et onde Q), la couche cachée était composée de sept neurones à fonction d'activation de type tangente-sigmoïde et la couche de sortie est composée d'un neurone de sortie à fonction d'activation de type sigmoïde.

Tableau 3. Analyse de sensibilité des variables d'entrée du réseau de neurones utilisé pour le diagnostic de l'IDM

Variables d'entrée	Erreur après suppression
Age	0,285
Sexe	0,260
Diabète	0,260
HTA	0,271
Dyslipidémie	0,275
Antécédents coronariens personnels	0,267
Antécédents coronariens familiaux	0,261
Statut tabagique	0,271
Durée des symptômes	0,265
PAS	0,238
PAD	0,252
Fréquence cardiaque	0,284
Sus-décalage du ST	0,281
Sous-décalage du ST	0,268
Onde Q	0,308

Le tableau 4 donne les poids synaptiques du perceptron multicouche utilisé pour la prédiction de l'IDM. Ces coefficients représentent les connections entre les sept

neurones de la couche cachée (h1, ..., h7) et tous les neurones d'entrée.

Tableau 4. Poids synaptiques de la couche d'entrée du réseau de neurones utilisé pour le diagnostic de l'IDM

Variables	h1	h2	h3	h4	h5	h6	h7
Age	0,001286	0,005946	0,02182	0,019058	0,022109	-0,00558	-0,0423
HTA	-0,53776	0,99226	0,642945	-0,03336	0,048452	-0,10124	-0,02713
Dyslipidémie	0,469631	0,194498	-0,23755	0,408117	-0,17321	-0,58622	0,31005
Antécédents coronariens personnels	-0,81086	-1,82473	-2,11179	0,976973	-0,99473	-0,4927	1,035461
Consommation de tabac	-0,00402	-0,14336	0,294542	-0,64644	-0,34212	-0,67523	0,108493
Fréquence cardiaque	0,000552	0,012145	0,028881	-0,00688	-0,01353	-0,03798	0,007771
Sus-décalage ST	-0,65046	0,56405	0,580157	-0,55495	-0,61382	-1,0432	0,009063
Sous-décalage ST	0,192241	-0,57649	-0,01765	1,210602	0,667809	-0,81258	-0,66893
Onde Q	0,111617	0,419856	0,742007	-1,29188	0,221202	-2,01517	0,2639
Biais	0,222337	-1,8099	-4,85774	0,608589	0,515256	6,366217	1,51612

A chaque nœud caché, une combinaison linéaire pondérée des entrées est additionnée, puis une transformation tangente-sigmoïde est appliquée.

Les poids reliant la couche cachée et le neurone de sortie sont indiqués dans le tableau 5.

Tableau 5. Poids synaptiques de la couche cachée du réseau de neurones utilisé pour le diagnostic de l'IDM.

Neurones cachés	Coefficients
h1	-0,12829
h2	0,07448
h3	0,0584
h4	0,15043
h5	0,039864
h6	-0,59325
h7	0,047955
Biais	0,49511

La probabilité d'avoir l'IDM = $1 / [1 + \exp(-(\text{biais} + \sum h_i \text{nœud caché}_i))]$

La performance diagnostique, exprimée par la surface sous la courbe ROC, est évaluée à 97,2 % (IC 95 % = [95 % - 99,3 %]) (Fig. 2). La sensibilité et la spécificité étaient respectivement de 81,1 % (IC 95 % = [64,5 - 92,0]) et 96 % (IC 95 % = [92,6 - 99,4]). Les VPP et VPN étaient respectivement de 85,7 % (IC 95 % = [69,7 - 95,2]) et 94,5 % (IC 95 % = [90,5 - 98,4]). L'analyse du test de Hosmer-Lemeshow donne une valeur du Chi² de 8,41 (p= 0,394) traduisant une bonne calibration.

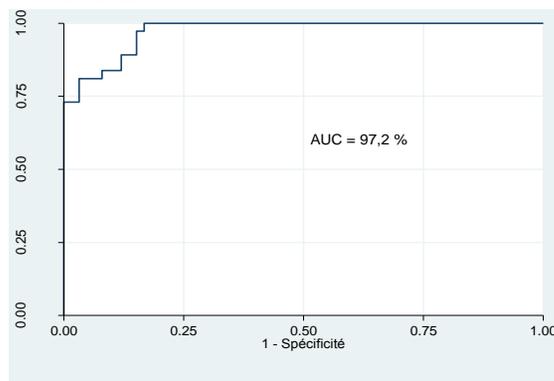


Figure 2 : Performance diagnostique du modèle neuronal.

Discussion

Dans le cadre de l'intelligence artificielle, de nouvelles approches apparaissent donc pour tirer parti de l'évolution des coûts et de la puissance de traitement des ordinateurs face à des problèmes définis de manière plus floue. Des algorithmes non linéaires sont apparus comme le recuit simulé, les algorithmes génétiques, la recherche sur tabous et les machines à vecteur de support [32, 33].

L'intelligence artificielle se propose de reconstituer le raisonnement humain soit à partir d'un apport externe de la connaissance et de l'utilisation de règles dans les systèmes experts, soit par l'imitation du fonctionnement du cerveau dans les réseaux neuronaux [34]. Ce dernier est un graphe orienté et pondéré dont les nœuds sont des automates simples dotés d'un état interne d'activation par lequel ils influencent les autres neurones du réseau.

La famille de RNA la plus utilisée comme outil d'aide à la décision médicale est le perceptron multicouche. À lui seul, ce type de réseau recouvre plus de 95 % des applications scientifiques. Ce type de réseau est très performant pour les problèmes de classification.

Le perceptron multicouche est un modèle de réseau à propagation par couche. L'idée consiste à regrouper les neurones par couches interconnectées. Une première couche appelée couche d'entrée est composée d'un certain nombre de neurones dont la tâche est de recevoir l'information de l'extérieur. Ces informations sont transformées puis transmises aux neurones de la couche intermédiaire, appelée aussi couche cachée, qui vont effectuer certains traitements puis envoyer les résultats vers une dernière couche qui code la variable de discrimination appelée couche de sortie. La présence d'une couche cachée permet de modéliser la fonction de discrimination non linéaire entre les entrées et la sortie.

La difficulté de développer des modèles employant les RNA est qu'aucune théorie généralement admise n'existe pour définir l'architecture du réseau. En particulier la détermination du nombre de couches cachées et du nombre de neurones sur chacune d'entre elle relève encore largement de l'intuition

l'utilisateur et de sa capacité à expérimenter plusieurs architectures afin de retenir celle qui donne les résultats les meilleurs. Plus le réseau est complexe, c'est-à-dire plus il comporte de neurones cachés et de liens synaptiques, plus il est capable de reconnaître les formes qui lui sont présentées à travers l'échantillon d'apprentissage. Mais malheureusement, l'accroissement de la complexité n'améliore pas nécessairement le taux de reconnaissance sur l'échantillon test.

Il faut cependant remarquer que cette critique doit être nuancée, car plusieurs méthodes ont été proposées pour aider l'utilisateur dans cette tâche. Dreiseitl et Ohno-Machado[35] ont montré qu'une seule couche cachée est généralement suffisante pour classer la plus part des données. Quant à la taille de la couche cachée, elle doit être soit égale à celle de la couche d'entrée [36], soit 75 % de celle-ci [37]. Ainsi la couche d'entrée de notre réseau était composée de neuf neurones et la couche cachée de sept neurones (77,8 % de la taille de la couche d'entrée).

Après avoir déterminé l'architecture du réseau, sa mise en œuvre nécessite trois échantillons de données. Le premier, destiné à l'apprentissage servira au paramétrage, le second servira à la validation et le troisième sera un échantillon test destiné à évaluer les capacités de généralisation du réseau.

Durant la phase d'apprentissage, l'erreur commise diminue, jusqu'à tendre asymptotiquement vers zéro si l'architecture du réseau a été correctement choisie. Mais il convient de remarquer que plus l'erreur est faible, c'est-à-dire plus le réseau apprend à reconnaître les formes qui lui sont présentées, moins il risque d'être capable de généralisation. En effet au-delà d'un certain point, si le réseau reconnaît de mieux en mieux les formes appartenant à l'échantillon d'apprentissage, il donne de piètres résultats sur un échantillon inconnu. On dit alors qu'il se spécialise. Or le but d'un tel système est bien de reconnaître des formes inconnues jusqu'alors. Il convient donc d'arrêter la phase d'apprentissage lorsque le

réseau semble donner les meilleurs résultats en généralisation. C'est l'échantillon de validation qui permet de suivre la capacité du système à généraliser.

Pour limiter ce phénomène, notre échantillon global a été découpé en deux échantillons d'apprentissage et du test pour lequel on a prélevé un troisième échantillon à partir de celui d'apprentissage pour la validation. L'apprentissage est stoppé lorsque l'amélioration des résultats n'est plus perceptible que sur le seul échantillon d'apprentissage.

Une autre limite des modèles de réseau neuronal est que des coefficients normalisés et des rapports de chance correspondant à chaque variable ne peuvent pas être facilement calculés et présentés. Le réseau de neurones doit être assimilé à une boîte noire dont seules les informations à l'entrée et à la sortie sont observables [38]. L'utilisateur n'a pas la possibilité, comme dans le cas de la régression logistique, de comprendre l'influence des variables indépendantes sur le classement des observations [39].

Cette difficulté d'interprétation des différentes variables est une des caractéristiques les plus critiquées des réseaux neuronaux. Par ailleurs, comparés aux modèles de régression linéaire et logistique, les RNA sont plus susceptibles de surestimer si trop de paramètres sont laissés libres. Pour contourner cet inconvénient, Baetge et al. [40] préconisent d'utiliser l'analyse de sensibilité. Par cette dernière, l'utilisateur peut explorer la sensibilité de l'information à la sortie en variant une ou plusieurs informations à l'entrée. Le présent travail confirme qu'un RNA peut identifier, avec un degré élevé de précision, à la fois la présence et l'absence d'un IDM chez les patients qui consultent pour douleur thoracique. La nécessité d'améliorer la sensibilité diagnostique d'IDM est devenue primordiale du fait que la morbidité et la mortalité peuvent être réduites par une thérapie bien conduite. Une spécificité améliorée pourrait réduire considérablement les coûts des soins de santé en éliminant les admissions inutiles aux unités de soins intensifs.

Le modèle élaboré avec le RNA avait une bonne valeur diagnostique avec une aire sous la

courbe ROC de 97,2 %. Le modèle était bien calibré. En 2006, Green et al. [41] ont diagnostiqué le syndrome coronarien aigu avec les RNA et la régression logistique. Ils ont constaté que les RNA prédisaient le SCA avec une aire sous la courbe ROC de 81,1 %, une sensibilité de 95 % et une spécificité de 40 %. En utilisant les RNA pour la prédiction d'IDM, Baxt et al. [42] ont trouvé une sensibilité de 94,5 % et une spécificité de 95,9 %.

Dans une récente étude, la sensibilité et la spécificité de l'algorithme neuronal de Berikol et al. [43] étaient de 90 % et 92,86 %, respectivement. Notre modèle neuronal est comparable à ceux de la littérature avec une sensibilité de 81,1 % et une spécificité de 96 %. Par contre, les performances diagnostiques de notre modèle prédictif dépassent celles des médecins aux urgences, dont leur sensibilité est de 73,3 % et leur spécificité est de 81,1 % [44]. Aussi, la VPP de notre modèle (85,7 %) est nettement meilleure que le seuil de décision du clinicien, où 70 % des patients admis à des indications suggérant un IDM pouvaient ne le pas avoir [45].

Notre modèle prédictif n'a pas inclus les enzymes cardiaques. Les marqueurs cardiaques sont utilisés pour le diagnostic final et les incorporer dans le modèle prédictif peut conduire à un biais de prédiction. Comme le démontrent les études de Conti et al. [46], Sanchis et al. [47, 48] et Goldman et al. [49], les critères cliniques (examen clinique et ECG) suffisent à poser le diagnostic de l'IDM.

La combinaison des données de la troponine I avec l'algorithme de Goldman n'a pas amélioré sa capacité prédictive [50]. Aussi, l'utilisation des marqueurs biologiques (CPK-MB, troponine I) n'ont pas amélioré les performances diagnostiques du modèle de Berikol et al [43]. Selon Collinson et al. [51], la troponine T ultrasensible a une excellente spécificité de 96% mais une sensibilité de 78% pour discriminer un IDM d'une douleur thoracique non angineuse. Bien que notre modèle n'ait pas incorporé les données des biomarqueurs cardiaques, sa performance était très bonne.

Notre travail avait des limites. Les patients ont été inclus rétrospectivement et dans un seul

centre, ce qui pouvait affecter la généralisation de notre modèle prédictif. Une étude multicentrique de plus grande échelle aurait plus de puissance. En outre, avant la mise en œuvre pratique, le modèle doit être validé prospectivement dans plusieurs centres. Aussi, Le pourcentage des différents facteurs de risque cardiovasculaire dépend de leur transcription dans les dossiers médicaux aux urgences, car les fiches de recueil ont été remplies à partir de ces données, ce qui constitue un biais de recueil. En effet, le poids et la taille (et donc l'IMC) n'étaient pas notés dans les dossiers médicaux aux urgences et l'interrogatoire des patients est parfois fluctuant. Ceci peut également être dû à un manque de puissance de notre étude.

En conclusion, cette étude a montré qu'un RNA peut prédire l'IDM au service des urgences avec une sensibilité de 81,1 %, une spécificité de 96 % et une surface sous la courbe ROC de 97,2 %. Ainsi, Le réseau de neurones de la présente étude peut être incorporé dans des programmes sur ordinateur et peut détecter au moyen des variables d'entrée la présence de pathologie.

Conflits d'intérêts

Les auteurs déclarent n'avoir aucun conflit d'intérêt.

4 Références

1. Boukili Y. Troponines hypersensibles: vers une nouvelle définition de l'infarctus du myocarde? La Presse Médicale. 2012;41:634-7.
2. Charpentier S, Chenevier-Gobeaux C. Recommandations ESC 2015: exclusion ou confirmation du diagnostic d'infarctus du myocarde en 1 heure avec la troponine T hypersensible. La Presse Médicale. 2016;45:859-64.
3. Charpentier S, Savary D, Lapostolle F, Chouihed T, Bonnefoy E, Manzo-Silberman S, et al. Recommandations de la Société européenne de cardiologie pour la prise en charge des patients présentant un syndrome coronaire aigu sans sus-décalage du segment ST. Annales françaises de médecine d'urgence. 2014;4:56-64

4. Charpentier S, Lauque D. Douleur thoracique et syndromes coronariens aigus: stratégie diagnostique. *Annales françaises de médecine d'urgence*. 2011;1.
5. Lee TH, Rouan GW, Weisberg MC, Brand DA, Acampora D, Stasiulewicz C, et al. Clinical characteristics and natural history of patients with acute myocardial infarction sent home from the emergency room. *The American journal of cardiology*. 1987;60:219-24.
6. Emerson PA, Russell NJ, Wyatt J, Crichton N, Pantin CF, Morgan AD, et al. An audit of doctor's management of patients with chest pain in the accident and emergency department. *The Quarterly journal of medicine*. 1989;70:213-20.
7. Puleo PR, Meyer D, Wathen C, Tawa CB, Wheeler S, Hamburg RJ, et al. Use of a rapid assay of subforms of creatine kinase-MB to diagnose or rule out acute myocardial infarction. *The New England journal of medicine*. 1994;331:561-6.
8. Roberts R, Kleiman NS. Earlier diagnosis and treatment of acute myocardial infarction necessitates the need for a 'new diagnostic mind-set'. *Circulation*. 1994;89:872-81.
9. Luepker RV. Delay in acute myocardial infarction: why don't they come to the hospital more quickly and what can we do to reduce delay? *American heart journal*. 2005;150:368-70.
10. McGinn AP, Rosamond WD, Goff DC, Jr., Taylor HA, Miles JS, Chambless L. Trends in prehospital delay time and use of emergency medical services for acute myocardial infarction: experience in 4 US communities from 1987-2000. *American heart journal*. 2005;150:392-400.
Azzaz S, Charbonnel C, Ajlani B, Cherif G, Convers R, Blicq E, et al., editors. Évolution de la prise en charge interventionnelle et des délais de reperfusion à la phase aiguë de l'infarctus du myocarde avec sus-
- décalage du segment ST. *Annales de cardiologie et d'angiologie*; 2015: Elsevier.
11. Le Breton H. Prise en charge de l'infarctus du myocarde: les délais. *La Presse Médicale*. 2011;40:600-5.
12. Atoui H. Conception de systèmes intelligents pour la télémédecine citoyenne: Villeurbanne, INSA; 2006.
13. Danchin N, Coste P, Ferrières J, Steg P-G, Cottin Y, Blanchard D, et al. Comparison of thrombolysis followed by broad use of percutaneous coronary intervention with primary percutaneous coronary intervention for ST-segment-elevation acute myocardial infarction. *Circulation*. 2008;118:268-76.
14. Wijns W, Kolh P, Danchin N, Di Mario C, Falk V, Folliguet T, et al. Guidelines on myocardial revascularization. *European heart journal*. 2010;31:2501-55.
15. Antman EM, Cohen M, Bernink PJ, McCabe CH, Horacek T, Papuchis G, et al. The TIMI risk score for unstable angina/non-ST elevation MI: A method for prognostication and therapeutic decision making. *Jama*. 2000;284:835-42.
16. Ohman EM, Granger CB, Harrington RA, Lee KL. Risk stratification and therapeutic decision making in acute coronary syndromes. *Jama*. 2000;284:876-8.
17. Xue J, Aufderheide T, Scott Wright R, Klein J, Farrell R, Rowlandson I, et al. Added value of new acute coronary syndrome computer algorithm for interpretation of prehospital electrocardiograms. *Journal of electrocardiology*. 2004;37 Suppl:233-9.
18. Harrison RF, Kennedy RL. Artificial neural network models for prediction of acute coronary syndromes using clinical data from the time of presentation. *Annals of emergency medicine*. 2005;46:431-9.
19. Green M, Björk J, Hansen J, Ekelund U, Edenbrandt L, Ohlsson M, editors. Detection of acute coronary syndromes in chest pain patients using neural network ensembles. *Second International*

21. Conference on Computational Intelligence in Medicine and Healthcare; 2005.
22. Kennedy RL, Harrison RF. Identification of patients with evolving coronary syndromes by using statistical models with data from the time of presentation. *Heart*. 2006;92:183-9.
23. Engle RL, Jr., Flehinger BJ. Why expert systems for medical diagnosis are not being generally used: a valedictory opinion. *Bulletin of the New York Academy of Medicine*. 1987;63:193-823.
24. Kennedy RL, Harrison RF, Marshall SJ. Do we need computer-based decision support for the diagnosis of acute chest pain: discussion paper. *Journal of the Royal Society of Medicine*. 1993;86:31-4.
25. Drew PJ, Monson JR. Artificial neural networks. *Surgery*. 2000;127:3-11.
26. Zribi M, Boujelbene Y. Les réseaux de neurones un outil de sélection de variables: Le cas des facteurs de risque de la maladie du cancer du sein. 2012.
27. Baxt WG. Application of artificial neural networks to clinical medicine. *Lancet*. 1995;346:1135-8.
28. Philippe F. Épidémiologie et définition de l'infarctus du myocarde. *EMC - Cardiologie*. 2014;9:1-6 [Article 11-030-P-04].
29. Rumelhart DE, Hinton GE, Williams RJ. Learning representation by back-propagation errors. *Nature*. 1986;323:533-6.
30. Hunter A, Kennedy L, Henry J, Ferguson I. Application of neural networks and sensitivity analysis to improved prediction of trauma survival. *Computer methods and programs in biomedicine*. 2000;62:11-9.
31. Hosmer D, Lemeshow S. *Applied logistic regression*. New York: Wiley; 1989.
32. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29-36.
33. Coates D, Doherty N, French A. The new multivariate jungle: Computer intensive methods in database marketing. *Journal of Marketing Management*. 1994;10:207-20.
- 34.
35. Bibault J-E, Burgun A, Giraud P. Intelligence artificielle appliquée à la radiothérapie. *Cancer/Radiothérapie*. 2017.
36. Borne P, Benrejeb M, Haggège J. *Les réseaux de neurones: présentation et applications*: Editions OPHRYS; 2007.
37. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*. 2002;35:352-9.
38. Wierenga B, Kluytmans J. Neural nets versus marketing models in time series analysis: A simulation study. *EMAC*; 1994. p. 1139-53.
39. Venugopal V, Baets W. Neural networks and statistical techniques in marketing research: A conceptual comparison. *Marketing Intelligence & Planning*. 1994;12:30-8.
40. Hart A, Wyatt J. Evaluating black-boxes as medical decision aids: issues arising from a study of neural networks. *Medical informatics = Medecine et informatique*. 1990;15:229-36.
41. Baxt WG, White H. Bootstrapping confidence intervals for clinical input variable effects in a network trained to identify the presence of acute myocardial infarction. *Neural Comput*. 1995;7:624-38.
42. Baetge J, Manolopoulos PR. Bilanz-Ratings zur Beurteilung der Unternehmensbonität-Entwicklung und Einsatz des BBR Baetge-Bilanz-Rating im Rahmen des Benchmarking. *Die Unternehmung*. 1999;53:351-71.
43. Green M, Bjork J, Forberg J, Ekelund U, Edenbrandt L, Ohlsson M. Comparison between neural networks and multiple logistic regression to predict acute coronary syndrome in the emergency room. *Artificial intelligence in medicine*. 2006;38:305-18.
44. Baxt WG, Shofer FS, Sites FD, Hollander JE. A neural computational aid to the diagnosis of acute myocardial infarction. *Annals of emergency medicine*. 2002;39:366-73.
45. Berikol GB, Yildiz O, Ozcan IT. Diagnosis of Acute Coronary Syndrome with a

- Support Vector Machine. Journal of medical systems. 2016;40:84.
46. Baxt WG, Skora J. Prospective validation of artificial neural network trained to identify acute myocardial infarction. Lancet. 1996;347:12-5.
 47. Karlson BW, Herlitz J, Wiklund O, Richter A, Hjalmarson A. Early prediction of acute myocardial infarction from clinical history, examination and electrocardiogram in the emergency room. The American journal of cardiology. 1991;68:171-5.
 48. Conti A, Vanni S, Del Taglia B, Paladini B, Magazzini S, Grifoni S, et al. A new simple risk score in patients with acute chest pain without existing known coronary disease. The American journal of emergency medicine. 2010;28:135-42.
 49. Sanchis J, Bodí V, Núñez J, Bertomeu-González V, Gómez C, Bosch MJ, et al. New risk score for patients with acute chest pain, non-ST-segment deviation, and normal troponin concentrations: a comparison with the TIMI risk score. Journal of the American College of Cardiology. 2005;46:443-9.
 50. Sanchis J, Bodí V, Núñez J, Núñez E, Bosch X, Pellicer M, et al. Identification of very low risk chest pain using clinical data in the emergency department. International journal of cardiology. 2011;150:260-3.
 51. Goldman L, Cook EF, Johnson PA, Brand DA, Rouan GW, Lee TH. Prediction of the need for intensive care in patients who come to emergency departments with acute chest pain. New England Journal of Medicine. 1996;334:1498-504.
 52. Limkakeng A, Jr., Gibler WB, Pollack C, Hoekstra JW, Sites F, Shofer FS, et al. Combination of Goldman risk and initial cardiac troponin I for emergency department chest pain patient risk stratification. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine. 2001;8:696-702.
 53. Collinson P, Gaze D, Thokala P, Goodacre S. Randomised assessment of treatment using panel assay of cardiac markers—contemporary biomarker evaluation (RATPAC CBE). 2013.