

UNIVERSITÉ DE LILLE
FACULTÉ DE MÉDECINE HENRI WAREMBOURG
Année : 2024

THÈSE POUR LE DIPLÔME D'ÉTAT
DE DOCTEUR EN MÉDECINE

Validation externe d'un modèle obtenu par Machine Learning sur des données échographiques pour évaluer le pronostic des patients admis enUSIC : cohorte lilloise prospective.

Présentée et soutenue publiquement le 27 juin 2024 à 18h
au Pôle Recherche
par **Jean GRENIER-BOLEY**

JURY

Président :

Monsieur le Professeur Christophe BAUTERS

Assesseurs :

Monsieur le Professeur Gilles LEMESLE

Monsieur le Professeur Augustin COISNE

Directeur de thèse :

Monsieur le Docteur Samy AGHEZZAF

Avertissement

La Faculté n'entend donner aucune approbation aux opinions émises dans les thèses :
celles-ci sont propres à leurs auteurs.

Table des matières

ABREVIATIONS	6
ABSTRACT FRANÇAIS	9
I. INTRODUCTION	13
A. LES SOINS INTENSIFS CARDIOLOGIQUES	13
1. Bref historique	13
2. Épidémiologie	14
<i>a. Épidémiologie initiale dans les coronary care units et son évolution</i>	
<i>b. Épidémiologie actuelle des unités de soins intensifs cardiologiques</i>	
B. ÉVALUER LA GRAVITE DES PATIENTS	16
1. Exemples de score	16
<i>a. Score de GRACE</i>	
<i>b. Score TIMI</i>	
<i>c. Score ACUTE-HF</i>	
<i>d. Score PESI</i>	
<i>e. Scores SOFA et APACHE IV</i>	
<i>f. Score M-CARS</i>	
2. Apport de l'échographie transthoracique	25
C. INTELLIGENCE ARTIFICIELLE ET MACHINE LEARNING	27
1. Définition et principes	27
<i>a. Apprentissage automatique supervisé</i>	
<i>b. Apprentissage automatique non supervisé</i>	
2. Apport du Machine Learning en cardiologie	31
D. VERS UN NOUVEAU SCORE	32
1. ADDICT-ICCU	32
<i>a. Description de la cohorte et objectif principal</i>	
<i>b. Caractéristiques de la population</i>	

2.	Élaboration du score	33
	<i>a. Sélection des variables</i>	
	<i>b. Construction du modèle</i>	
	<i>c. Poids des variables dans la prédiction finale</i>	
3.	Objectif de travail	44
II.	METHODES	44
A.	LA POPULATION DE L'ETUDE	44
B.	COLLECTE DES DONNEES	45
C.	CRITERE DE JUGEMENT PRINCIPAL	46
D.	STATISTIQUES	46
III.	RESULTATS	47
A.	DESCRIPTION DE LA POPULATION	47
B.	CARACTERISTIQUES ECHOCARDIOGRAPHIQUES DE LA POPULATION	50
C.	OCCURRENCE DU CRITERE DE JUGEMENT PRINCIPAL	51
D.	MODELE DE MACHINE LEARNING	51
IV.	DISCUSSION	53
V.	CONCLUSION	57
VI.	BIBLIOGRAPHIE	58
VII.	ANNEXES	62

Abréviations

AHF	<i>Acute heart failure</i> , soit insuffisance cardiaque aiguë
AIT	Accident ischémique transitoire
APACHE	<i>Acute Physiology and Chronic Health Evaluation</i>
AUC	<i>Area under curve</i> , soit aire sous la courbe
AVC	Accident vasculaire cérébral
BNP	<i>Brain natriuretic peptide</i> , soit peptide natriurétique cérébral
BUN	<i>Blood urea nitrogen</i> , soit azote uréique sanguin
CCU	<i>Coronary care unit</i>
CHU	Centre Hospitalier Universitaire
CI	<i>Confidence interval</i> , soit intervalle de confiance
CICU	<i>Cardiac intensive care unit</i>
CJP	Critère de jugement principal
CPK	Créatine phosphokinase
DS	Déviations standard
ECG	Électrocardiogramme
ESC	<i>European Society of Cardiology</i> , soit Société Européenne de Cardiologie
ETT	Échocardiographie transthoracique
FEVG	Fraction d'éjection ventriculaire gauche
FiO ₂	<i>Fraction of inspired oxygen</i> , soit fraction inspirée en oxygène
GRACE	<i>Global Registry of Acute Coronary Events</i>
HF	<i>Heart failure</i> , soit insuffisance cardiaque

IMC	Indice de masse corporelle
IT	Insuffisance tricuspide
ITV	Intégrale temps-vitesse
LVEF	<i>Left ventricular ejection fraction</i> , soit fraction d'éjection ventriculaire gauche
LVOT	<i>Left ventricular outflow tract</i> , soit chambre de chasse sous-aortique
MAP	<i>Mean arterial pressure</i> , soit pression artérielle moyenne
M-CARS	<i>Mayo CICU Admission Risk Score</i>
NIV	<i>Non invasive ventilation</i> , soit ventilation non invasive
NPV	<i>Negative predictive value</i> , soit valeur prédictive négative
NSTEMI	<i>Non ST-elevation myocardial infarction</i> , soit infarctus du myocarde sans sus-décalage du segment ST
NTproBNP	<i>N-terminal prohormone of brain natriuretic peptide</i> , soit prohormone N-terminale du peptide natriurétique cérébral
PAD	Pression artérielle diastolique
PaO ₂	<i>Arterial partial pressure of oxygen</i> , soit pression partielle artérielle en oxygène
PAPs	Pression artérielle pulmonaire systolique
PAS	Pression artérielle systolique
PESI	<i>Pulmonary Embolism Severity Index</i>
pOD	Pression atriale droite
PPC	Pression positive continue
PPV	<i>Positive predictive value</i> , soit valeur prédictive positive
qSOFA	<i>Quick Sequential Organe Failure Assessment</i> , soit version rapide du SOFA
RDW	<i>Red blood cell distribution width</i> , soit indice de distribution des globules rouges
ROC	<i>Receiver operating characteristic</i> , soit caractéristique de performance
SHAP	Valeur de Shapley

SIV	Septum interventriculaire
SOFA	<i>Sequential Organ Failure Assessment</i>
sPAP	<i>Systolic pulmonary artery pressure</i>
sPESI	<i>Simplified Pulmonary Embolism Severity Index</i> , soit version simplifiée du PESI
STEMI	<i>ST-elevation myocardial infarction</i> , soit infarctus du myocarde avec sus-décalage du segment ST
SBP	<i>Systolic blood pressure</i> , soit pression artérielle systolique
TAPSE	<i>Tricuspid annular plane systolic excursion</i> , soit excursion systolique du plan de l'anneau tricuspide
TIMI	<i>Thrombolysis in Myocardial Infarction</i>
TOPCAT	<i>Treatment of Preserved Cardiac Function Heart Failure with an Aldosterone Antagonist</i>
USIC	Unité de soins intensifs cardiologiques
VD	Ventricule droit
VG	Ventricule gauche
Vmax	Vitesse maximale
VNI	Ventilation non invasive
VTDVG	Volume télédiastolique du ventricule gauche
VTI	<i>Velocity time integral</i>
VTSVG	Volume télésystolique du ventricule gauche
XG-Boost	<i>Extreme gradient boosting</i> , soit méthode d'amplification de gradient

Abstract (324 mots, 2151 caractères)

Introduction. Stratifier le risque d'évènement péjoratif majeur chez les patients admis en USIC reste difficile. Des scores pronostiques existent mais ceux-ci ont été développés pour des populations d'unités de soins intensifs non cardiologiques ou pour des pathologies cardiaques spécifiques. Par ailleurs ils n'intègrent que peu ou pas les paramètres échocardiographiques. Un modèle pronostique (ML XG-Boost) réalisable dès l'admission a été créé à l'aide de paramètres échocardiographiques seuls et construit via Machine Learning. Ce modèle a montré de bonnes performances dans la cohorte multicentrique française de dérivation initiale (ADDICT-ICCU). L'objectif de ce travail était d'en réaliser sa validation externe.

Matériel et méthodes. Il s'agit d'une étude prospective observationnelle monocentrique menée à l'USIC du CHU de Lille. Les patients adultes admis à l'USIC étaient incluables et l'ETT était réalisée à l'admission. Le critère de jugement principal était la survenue d'un évènement clinique intra-hospitalier majeur parmi le décès toute cause, l'arrêt cardiaque ressuscité ou la survenue d'un choc cardiogénique dans les 28 jours suivant l'admission.

Résultats. Dans notre cohorte de 200 patients inclus à l'USIC du CHU de Lille entre novembre 2022 et avril 2023 inclus, 135 (67,5%) étaient des hommes, l'âge moyen était de $62,6 \pm 16,6$ ans, 122 patients étaient hypertendus (61,3%), 60 (30,2%) étaient diabétiques. La FEVG moyenne était à $49,1 \pm 16,2$ %, le rapport E/E' moyen à $11,1 \pm 4,74$. L'ITV sous-aortique était en moyenne à $14,7 \pm 4,72$ cm, le TAPSE à $19 \pm 5,01$ mm, la PAPs était en moyenne à $41,7 \pm 15,6$ mmHg. 28 patients (14%) ont présenté un évènement cardiovasculaire grave à 28 jours. On note 10 décès (5%), 18 chocs cardiogéniques (9%) et 9 arrêts cardiaques récupérés

(4,5%). Dans cette population, les performances du modèle ML XG-Boost restaient supérieures aux scores préexistants (AUROC 0,74, $p < 0,001$).

Conclusion. Les performances du modèle ML XG-Boost dans une cohorte externe lilloise sont satisfaisantes et permettent sa validation externe. De plus amples études doivent permettre de corroborer ces résultats.

Abstract (298 words, 1892 characters)

Introduction. Nowadays, it remains difficult to stratify the risk of major adverse events in ICCU patients. Several prognostic scores exist, but they only apply to specific cardiac conditions, or to medical intensive care unit patients, and not specifically to ICCU patients. Furthermore, they barely use echocardiography data, if any. A new prognostic model (ML XG-Boost) has been built thanks to Machine Learning, using TTE data only. This model showed good performances in the French multicentric initial cohort (ADDICT-ICCU). The aim of this work was to confirm its good performances in an external validation cohort.

Materials & methods. This study is a monocentric prospective and observational cohort conducted in Lille University Center ICCU. Adult patients admitted to this Unit were eligible. The primary composite endpoint was the occurrence of an in-hospital major adverse event, including death of any cause, resuscitated cardiac arrest or cardiogenic shock, in the first 28 days following admission.

Results. 200 patients were included between November 2022 and April 2023. There were 135 male (67.5%), mean age was 62.6 ± 16.6 . Hypertension was found in 122 patients (61.3%), and diabetes in 60 cases (30.2%). Mean LVEF was $49.1 \pm 16.2\%$, mean E/E' was 11.1 ± 4.74 . Mean LVOT TVI was 14.7 ± 4.72 cm, mean TAPSE was 19 ± 5.01 mm, and mean sPAP was 41.7 ± 15.6 mmHg. Major adverse events at 28 days occurred in 28 patients (14%), including 10 deaths of any cause (5%), 18 cardiogenic shocks (9%) and 9 resuscitated cardiac arrests (4.5%). In this cohort, ML XG-Boost showed better performances compared to existing scores (AUROC 0.74, $p < 0.001$).

Discussion. In this independent cohort based in Lille University Center, ML XG-Boost continued to demonstrate good prognostic values, which allows its external validation. Further studies are still required to corroborate these results.

I. Introduction

A. Les Soins Intensifs Cardiologiques

1. Bref historique

En 1947, *Beck et al* décrivent un premier succès de réanimation d'un patient en arrêt cardiaque sur fibrillation ventriculaire par la technique de défibrillation (1). Ils souhaitent montrer, par ce *case report*, qu'il est possible de ramener à la vie des patients en état de mort apparente. La technique de réanimation cardio-pulmonaire a par la suite rapidement progressé, tout comme la méthode de défibrillation, passant d'une défibrillation interne exclusive (*open-chest defibrillation*) à une défibrillation externe (*closed-chest defibrillation*) (2).

Au milieu du XX^e siècle, la mortalité précoce des infarctus du myocarde est élevée et il est constaté qu'il existe un lien entre la mortalité et la survenue d'arythmies ventriculaires (fibrillation ventriculaire, tachycardie ventriculaire) (3–5). Dès lors, au vu des évolutions techniques, on imagine dans les années 1960 la création d'unités dont le but est la surveillance rapprochée des patients atteints d'infarctus du myocarde : le but est de diagnostiquer et surtout de traiter rapidement ces arythmies ventriculaires. Les premières unités de ce type, nommées « *Coronary Care Units* », ouvrent en 1962 à Kansas City, Sydney, Philadelphie et Toronto (6).

Les principes généraux de ces unités sont les suivants : une surveillance électrocardiographique continue et reliée à des alarmes, une rapidité de mise en œuvre des mesures de réanimation cardio-pulmonaire, du personnel entraîné et formé à l'utilisation de

matériels spécifiques dans cette unité et la présence d'une équipe paramédicale pouvant initier la réanimation cardio-pulmonaire avant même l'arrivée du personnel médical (7).

2. Épidémiologie

a. Épidémiologie initiale dans les *coronary care units* et son évolution

Comme leur nom l'indique, les *coronary care units* (CCU) ne s'intéressaient initialement qu'aux syndromes coronariens aigus et à leurs complications. Dans les années 1950, avant leur création, la mortalité intra-hospitalière en lien avec les infarctus du myocarde était aux alentours de 30 à 40%. Dans les années 1970, elle est évaluée à 15-20% (7). Selon Killip et Kimball, les CCU sont à l'origine d'une baisse de 20% de la mortalité dans les suites d'un infarctus (8).

Dans les années 1980, on remarque une nette évolution de l'épidémiologie des complications post-infarctus. La première cause de décès précoce en lien avec l'infarctus du myocarde n'est plus la survenue d'arythmie ventriculaire grave, mais l'insuffisance cardiaque et la survenue de choc cardiogénique (6,7).

La population a énormément évolué depuis les années 1960. Les soins prodigués en CCU sont devenus de plus en plus complexes, et les pathologies de plus en plus variées : on note une proportion croissante de patients porteurs de pathologies extracardiaques et de maladies chroniques. Parmi ces comorbidités, on retrouve par exemple le diabète, l'insuffisance rénale chronique, ou encore la broncho-pneumopathie chronique obstructive (7). Les patients hospitalisés dans ces unités ont en définitive plus de similitudes avec les patients des unités de soins intensifs médicales qu'avec les patients initiaux des CCU. Le terme de

CCU n'étant alors plus adapté, ces unités prennent progressivement le nom de *Cardiac Intensive Care Unit* (CICU), ou Unité de Soins Intensifs Cardiologiques (USIC) en français (9).

Les patients hospitalisés dans les USIC sont donc, pour partie, hospitalisés pour des pathologies aiguës extracardiaques, mais dont la condition cardiaque sous-jacente entraîne une particulière gravité. On observe d'ailleurs une aggravation globale de la sévérité des patients.

En 2012, Morrow *et al* étudient les évolutions concernant les diagnostics d'entrées au Duke University Coronary Care Unit entre 1989 et 2006 : on note une diminution de 7% des infarctus du myocarde avec sus-décalage du segment ST, une augmentation de 13% des infarctus du myocarde sans sus-décalage du segment ST et, concernant les diagnostics non cardiologiques, une augmentation de 26% des sepsis (7).

b. Épidémiologie actuelle des unités de soins intensifs cardiologiques

De nombreuses études se sont intéressées à l'épidémiologie des unités de soins intensifs cardiologiques. Il existe un certain degré de variation entre les pays, les régions et les types d'établissement. Les chiffres retrouvés ne sont donc jamais parfaitement généralisables. Sur la période 2013-2014, dans l'USIC de l'Université de Virginie, Holland et Moss trouvaient 25% d'hospitalisations pour syndrome coronarien aigu, 15% pour insuffisance cardiaque aiguë, 6% d'arythmies ventriculaires et 4% de troubles conductifs significatifs. On retrouvait également 26% d'admissions dans les suites de procédures interventionnelles planifiées. L'âge moyen des patients était de 65,9 ans, avec 62% d'hommes. L'indice de masse corporelle moyen était de 27,4 kg/m² (10).

Des données existent pour la France. Sur l'année 2014 on comptait plus de 277 000 admissions en USIC. L'âge moyen des patients hospitalisés était de 71 [59 ; 81] ans, avec 63% d'hommes. La première cause d'hospitalisation correspondait aux maladies coronaires (49%), puis venaient les arythmies (15,2%) et l'insuffisance cardiaque (10%) (11).

B. Évaluer la gravité des patients

Au vu de la gravité des patients et du nombre important d'admissions en USIC, l'un des enjeux majeurs actuels est la stratification du risque des patients : trouver le patient dont le risque d'évènement grave est plus important que les autres.

Des scores ont été publiés par le passé (score de GRACE, score TIMI, score ACUTE-HF, score PESI...). Ceux-ci se concentrent habituellement sur des sous-groupes de patients de cardiologie ou d'USIC, voire sur des populations d'unités de soins intensifs médicales, et leur utilisation généralisée sur une population hétérogène est donc difficile.

1. Exemples de score

a. Score de GRACE

Le registre GRACE (*Global Registry of Acute Coronary Events*) correspond à l'une des plus importantes cohortes de patients hospitalisés pour syndrome coronarien aigu, avec ou sans sus-décalage du segment ST. De nombreux centres y ont participé, dans 14 pays. La population de l'étude comportait plus de 11 000 patients et on notait 4,6% de décès intra-hospitaliers. De nombreux paramètres étudiés étaient statistiquement associés à la mortalité

intra-hospitalière. Parmi eux, 8 paramètres sont les plus déterminants pour expliquer la mortalité intra-hospitalière. Ces paramètres proviennent d'un modèle de régression logistique multivariée et ont été intégrés dans le score de GRACE (figure 1).

Figure 1 : Score de GRACE

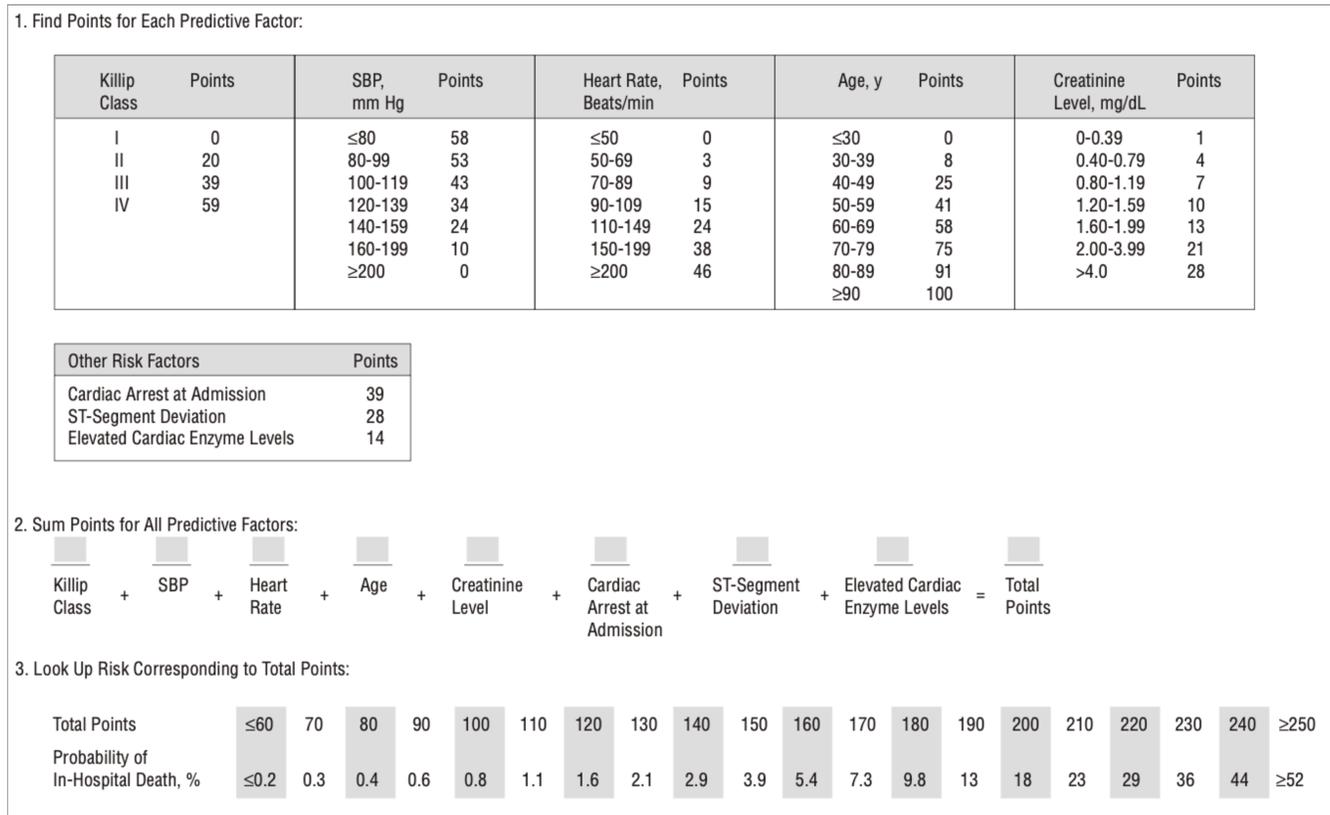


Illustration tirée de l'article de Granger et al (12). SBP = systolic blood pressure.

Celui-ci comporte l'âge, la classification Killip, la pression artérielle systolique, la présentation initiale comme arrêt cardiaque ressuscité, le niveau de créatininémie, la fréquence cardiaque, la déviation du segment ST, ainsi que la positivité des enzymes cardiaques à l'admission (12).

Au moment de l'établissement de ce score, il s'agissait alors du premier modèle s'intéressant aux syndromes coronariens aigus dans leur ensemble, et non seulement aux infarctus du myocarde avec sus-décalage du segment ST. Il est cependant assez peu utilisé

en pratique clinique : dans les recommandations de 2023 sur les syndromes coronariens aigus (SCA), l'*European Society of Cardiology* (ESC) mentionne la possibilité de calculer le score de GRACE afin de stratifier le risque des patients. En revanche, elle dégrade son niveau de recommandation, passant d'un grade I à un grade IIa (13).

b. Score TIMI

Le score TIMI (*Thrombolysis in Myocardial Infarction*) s'intéresse aux patients victimes de syndromes coronariens aigus sans sus-décalage du segment ST. Cette entité est vaste puisqu'elle regroupe deux diagnostics distincts : les infarctus du myocarde sans sus-décalage du segment ST et l'angor instable. C'est cette hétérogénéité qui provoque une différence de pronostic au sein de cette population. Le score TIMI a donc eu pour objectif de stratifier leur risque. Il a été développé en utilisant les patients de l'étude TIMI, dont l'objectif était de comparer des stratégies thérapeutiques antithrombotiques dans les syndromes coronariens aigus sans sus-décalage du segment ST.

Sur les 1957 patients inclus pour la réalisation du score, 16,7% ont présenté un évènement du critère de jugement principal composite (mortalité toute cause, infarctus du myocarde, revascularisation en urgence). 7 variables se sont avérées statistiquement associées à l'occurrence du critère de jugement principal dans l'analyse multivariée, avec des *odds ratio* globalement similaires. La présence de chacun de ces critères compte pour un point dans le calcul de ce score de risque (**figure 2**). Ces critères sont l'âge de 65 ans ou plus, la présence d'au moins trois facteurs de risque cardiovasculaire, la présence d'une sténose coronaire de plus de 50% déjà connue, la déviation du segment ST à l'arrivée, la survenue d'au moins 2 épisodes angineux dans les 24 heures précédant l'admission, l'utilisation d'aspirine

dans les sept jours précédents, les biomarqueurs cardiaques élevés (troponinémie ou dosage de la créatine phosphokinase [CPK]). A noter que les auteurs mettent en avant le fait que le score se calcule très facilement et ne nécessite pas d'aide informatique, ce qui est un avantage par rapport au score de GRACE (14).

Figure 2 : Score TIMI

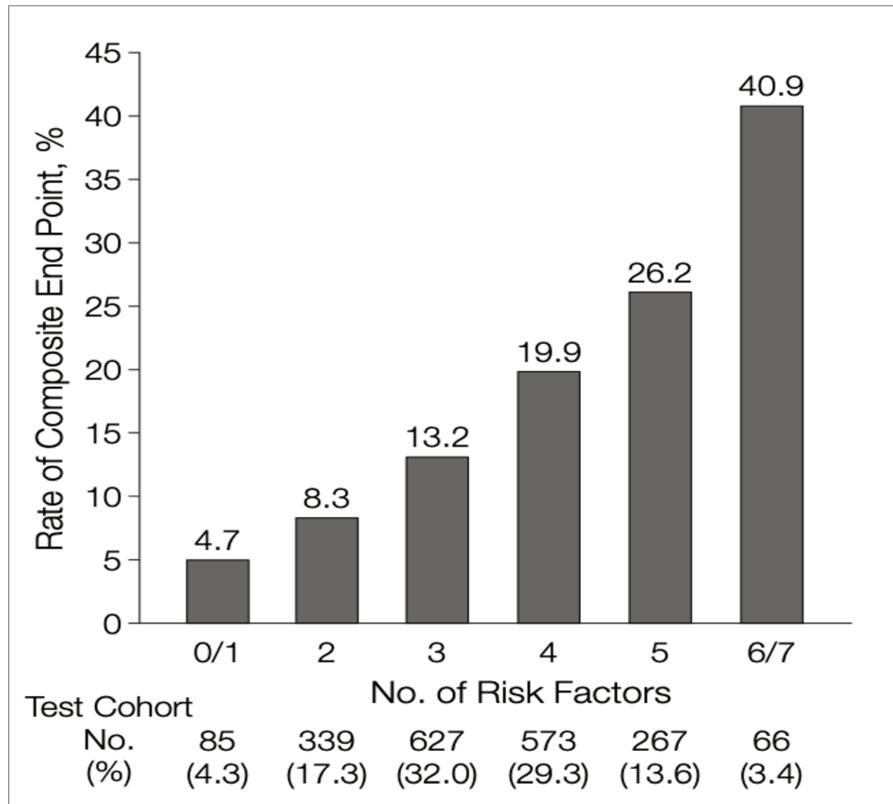


Illustration adaptée de l'article de Antman et al (14).

c. Score ACUTE-HF

ACUTE-HF est un modèle plus récent (**tableau 1**), s'intéressant spécifiquement aux patients présentant une décompensation cardiaque. L'insuffisance cardiaque est la première cause d'hospitalisation chez les patients de plus de 65 ans. Elle est très clairement associée à la mortalité et son taux important de ré-hospitalisation explique la morbidité qui lui est associée.

Les paramètres associés à la mortalité (à 30 jours, 6 mois et 5 ans) et à la durée d'hospitalisation ont été étudiés sur 771 patients. Les variables les mieux corrélées à la mortalité à 6 mois ont été intégrées dans le score ACUTE-HF. Il s'agit de la créatininémie supérieure à 20 mg/l, la fraction d'éjection ventriculaire gauche (FEVG) inférieure à 30%, l'âge de plus de 76 ans, l'antécédent d'hospitalisation pour insuffisance cardiaque, pour accident vasculaire cérébral ou accident ischémique transitoire, la présence d'une insuffisance mitrale au moins modérée et la nécessité de recourir à la ventilation non invasive. Plus le score est élevé, plus le risque de décès à 6 mois est important : pour un score supérieur à 3, le risque est estimé à 62,9% (15).

Ce score a montré de bonnes valeurs prédictives concernant la mortalité dans des cohortes de validation externe (16).

Tableau 1 : Score ACUTE-HF

	Points
Age > 76 ans	0,7
Créatinine > 2 mg/dL	1,4
Utilisation de VNI	0,8
AIT ou AVC	0,9
FEVG < 30%	0,8
Antécédent d'hospitalisation pour décompensation cardiaque	0,7
Insuffisance mitrale au moins modérée	0,5
Total	5,8

*Illustration adaptée et traduite de l'article de Cameli et al (15).
VNI = ventilation non invasive. AIT = accident ischémique cérébral.
FEVG = fraction d'éjection ventriculaire gauche.*

d. Score PESI

L'embolie pulmonaire est une pathologie fréquente et particulièrement variée quant à sa présentation clinique. Elle s'étend de cas pauci voire asymptomatiques, à des états de choc voire des arrêts cardiaques d'emblée. De fait, la mortalité est très variable, passant de moins de 2% dans les cas les moins graves, à 95% en cas d'embolie pulmonaire compliquée d'arrêt cardiaque (17). Le score PESI (*Pulmonary Embolism Severity Index*), publié en 2005, est un outil de stratification du risque chez les patients présentant une embolie pulmonaire (**figure 3**). Ce score participe à la stratégie thérapeutique, détermine la nécessité d'une hospitalisation, et contribue à définir la durée d'hospitalisation. Onze paramètres, cliniques, biologiques, et d'imagerie cardiaque, sont pris en compte et intégrés dans ce score (17).

Figure 3 : Scores PESI et sPESI

Variable	Score	
	Original PESI ^a	Simplified PESI ^b
Age >80 y	Age in years	1
Male sex	+10	
History of cancer	+30	1
History of heart failure	+10	1 ^c
History of chronic lung disease	+10	
Pulse ≥110 beats/min	+20	1
Systolic blood pressure <100 mm Hg	+30	1
Respiratory rate ≥30 breaths/min	+20	
Temperature <36°C	+20	
Altered mental status	+60	
Arterial oxyhemoglobin saturation level <90%	+20	1

Illustration adaptée de l'article de Jiménez et al (18).

Une version simplifiée (nommée sPESI) a été publiée en 2010 (**figure 3**), n'incluant que 5 paramètres : l'âge de plus de 80 ans, la pression artérielle systolique de moins de 100 mmHg, la fréquence cardiaque de plus de 110 bpm, la saturation en oxygène de moins de 90% en air ambiant, les antécédents de cancer, d'insuffisance cardiaque ou de maladie pulmonaire chronique. Ses performances ont été évaluées comme similaires au score PESI classique.

La mortalité des patients classés à faible risque par le score sPESI est d'environ 1% alors qu'elle est de 10,9% chez ceux classés à haut risque (18). Du fait de sa facilité d'utilisation, le score sPESI a été rapidement adopté et il est utilisé en pratique quotidienne. L'ESC recommande par ailleurs son utilisation avec un grade IIa et un niveau de preuve B (19).

En plus du PESI simplifié, la stratification du risque dans l'embolie pulmonaire repose également sur des critères d'imagerie et de biologie, en particulier le retentissement sur les cavités cardiaques droites (dilatation ventriculaire droite à l'angioscanner thoracique ou en échocardiographie transthoracique [ETT]) et le dosage des enzymes cardiaques (BNP ou NTproBNP, et troponinémie) (19).

e. Scores SOFA et APACHE IV

Les scores SOFA (*Sequential Organ Failure Assessment*) et APACHE IV (*Acute Physiology and Chronic Health Evaluation*) ont été développés pour estimer la gravité des patients admis dans les unités de soins intensifs médicales. Ils ont été testés par la suite dans des populations d'USIC et ont montré de manière générale de bonnes performances, avec par exemple une aire sous la courbe (AUC) pour le score APACHE IV à 0,82 dans une étude de Bennett *et al* en 2019 (20) et une AUC de 0,83 pour le score SOFA dans une étude de 2018 (21).

L'avantage principal du SOFA (**figure 4**) et de sa version rapide (le quick SOFA [qSOFA]) par rapport à l'APACHE IV est sa simplicité d'utilisation, avec nettement moins de variables nécessaires à son calcul.

Figure 4 : score SOFA

SOFA score	1	2	3	4
<i>Respiration</i>				
PaO ₂ /FiO ₂ , mmHg	< 400	< 300	< 200 —— with respiratory support ——	< 100
<i>Coagulation</i>				
Platelets × 10 ³ /mm ³	< 150	< 100	< 50	< 20
<i>Liver</i>				
Bilirubin, mg/dL (μmol/L)	1.2 – 1.9 (20 – 32)	2.0 – 5.9 (33 – 101)	6.0 – 11.9 (102 – 204)	> 12.0 (< 204)
<i>Cardiovascular</i>				
Hypotension	MAP < 70 mmHg	Dopamine ≤ 5 or dobutamine (any dose) ^a	Dopamine > 5 or epinephrine ≤ 0.1 or norepinephrine ≤ 0.1	Dopamine > 15 or epinephrine > 0.1 or norepinephrine > 0.1
<i>Central nervous system</i>				
Glasgow Coma Score	13 – 14	10 – 12	6 – 9	< 6
<i>Renal</i>				
Creatinine, mg/dL (μmol/L) or urine output	1.2 – 1.9 (110 – 170)	2.0 – 3.4 (171 – 299)	3.5 – 4.9 (300 – 440) or < 500 ml/day	> 5.0 (> 440) or < 200 mL/day

^a Adrenergic agents administered for at least 1 h (doses given are in μg/kg·min)

Illustration adaptée de l'article de Vincent et al (22). PaO₂ = arterial partial pressure of oxygen. FiO₂ = fraction of inspired oxygen. MAP = mean artérielle pressure.

Ces scores présentent toutefois un inconvénient commun : leur calcul est basé sur les données cumulées des 24 premières heures d'hospitalisation. Ils ne sont donc pas réalisables d'emblée à l'admission, ce qui limite leur utilisation à des fins d'orientation des patients (21).

f. Score M-CARS

En 2019, *Jentzer et al* ont pour but de créer un score pronostique de mortalité s'intéressant spécifiquement à la population des USIC : le score M-CARS (*Mayo CICU Admission Risk Score*) (23). L'intérêt de ce score (**figure 5**) réside dans le fait qu'il ne se limite

pas à une pathologie précise comme les scores sus-cités mais qu'il a l'ambition de donner le pronostic d'un patient admis en USIC quel que soit le diagnostic final.

Pour ce faire, ont été inclus de manière rétrospective et monocentrique plus de 10 000 patients hospitalisés en USIC. Les facteurs les plus fortement associés à la mortalité toute cause intra-hospitalière ont été recherchés par un modèle statistique de régression logistique. Afin de s'affranchir des difficultés de calcul du SOFA et de l'APACHE IV, ils n'ont retenu que les sept facteurs les plus significatifs. Ces facteurs étaient l'échelle de Braden, l'indice de distribution des globules rouges, le taux d'azote uréique sanguin, le trou anionique sanguin, l'arrêt cardiaque à l'admission, l'état de choc (quel qu'il soit) à l'admission et la défaillance respiratoire à l'admission. Leur modèle final montrait une bonne AUC à 0,87 dans leur cohorte initiale et 0,86 dans la cohorte de validation interne (23).

Figure 5 : score M-CARS

Variable	Value	Points Assigned
Admission value of BUN	>23 mg/dL	1
	≤23 mg/dL	0
Admission value of anion gap	>14	1
	≤14	0
Admission Braden skin score	≤12	2
	13–15	1
	>15	0
Admission value of RDW	>14.3	1
	≤14.3	0
Admission diagnosis of cardiac arrest	Yes	2
	No	0
Admission diagnosis of shock	Yes	2
	No	0
Admission diagnosis of respiratory failure	Yes	1
	No	0

Illustration adaptée de l'article de Jentzer et al (23). M-CARS = Mayo CICU Admission Risk Score. BUN = blood urea nitrogen. RDW = red blood cell distribution width.

Il est intéressant de noter que les paramètres échocardiographiques sont peu utilisés voire absents de l'ensemble des scores précédemment cités alors qu'ils sont un élément primordial de l'évaluation des patients à l'admission enUSIC.

2. Apport de l'échographie transthoracique

L'échocardiographie transthoracique fait partie intégrante du quotidien des cardiologues et cela est particulièrement vrai en unité de soins critiques.

Il semble exister un lien fort entre certains paramètres d'échocardiographie transthoracique et la mortalité intra-hospitalière. En 2022, *Jentzer et al* ont publié une étude visant à trouver les paramètres d'ETT les plus fortement associés à la mortalité intra-hospitalière dans une population de patients d'USIC (24). Ils ont, pour ce faire, inclus de manière rétrospective un peu moins de 7000 patients hospitalisés enUSIC et pour lesquels les variables d'ETT dans les 24 heures d'admission étaient disponibles. Sur cette population, 609 patients sont décédés. La population de survivants était très différente des patients décédés en intra-hospitalier, tant sur les paramètres d'ETT que sur les caractéristiques épidémiologiques habituelles ou sur les diagnostics. Après ajustement sur les marqueurs prédictifs de mortalité déjà connus, certains paramètres d'ETT restaient très fortement associés à la mortalité.

Les deux paramètres les plus fortement associés étaient l'intégrale temps-vitesse sous-aortique, en particulier lorsque sa valeur est inférieure ou égale à 16 cm, suivi du rapport E/E' médial à l'anneau mitral, surtout s'il est de 15 ou plus (**figure 6**) (24).

Figure 6 : Arbre de classification pronostique basé sur l'ETT

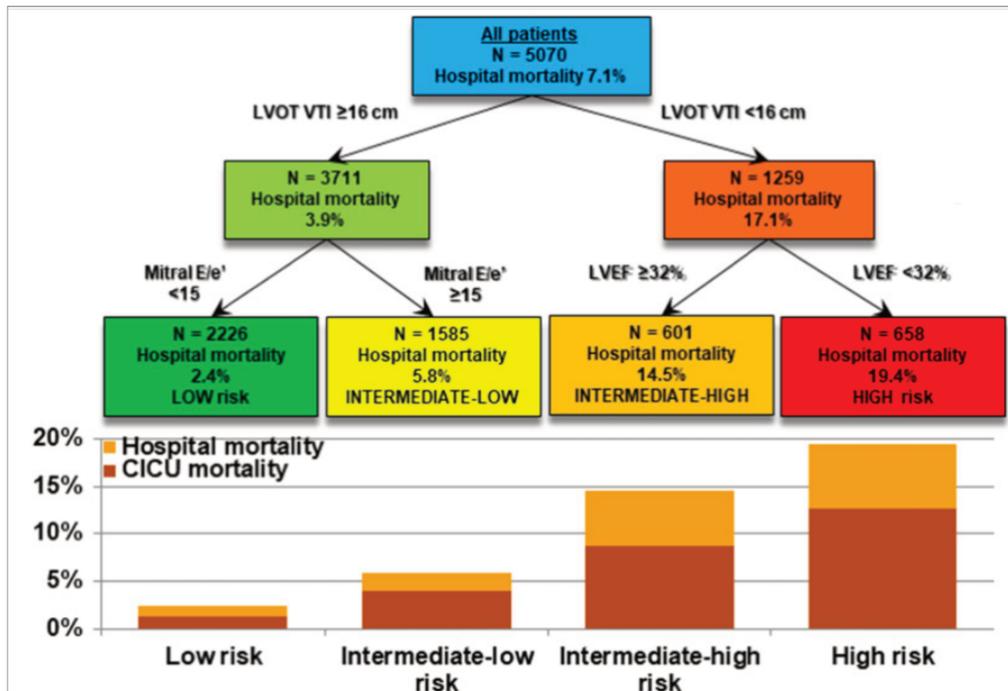


Illustration adaptée de l'article de Jentzer (24). LVOT = left ventricular outflow tract. VTI = velocity time integral. LVEF = left ventricular ejection fraction. CICU = Cardiac intensive care unit.

A l'heure actuelle, les scores pronostiques n'utilisent que peu les paramètres d'échographie cardiaque, en dehors de la FEVG. Des paramètres d'ETT ont été utilisés pour tenter de prédire le pronostic de patients admis enUSIC mais il y a eu peu de scores échographiques proposés (25). Un score utilisant les paramètres d'ETT recueillis à l'admission pourrait être plus à même de détecter les patients à risque d'évènement grave dans lesUSIC.

Par ailleurs, l'ensemble de ces scores a été obtenu en utilisant une méthode statistique : la régression logistique. Or, de nouvelles modalités de création de score sont apparues ces dernières années.

C. Intelligence artificielle et Machine Learning

La simple recherche « *Artificial Intelligence* » dans le moteur de recherche PubMed fait ressortir plus de 220 000 publications. Son association au mot « *Cardiology* » donne accès à plus de 4000 résultats. Sur les dernières années, son utilisation a été exponentielle. A titre d'illustration, la première version de ChatGPT®, créée par la société OpenAI, a été publiée le 30 novembre 2022. En avril 2023, on notait déjà plus de 120 millions de visites mensuelles sur le site.

1. Définition et principes

L'intelligence artificielle a été définie par Alan Turing comme une branche des sciences informatiques ayant pour objectif d'analyser des situations et de réaliser des tâches cognitives avec un degré proche des capacités de réflexion humaine (26,27). Le Machine Learning fait partie de l'intelligence artificielle et regroupe différentes méthodes permettant de tirer des informations depuis des bases de données par apprentissage automatique : il mélange informatique et statistiques. Arthur Samuel définit le Machine Learning en 1959 comme le fait de donner à un ordinateur la capacité d'apprendre sans qu'il ne soit explicitement programmé pour cela (28).

Dans un algorithme conventionnel, des instructions sont appliquées de manière stricte à des données en entrées, tandis que dans le cadre du Machine Learning les paramètres de l'algorithme vont s'affiner au fur et à mesure des entrées afin d'améliorer son fonctionnement. Cela implique un entraînement de l'algorithme sur une base de données qui

doit être représentée de manière compréhensible pour un ordinateur (matrice ou *dataset*) : chaque ligne y représente un objet (en l'occurrence un patient) et chaque colonne une caractéristique en lien avec cet objet (29).

Il existe deux principaux types de Machine Learning : l'apprentissage « supervisé » et l'apprentissage « non supervisé ». La création d'un modèle de Machine Learning supervisé comporte plusieurs étapes : une phase d'entraînement, une étape de validation, puis une phase de test. L'étape d'entraînement correspond à la création initiale du modèle, réalisée sur une partie du *dataset*, nommée « *training set* » ou jeu de données d'apprentissage, et qui représente habituellement 70-80% de la base de données. L'étape de validation est conduite sur une partie du *training set* nommée « *validation subset* », ou jeu de données de validation, qui correspond à une partie de la base de données initiale mais qui était inconnue du modèle. Cela permet d'en évaluer précocement les performances et d'en affiner certains paramètres (hyperparamètres). Enfin, les performances du modèle sont à nouveau évaluées sur le « *test set* », ou jeu de données de test, constituée des 20-30% restants de la base de données initiale. Il est impératif que les patients du « *test set* » n'aient pas été utilisés pour entraîner le modèle initial sous peine d'en surévaluer les performances. Ensuite la validation externe du modèle est proposée sur une autre cohorte (29,30).

Il existe de nombreuses méthodes d'apprentissage automatique, chacune ayant ses particularités et ses domaines d'application préférentiels. Les réseaux neuronaux convolutifs sont par exemple bien adaptés à l'analyse d'image : c'est la méthode de référence en ce qui concerne l'extraction de données en imagerie cardiovasculaire (29).

a. Apprentissage automatique supervisé

Les applications classiques des algorithmes de Machine Learning supervisé sont la réalisation de tâches de classification qui sont habituellement bien réalisées par les humains (29,30). Pour citer un exemple concret en médecine : fournir une banque d'images de radiographies pulmonaires classées par des humains en fonction de la présence ou non d'un nodule pulmonaire afin que l'algorithme puisse ensuite réaliser cette tâche en autonomie. Un autre exemple, centré cette fois sur la cardiologie, est l'interprétation automatique des électrocardiogrammes (ECG), où l'on apprend à un modèle à trouver des motifs électrocardiographiques reconnaissables (sus-décalage du segment ST, onde T négative, allongement de l'espace PR, élargissement des QRS) sur une base de données existante. Le but est que le modèle puisse ensuite reconnaître des motifs similaires et rendre, en autonomie, une interprétation correcte d'un ECG sur lequel il n'a pas été entraîné.

En résumé, les algorithmes d'apprentissage automatique supervisé apprennent sur la base d'exemples auxquels ils sont exposés.

b. Apprentissage automatique non supervisé

Dans le cadre du Machine Learning non supervisé, le but est de trouver des associations entre différentes variables, par exemple entre des supposés facteurs de risque et une pathologie précise (30). L'exemple type est la réalisation de tâches de *clustering*, c'est-à-dire la caractérisation de sous-groupes au sein d'une population hétérogène, en particulier pour des maladies complexes multifactorielles où les statistiques conventionnelles peuvent être

prises en défaut. Les clusters ne sont pas connus avant l'apprentissage, il s'agit un produit de l'analyse des données : le problème est donc dit « non supervisé ».

En effet, imaginons une population porteuse d'une pathologie complexe A, et une population saine appariée sur les facteurs de risque habituels de cette pathologie. Les modèles statistiques trouveront les paramètres significativement associés à la pathologie A. Toutefois, des sous-groupes de patients porteurs de cette pathologie présentent probablement des caractéristiques communes qu'il serait intéressant de relever. Dans le cas d'un modèle statistique conventionnel, ces caractéristiques seraient considérées comme non significativement associées à la pathologie. Par exemple, dans l'insuffisance cardiaque à FEVG préservée, les études d'intervention thérapeutique sont dans leur grande majorité négatives, notamment l'étude TOPCAT qui évaluait l'efficacité de la Spironolactone (**figure 7A**) (31). En 2020, un article est publié, basé sur les patients de l'étude TOPCAT, et dans lequel les auteurs identifient trois phénogroupes de patients au pronostic différent : le bénéfice de la Spironolactone est suggéré dans l'un de ces phénogroupes (**figure 7B**) (32).

Figure 7A : courbe de mortalité de l'étude TOPCAT

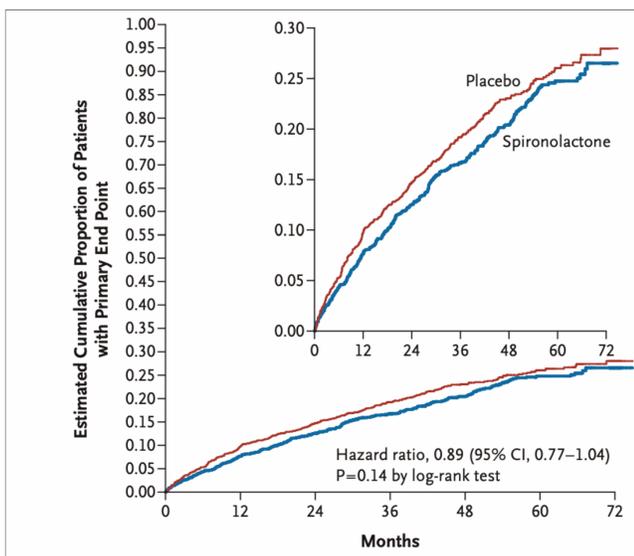


Illustration adaptée de l'article de Pitt et al (31).
CI = confidence interval.

Figure 7B : courbe de survie selon les phénogroupes d'insuffisance cardiaque

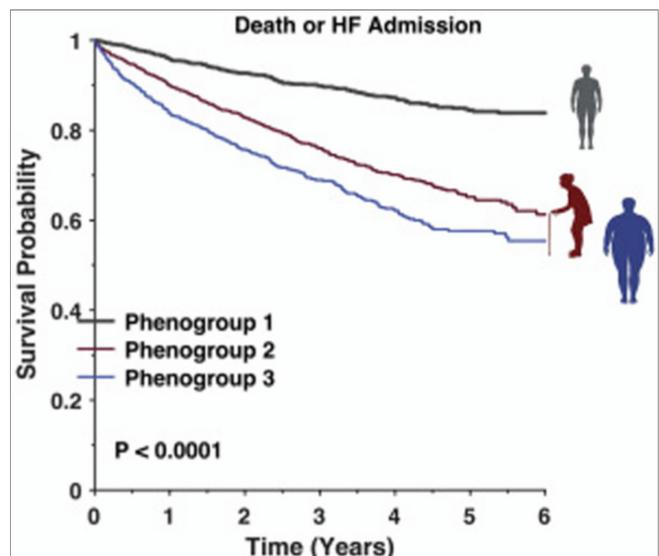


Illustration adaptée de l'article de Cohen et al (32).
HF = insuffisance cardiaque.

2. Apport du Machine Learning en cardiologie

Les scores pronostiques basés sur les statistiques conventionnelles n'utilisent qu'un nombre restreint de variables, habituellement celles suggérées par des modèles univariés et confirmées dans des modèles multivariés, mais cela ne permet pas toujours de capturer l'hétérogénéité des patients au sein d'une population d'intérêt. A l'aide de différentes méthodes de Machine Learning, il est plus aisé d'identifier les paramètres à intégrer dans un modèle prédictif et il est considéré qu'un des avantages majeurs du Machine Learning réside dans sa bonne performance en matière de choix des variables (*feature selection*). La méthode la plus utilisée à ce sujet est la méthode dite des forêts aléatoires en survie (30,33).

Plusieurs articles ont évalué l'utilité potentielle du Machine Learning en cardiologie. L'interprétation automatique des électrocardiogrammes, née dans les années 1960, est aujourd'hui quasiment universelle. Initialement son fonctionnement était assez rudimentaire, basé sur un système de règles assez simple. Plus récemment, un modèle de réseau neuronal convolutif développé par le département de sciences informatiques de l'Université de Stanford a montré ses bonnes performances diagnostiques par rapport à l'interprétation faite par des cardiologues (34).

Concernant l'échographie transthoracique, l'apport de l'intelligence artificielle est évident. Le caractère très chronophage de cet examen fait que l'on utilise régulièrement des estimations visuelles plutôt que des mesures précises (pour les limites anatomiques, les valvulopathies...). Il existe des logiciels pouvant réaliser des mesures automatiques, tels que EchoPAC® ou QLAB®, mais les résultats sont très dépendants de la manière dont les données sont collectées. En 2015, il a été montré que la FEVG et le *strain* longitudinal pouvaient être analysés de manière fiable et reproductible par un logiciel commercialement disponible

(AutoLV, TomTec Imaging Systems®). En comparaison avec l'estimation visuelle et la trace manuelle, le logiciel présentait une meilleure rapidité, sans perte de précision (35).

D. Vers un nouveau score

1. ADDICT-ICCU

a. Description de la cohorte et objectif principal

La cohorte ADDICT-ICCU a inclus de manière prospective des patients admis en Unité de Soins Intensifs Cardiologiques, du 7 au 22 avril 2021, dans 39 centres français, dont le Centre Hospitalier Universitaire (CHU) de Lille.

Tout patient de 18 ans ou plus, admis enUSIC, était incluable. Les patients exclus étaient les patients refusant de participer à l'étude ou ne pouvant consentir, les patients admis enUSIC dans les suites d'une procédure interventionnelle planifiée et les patients transférés d'un autre centre hospitalier dans lequel ils étaient restés 24 heures ou plus.

L'objectif principal de ce registre était d'étudier la prévalence des drogues à usage récréatif chez les patients admis enUSIC. De nombreux paramètres ont été recueillis, notamment des paramètres échocardiographiques (36–38). Du fait de son design, la population de cette cohorte était très peu sélectionnée, et par conséquent elle est un bon reflet de la population actuelle desUSIC françaises.

b. Caractéristiques de la population

Nous ne détaillons pas ici les caractéristiques de la population en lien avec l'objectif principal initial (prévalence des drogues à usage récréatif et leur association à la survenue d'évènements indésirables majeurs intra-hospitaliers), cela n'étant pas l'objet de ce travail.

En tout, 1499 patients ont été inclus dans cette cohorte, dont 69,6% d'hommes. L'âge moyen était de $63,3 \pm 14,9$ ans. 21,7% de ces patients étaient diabétiques, 53,0% hypertendus, 38,8% dyslipidémiques, 64,7% fumeurs (actifs ou sevrés). 15,6% de l'effectif avait déjà présenté un syndrome coronarien aigu par le passé.

51,5% des patients étaient admis pour infarctus du myocarde, avec ou sans sus-décalage du segment ST. L'insuffisance cardiaque concernait 14,1% de l'effectif. Les troubles conductifs représentaient 5,54% des patients, les embolies pulmonaires 3,07%, les arythmies supraventriculaires et ventriculaires respectivement 4,0% et 2,27%.

Des évènements indésirables intra-hospitaliers majeurs définis comme la mortalité toute cause, la survenue d'un choc cardiogénique, ou l'arrêt cardiaque ressuscité sont survenus dans 67 cas (4,47% de la population), dont 27 décès (1,81%) (39).

2. Élaboration du score

ADDICT-ICCU a servi de base de données afin d'évaluer la faisabilité et les paramètres d'un modèle prédictif dont l'objectif est de déterminer *a priori* la gravité et le devenir des patients admis en USIC (39). Les particularités de ce modèle résident dans le fait que la sélection des variables et sa construction ont été confiées à des algorithmes de Machine Learning, et dans le fait qu'il n'utilise que des paramètres d'échocardiographie transthoracique.

a. Sélection des variables

Seize paramètres échocardiographiques étaient disponibles sur les patients de ADDICT-ICCU : le volume télédiastolique du ventricule gauche (VTDVG) indexé à la surface corporelle, la fraction d'éjection ventriculaire gauche (FEVG), le volume atrial gauche indexé à la surface corporelle, l'épaisseur maximale du septum interventriculaire en télédiastole, les vitesses des ondes E et A sur le flux mitral, la vitesse de l'onde E' à l'anneau mitral en Doppler pulsé ainsi que le rapport E/E', l'intégrale temps-vitesse (ITV) sus-aortique, la dilatation du ventricule droit définie par un rapport VD/VG > 1 en coupe 4 cavités, l'excursion systolique du plan de l'anneau tricuspide (TAPSE), la vitesse maximale (Vmax) du flux d'insuffisance tricuspide, la pression artérielle pulmonaire systolique (PAPs), la présence d'une valvulopathie significative c'est-à-dire de grade supérieur ou égal à 2 sur 4 d'après les recommandations d'échocardiographie, la présence d'une endocardite confirmée par les critères de Duke modifiés, la présence d'un épanchement péricardique d'au moins 5mm en télédiastole, le diamètre maximal de la veine cave inférieure (39).

Parmi ces paramètres, 5 ont été sélectionnés par une méthode de Machine Learning dite des « forêts aléatoires en survie », ou *Random Survival Forest* en anglais. Ces paramètres sont les suivants : la FEVG, le rapport E/E', l'ITV sous-aortique, la PAPs et le TAPSE (39).

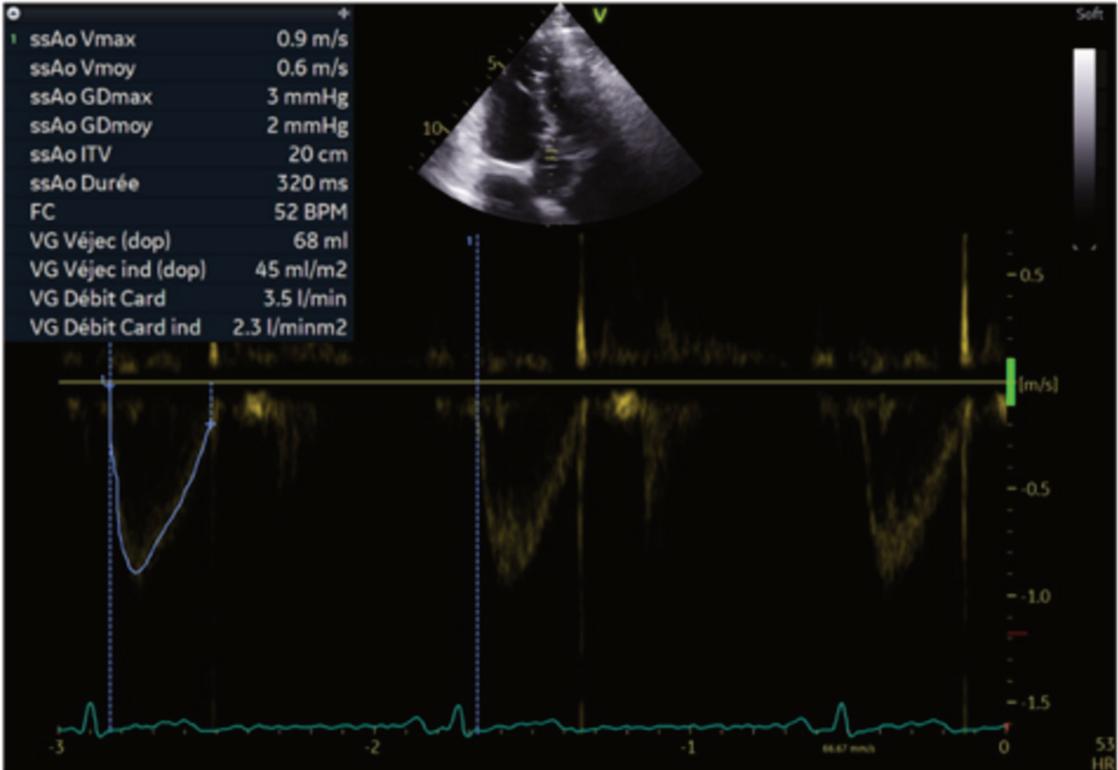
La FEVG (en pourcentage) se calcule de la manière suivante via la méthode de Simpson biplan :

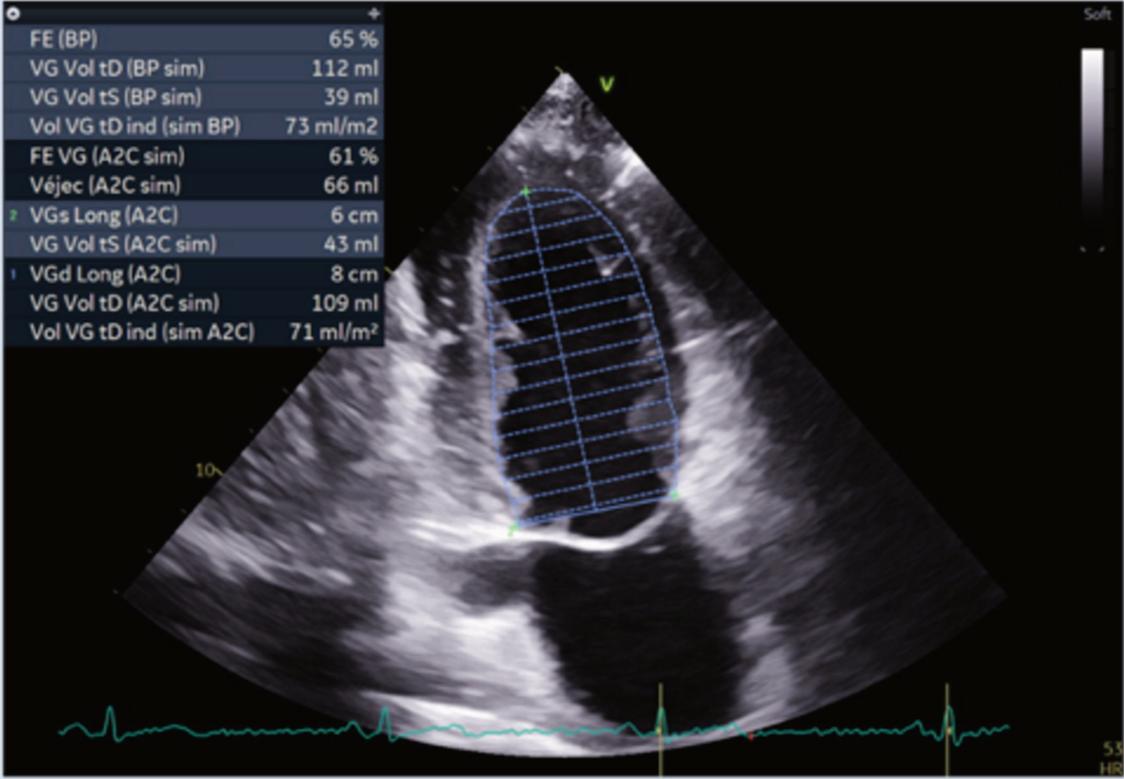
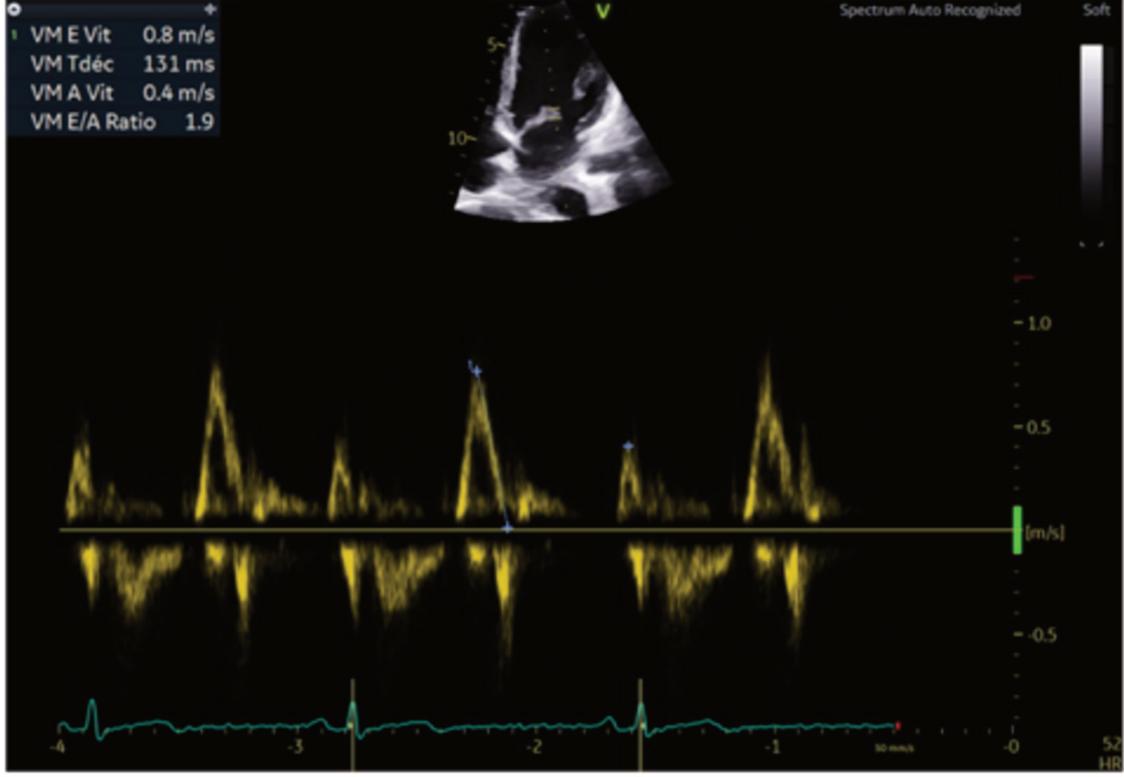
$$FEVG = \frac{(VTDVG - VTSVG)}{VTDVG}$$

où VTDVG et VTSVG sont les volumes télédiastolique et télésystolique du ventricule gauche (en mL). Le rapport E/E' se calcule à l'aide de plusieurs paramètres : la vitesse de l'onde E obtenue en Doppler pulsé sur le flux mitral (en m/s) et la moyenne des vitesses des ondes E' à l'anneau mitral septal et latéral en Doppler tissulaire (en cm/s). L'ITV sous-aortique se mesure par le contour du spectre Doppler pulsé en

systole au niveau de la chambre de chasse du ventricule gauche (en cm). La pression artérielle pulmonaire systolique (en mmHg) se calcule via l'équation de Bernoulli simplifiée : $PAPS = 4V^2 + pOD$, où V est la vitesse maximale de l'insuffisance tricuspide (en m/s), et pOD la pression atriale droite (en mmHg) estimée par le diamètre et le degré de collapsibilité de la veine cave inférieure. Le TAPSE se mesure en mode temps-mouvement par le déplacement de la partie latérale de l'anneau tricuspide en coupe apicale 4 cavités (en mm). Le **tableau 2** illustre chacun de ces paramètres à l'aide d'exemples tirés d'ETT.

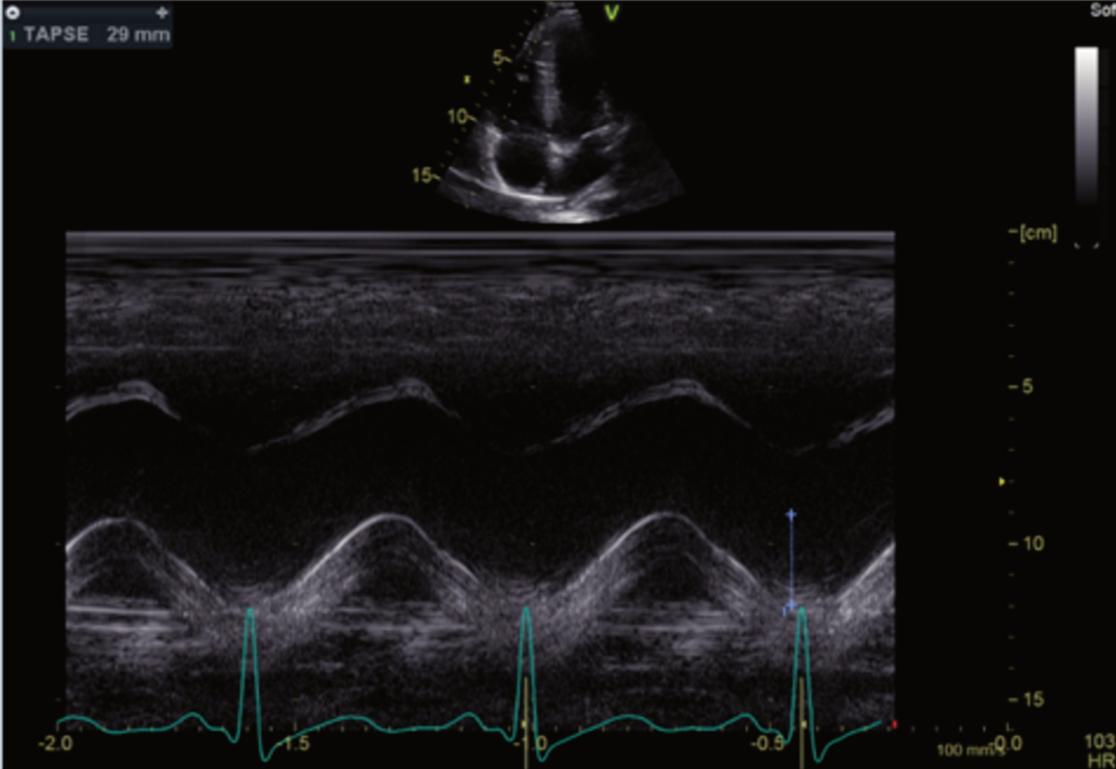
Tableau 2 : Illustrations des paramètres ETT sélectionnés par Machine Learning

Paramètre	Illustration
ITV sous aortique en Doppler pulsé	 <p>The illustration shows a Doppler ultrasound of the aortic valve. At the top, a color Doppler image shows regurgitant flow. Below it, a spectral Doppler waveform shows the regurgitant flow velocity over time. A list of parameters is displayed on the left side of the image:</p> <ul style="list-style-type: none"> ssAo Vmax: 0.9 m/s ssAo Vmoy: 0.6 m/s ssAo GDmax: 3 mmHg ssAo GDmoy: 2 mmHg ssAo ITV: 20 cm ssAo Durée: 320 ms FC: 52 BPM VG Véjec (dop): 68 ml VG Véjec ind (dop): 45 ml/m2 VG Débit Card: 3.5 l/min VG Débit Card ind: 2.3 l/minm2 <p>The spectral Doppler waveform shows a characteristic regurgitant flow pattern with a peak velocity of 0.9 m/s. The x-axis represents time in seconds, ranging from -3 to 0. The y-axis represents velocity in m/s, ranging from -1.5 to 0.5. The heart rate (HR) is 53 bpm.</p>

Paramètre	Illustration																						
<p>FEVG par méthode de Simpson biplan</p>	 <table border="1" data-bbox="341 220 722 567"> <tr><td>FE (BP)</td><td>65 %</td></tr> <tr><td>VG Vol tD (BP sim)</td><td>112 ml</td></tr> <tr><td>VG Vol tS (BP sim)</td><td>39 ml</td></tr> <tr><td>Vol VG tD ind (sim BP)</td><td>73 ml/m²</td></tr> <tr><td>FE VG (A2C sim)</td><td>61 %</td></tr> <tr><td>Véjec (A2C sim)</td><td>66 ml</td></tr> <tr><td>VGs Long (A2C)</td><td>6 cm</td></tr> <tr><td>VG Vol tS (A2C sim)</td><td>43 ml</td></tr> <tr><td>VGd Long (A2C)</td><td>8 cm</td></tr> <tr><td>VG Vol tD (A2C sim)</td><td>109 ml</td></tr> <tr><td>Vol VG tD ind (sim A2C)</td><td>71 ml/m²</td></tr> </table>	FE (BP)	65 %	VG Vol tD (BP sim)	112 ml	VG Vol tS (BP sim)	39 ml	Vol VG tD ind (sim BP)	73 ml/m ²	FE VG (A2C sim)	61 %	Véjec (A2C sim)	66 ml	VGs Long (A2C)	6 cm	VG Vol tS (A2C sim)	43 ml	VGd Long (A2C)	8 cm	VG Vol tD (A2C sim)	109 ml	Vol VG tD ind (sim A2C)	71 ml/m ²
FE (BP)	65 %																						
VG Vol tD (BP sim)	112 ml																						
VG Vol tS (BP sim)	39 ml																						
Vol VG tD ind (sim BP)	73 ml/m ²																						
FE VG (A2C sim)	61 %																						
Véjec (A2C sim)	66 ml																						
VGs Long (A2C)	6 cm																						
VG Vol tS (A2C sim)	43 ml																						
VGd Long (A2C)	8 cm																						
VG Vol tD (A2C sim)	109 ml																						
Vol VG tD ind (sim A2C)	71 ml/m ²																						
<p>Onde E via le flux mitral en Doppler pulsé</p>	 <table border="1" data-bbox="341 1039 560 1176"> <tr><td>VM E Vit</td><td>0.8 m/s</td></tr> <tr><td>VM Tdéc</td><td>131 ms</td></tr> <tr><td>VM A Vit</td><td>0.4 m/s</td></tr> <tr><td>VM E/A Ratio</td><td>1.9</td></tr> </table>	VM E Vit	0.8 m/s	VM Tdéc	131 ms	VM A Vit	0.4 m/s	VM E/A Ratio	1.9														
VM E Vit	0.8 m/s																						
VM Tdéc	131 ms																						
VM A Vit	0.4 m/s																						
VM E/A Ratio	1.9																						

Paramètre	Illustration
<p>Onde E' latérale via le Doppler tissulaire à l'anneau mitral</p>	
<p>Onde E' médiale via le Doppler tissulaire à l'anneau mitral</p>	

Paramètre	Illustration
<p>Vmax d'insuffisance tricuspide</p>	
<p>Veine cave inférieure</p>	

Paramètre	Illustration
TAPSE	

Images d'ETT réalisées dans le service Explorations Fonctionnelles Cardiovasculaires, CHU de Lille. Avec l'accord du Pr D. MONTAIGNE, chef de service. ITV = intégrale temps-vitesse. FEVG = fraction d'éjection ventriculaire gauche. Vmax = vitesse maximale. TAPSE = tricuspid annular plane systolic excursion.

b. Construction du modèle

Les 5 paramètres ont été intégrés dans quatre algorithmes différents entraînés sur 70% de la cohorte (*training set*) puis leurs performances ont été comparées sur les 30% restants (*test set*). Il existait un modèle de régression logistique, c'est-à-dire des statistiques traditionnelles, et trois modèles de Machine Learning : le modèle LASSO (méthode de contraction des coefficients), le modèle Random Forest (méthode basée sur des arbres décisionnels), et le modèle XG-Boost (méthode d'amplification de gradient, ou *extreme gradient boosting*).

Différents paramètres ont été utilisés pour évaluer les performances de chaque modèle, notamment l'aire sous la courbe ROC et l'aire sous la courbe de *precision-recall*. La courbe ROC évalue la sensibilité des scores, c'est-à-dire leur capacité à détecter correctement les patients qui vont présenter le critère de jugement principal (CJP), en fonction de leur propension à prédire à tort qu'un patient est à risque de présenter le CJP. Il s'agit de la métrique d'évaluation la plus utilisée et la plus versatile pour juger de la qualité d'un modèle de classification binaire. La courbe de *precision-recall* est une autre métrique, particulièrement utilisée dans le cadre des modèles de Machine Learning. La précision (*precision*), autrement appelée valeur prédictive positive, correspond au taux de patients présentant réellement le CJP parmi tous les patients détectés par le test. Le rappel (*recall*) est un synonyme de la sensibilité. En d'autres termes, la précision évalue la pertinence d'un modèle sur l'ensemble des prédictions positives alors que le rappel évalue sa pertinence sur l'ensemble des patients réellement positifs.

Figure 8 : Performances des scores évalués dans la cohorte de dérivation initiale

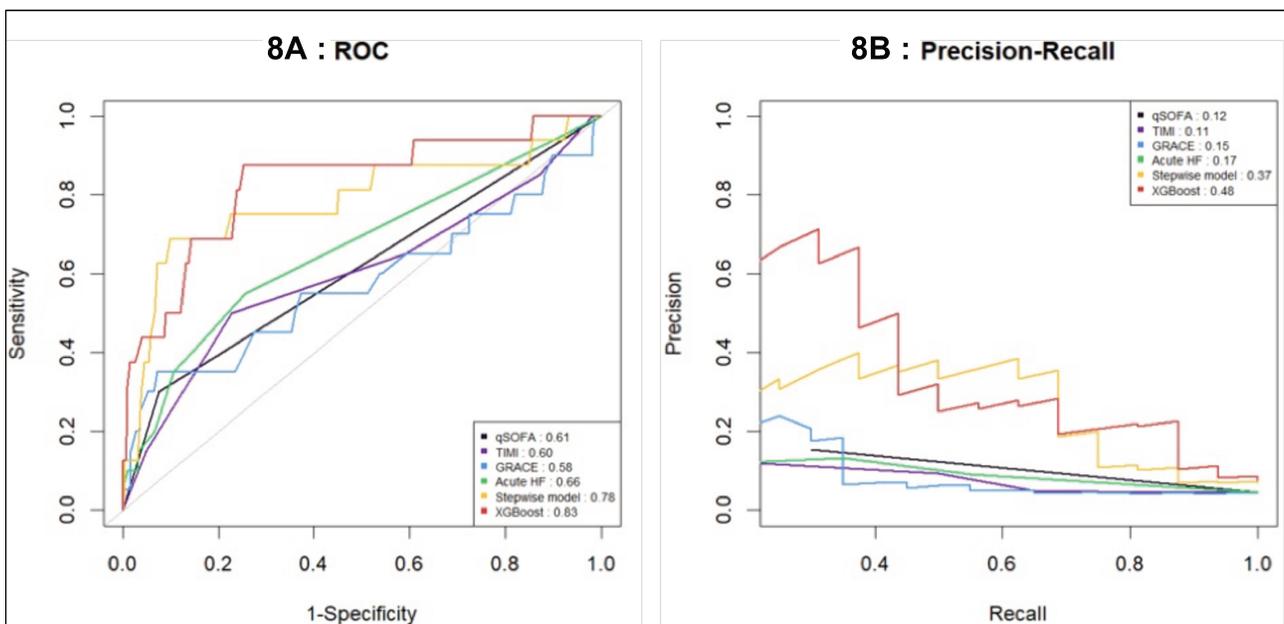


Illustration adaptée de l'article d'Aghezzi et al, under review (39).

En se basant sur ces différents paramètres, le modèle le plus performant était le modèle basé sur XG-Boost (nommé « ML XG-Boost »). Dans celui-ci, l'aire sous la courbe ROC était de 0,83 [0,75-0,86] (**figure 8A**) et l'aire sous la courbe de *precision-recall* était de 0,48 (**figure 8B**). Son score de Brier, représentant l'écart moyen entre les prédictions et les observations, était calculé à 0,83 (le plus bas parmi les quatre modèles testés). Les sensibilité et spécificité étaient respectivement de 0,38 et 0,98, et les valeurs prédictives positive et négative de 0,60 et 0,95 (39). Le **tableau supplémentaire 1** compare le ML XG-Boost aux autres scores testés sur différents paramètres statistiques.

XG-Boost (*extreme gradient boosting*) est une méthode populaire d'apprentissage automatique supervisé, utilisée pour répondre à des questions de classification ou de régression. Elle est dite « d'amplification de gradient », basée sur un ensemble de modèles d'arbre décisionnels. Lorsqu'un arbre est créé, il fournit des prédictions avec plus ou moins de décalage par rapport à la réalité observée. Ce décalage est nommé « résidu ». Le but de chaque arbre suivant est de corriger les prédictions incorrectes et de préserver les bonnes prédictions. Plutôt que de prédire de manière absolue le résultat, l'arbre va avoir pour but de prédire la différence entre chaque prédiction erronée et son observation. Ainsi, en additionnant le premier et le deuxième arbre, on affine la précision de la prédiction en diminuant le résidu : si le premier arbre était globalement performant, le second n'aura que peu d'impact dans la prédiction finale, et dans le cas contraire la prédiction sera améliorée. Par ailleurs, cette méthode permet de prévenir le surapprentissage (*overfitting*) via des techniques de régularisation (40).

c. Poids des variables dans la prédiction finale

Les modèles de Machine Learning apportent une couche de complexité par rapport aux statistiques conventionnelles et il est parfois difficile d'appréhender l'importance et la contribution de chaque variable dans les performances du modèle. C'est pourquoi les valeurs de Shapley ont été calculées. Plus la valeur est élevée, plus le paramètre présente de poids dans la prédiction finale (41). Dans le modèle ML XG-Boost, le paramètre contribuant le plus à la prédiction finale est en moyenne le rapport E/E' (figure 9).

Figure 9 : Valeurs de Shapley du modèle ML XG-Boost

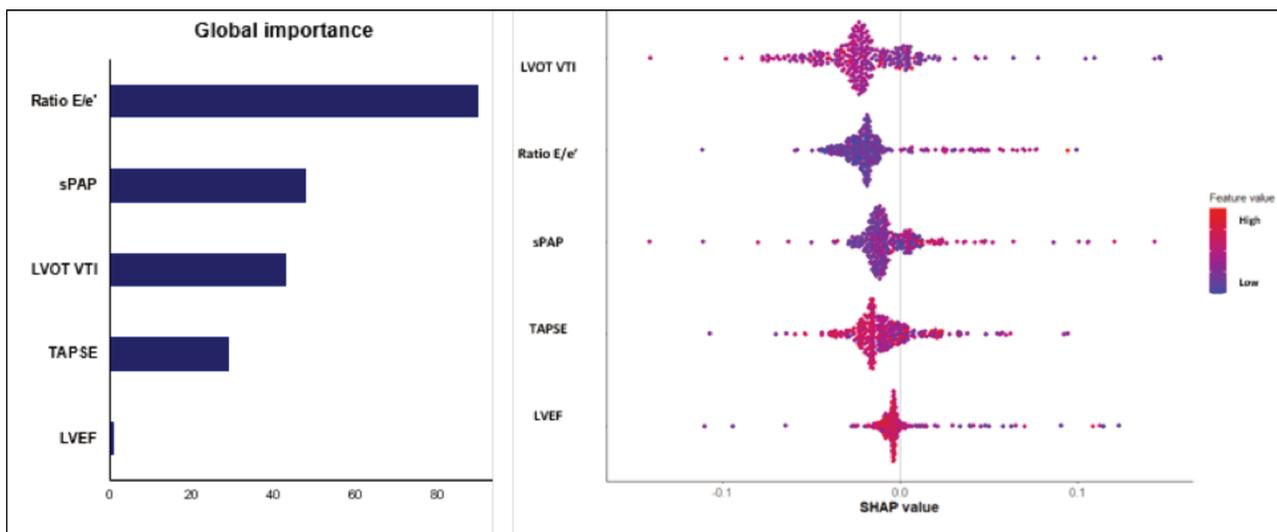
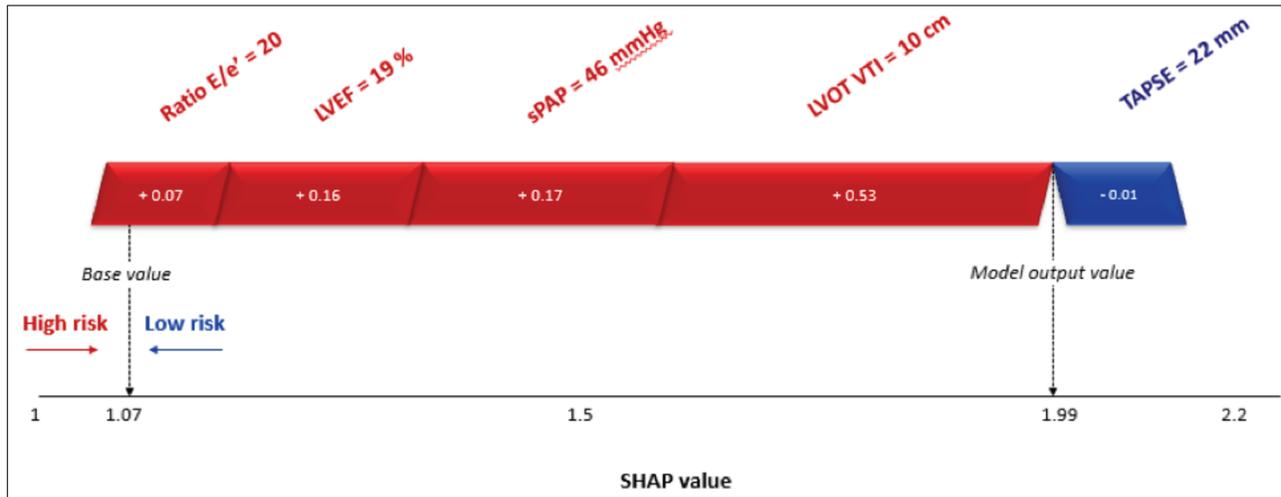


Illustration adaptée de l'article d'Aghezzaf et al, under review (39). sPAP = systolic pulmonary artery pressure. LVOT = left ventricular outflow tract. VTI = velocity time integral. TAPSE = tricuspid annular plane systolic excursion. LVEF = left ventricular ejection fraction.

Au-delà de l'importance globale des paramètres du modèle, il est également possible de représenter la contribution des variables pour chaque individu de la cohorte à l'aide de diagrammes dits *force plots*. La figure 10A illustre le *force plot* d'un patient ayant présenté le CJP, et la figure 10B correspond à un patient n'ayant pas présenté le CJP.

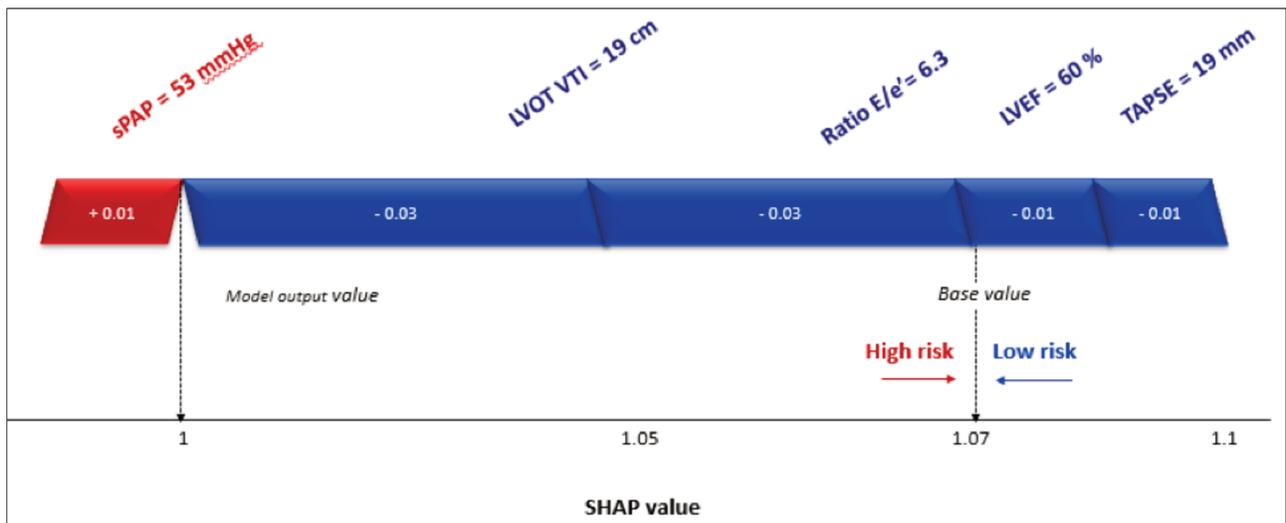
Concernant le patient de la **figure 10A**, ce sont l'ITV sous-aortique et la PAPs qui ont le plus contribué à dévier la prédiction finale à droite de la ligne de base (*base value*), dans la zone où le patient est considéré plus à risque de présenter le critère de jugement.

Figure 10A : Force plot des valeurs de Shapley d'un patient ayant présenté le CJP



Dans la situation du patient de la **figure 10B**, l'ITV sous-aortique et le rapport E/E' ont majoritairement contribué à dévier la prédiction finale à gauche de la ligne de base, dans la zone où l'on considère le patient moins à risque.

Figure 10B : Force plot des valeurs de Shapley d'un patient n'ayant pas présenté le CJP



Illustrations adaptées de l'article d'Aghezaf et al, under review (39). LVEF = left ventricular ejection fraction. sPAP = systolic pulmonary artery pressure. LVOT = left ventricular outflow tract. VTI = velocity time integral. TAPSE = tricuspid annular plane systolic excursion. SHAP = Shapley.

3. Objectif de travail

L'hypothèse de travail est que le modèle ML XG-Boost permet de prédire correctement le devenir des patients admis enUSIC.

L'objectif du travail est de réaliser la validation externe de ce modèle prédictif dans une cohorte lilloise de patients admis enUSIC. En effet, la création de tout score nécessite de vérifier son efficacité dans une cohorte de validation externe. Notre rôle a donc été de constituer une population de patients dont les caractéristiques devaient être proches de la population d'ADDICT-ICCU. Une fois cette cohorte constituée, les performances du modèle ML XG-Boost ont été évaluées et comparées aux scores préexistants.

II. Méthodes

A. La population de l'étude

La cohorte de validation externe a été réalisée dans le service d'Urgence et Soins Intensifs Cardiologiques, basé à l'Institut Cœur Poumon, CHU de Lille, France.

La période d'inclusion s'est étalée entre novembre 2022 et avril 2023. Les critères d'inclusion étaient similaires à ceux de la cohorte ADDICT-ICCU. Ont été exclus les patients dont l'échographie transthoracique réalisée dans les 24h de l'admission était incomplète, les patients ayant une durée de séjour de moins d'une journée, les patients mineurs et les patients victimes d'arrêt cardiaque extrahospitalier. Les patients admis au cours de périodes de garde

(soirs et week-ends) ont également été exclus. Une des particularités du service d'Urgence et de Soins Intensifs Cardiologiques du CHU de Lille est qu'il possède une aile dédiée à des consultations d'urgence cardiologiques. L'épidémiologie des patients de cette aile de consultation étant très différente et leur gravité étant tout autre, il a été considéré que ces patients n'étaient pas représentatifs du fonctionnement habituel d'un service d'USIC, et ces patients ont donc été exclus.

Les paramètres nécessaires au calcul des scores TIMI, GRACE, qSOFA et ACUTE-HF, ont également été collectés.

B. Collecte des données

Chaque patient hospitalisé dans le service bénéficiait systématiquement d'une échocardiographie transthoracique via l'un des trois échographes *General Electric*® de l'Unité. Les ETT étaient assurées par tous les médecins de l'Unité et étaient basées sur les dernières recommandations des sociétés savantes en matière d'imagerie cardiovasculaire (42,43). Les paramètres échocardiographiques d'intérêt ont été collectés avec l'aide d'une feuille imprimée dans l'Unité (**figure supplémentaire 2**) remplie par le médecin ayant réalisé l'ETT. Il s'agissait du volume télédiastolique du ventricule gauche, de la FEVG, de l'ITV sous-aortique, de la présence d'une valvulopathie supérieure ou égale à un grade 2, du rapport E/A, du rapport E/E' moyenné, du volume atrial gauche, de la présence d'une dilatation ventriculaire droite, du TAPSE et de la pression artérielle pulmonaire systolique.

Les données cliniques et paracliniques autres (diagnostic principal, pic de troponinémie, pression artérielle et saturation périphérique en oxygène, utilisation d'inotropes

ou de vasopresseurs...) étaient récupérées par consultation des dossiers médicaux au sein de l'Unité et avec l'aide de la lettre de sortie de l'Unité.

C. Critère de jugement principal

Le critère de jugement principal était la survenue d'un événement clinique intra-hospitalier majeur parmi le décès toute cause, l'arrêt cardiaque ressuscité, la survenue d'un choc cardiogénique nécessitant l'utilisation d'agents inotropes ou d'un système d'assistance circulatoire, dans les 28 jours suivant l'admission.

D. Statistiques

Les variables continues étaient exprimées en moyenne \pm déviation standard (DS) si leur distribution suivait une loi normale, et sinon en médiane et intervalle interquartile. Les variables catégorielles étaient exprimées en fréquence et pourcentage. La performance du modèle ML XG-Boost à prédire les événements cardiovasculaires majeurs à 28 jours a été testée et comparée à d'autres scores.

Les logiciels R version 4.0.2 (*R Foundation for Statistical Computing*) et Python version 3.5 ont été utilisés pour la réalisation des statistiques. Une valeur de $p \leq 0,05$ était considérée comme statistiquement significative.

III. Résultats

A. Description de la population

200 patients ont été inclus entre novembre 2022 et avril 2023. Les caractéristiques détaillées sont retranscrites dans le **tableau 3**. Parmi ces 200 patients, 135 étaient des hommes (67,5%), l'âge moyen était de $62,6 \pm 16,6$ ans.

Concernant les facteurs de risque cardiovasculaire, 122 patients étaient hypertendus (61,3%), 97 connus pour une dyslipidémie (48,7%), 27 présentaient une hérédité cardiovasculaire (13,6%), 60 étaient diabétiques (30,2%), 52 obèses (26%), 53 tabagiques actifs (26,6%).

Sur le plan des antécédents personnels cardiovasculaires, 60 étaient connus pour de l'insuffisance cardiaque (30%), 44 avaient fait un épisode de décompensation cardiaque par le passé (22%), 41 avaient déjà présenté un infarctus du myocarde (20,5%).

Concernant les données cliniques, les pressions artérielles systolique et diastolique moyennes étaient à $128 \pm 25,5$ mmHg et $74,7 \pm 16,3$ mmHg respectivement, la fréquence cardiaque moyenne à $87,7 \pm 27,3$ battements par minute, la saturation périphérique moyenne en oxygène à $97,4 \pm 2,56$ %.

Les caractéristiques biologiques moyennes étaient les suivantes : l'hémoglobine était à $13,2 \pm 2,22$ g/dL, la créatininémie à $111 \pm 93,5$ μ mol/L, le pic de troponinémie à 1152 ± 3756 ng/L, le NTproBNP à $5982 \pm 14\ 009$ pg/mL.

Des coronarographies ont été réalisées pour 136 patients soit 68% de notre effectif. Nous avons eu recours à une ventilation non invasive pour 17 patients (8,5%).

Le diagnostic principal le plus fréquent sur notre population correspond aux infarctus du myocarde, qu'ils soient avec ou sans sus-décalage du segment ST, respectivement pour 35 cas (17,8%) et 22 cas (11,2%). Par la suite viennent les décompensations cardiaques, avec 43 cas soit 21,8%. Les maladies valvulaires représentent 9,64% des cas (19 patients). Les arythmies ventriculaires concernent 6,09% des cas (12 patients), tout comme les arythmies supraventriculaires (12 cas également). Les troubles conductifs de haut degré étaient retrouvés dans 13 cas soit 6,60%. L'embolie pulmonaire représentait 4,57% des cas soit 9 patients. 6 patients ont été hospitalisés pour une myocardite et/ou une péricardite (3,05%). Enfin on note 10,7% de diagnostics autres que ceux sus-cités, et 2,54% de douleurs thoraciques inexpliquées.

Tableau 3 : Caractéristiques de la cohorte lilloise

	Total (n=200)	CJP - (n=172)	CJP + (n=28)	p-value
Données démographiques				
Hommes, n (%)	135 (67.5)	116 (67.4)	19 (67.9)	0.925
Âge, années	62.6 ±16.6	62.4 ±16.8	63.8 ±15.9	0.663
IMC, kg/m ²	27,7±6,44	27,9±6,66	26,5±4,78	0,178
Facteurs de risque cardiovasculaire, n (%)				
Hypertension artérielle	122 (61.3)	104 (60.8)	18 (64.3)	0.889
Dyslipidémie	97 (48.7)	87 (50.9)	10 (35.7)	0.199
Hérédité	27 (13.6)	22 (12.9)	5 (17.9)	0.551
Diabète	60 (30.2)	51 (29.8)	9 (32.1)	0.980
Obésité	52 (26.0)	47 (27.3)	5 (17.9)	0.408
Fumeur	53 (26.6)	42 (24.6)	11 (39.3)	0.161
Antécédents de maladie cardiovasculaire, n (%)				
Décompensation cardiaque	44 (22.0)	37 (21.5)	7 (25.0)	0.867
Insuffisance cardiaque	60 (30.0)	51 (29.7)	9 (32.1)	0.965
Infarctus du myocarde	41 (20.5)	36 (20.9)	5 (17.9)	0.904

	Total (n=200)	CJP - (n=172)	CJP + (n=28)	p-value
Données cliniques à l'admission				
PAS, mmHg	128 ±25.5	130 ±24.7	116 ±27.6	0.019
PAD, mmHg	74.7 ±16.3	75.8 ±15.9	67.9 ±17.4	0.034
Fréquence cardiaque, bpm	87.7 ±27.3	84.9 ±24.7	106 ±35.9	0.007
Saturation en oxygène, %	97.4 ±2.56	97.6 ±2.39	96.6 ±3.40	0.182
Données biologiques				
Hémoglobine, g/dL	13.2 ±2.22	13.3 ±2.23	12.5 ±2.08	0.007
Créatinine, µmol/L	111 ±93.5	109 ±96.6	125 ±71.6	0.035
Pic de troponinémie, UL	1152 ±3756	984 ±3060	2177 ±6576	0.011
NTproBNP, pg/mL	5982 ±14009	4986 ±12067	11994 ±21792	<0.001
Interventions et traitements reçus au cours de l'hospitalisation, n (%)				
Pas de coronarographie	136 (68.0)	116 (67.4)	20 (71.4)	0.841
VNI ou PPC	17 (8.50)	10 (5.81)	7 (25.0)	0.004
Diagnostic principal final, n (%)				0.005
Douleur thoracique inexplicée	5 (2.54)	5 (2.96)	0 (0.00)	
Décompensation cardiaque	43 (21.8)	36 (21.3)	7 (25.0)	
Myocardite ou péricardite	6 (3.05)	6 (3.55)	0 (0.00)	
Autres	21 (10.7)	20 (11.8)	1 (3.57)	
NSTEMI	22 (11.2)	22 (13.0)	0 (0.00)	
STEMI	35 (17.8)	28 (16.6)	7 (25.0)	
Arythmie supraventriculaire	12 (6.09)	10 (5.92)	2 (7.14)	
Embolie pulmonaire	9 (4.57)	9 (5.33)	0 (0.00)	
Arythmie ventriculaire	12 (6.09)	9 (5.33)	3 (10.7)	
Trouble de conduction	13 (6.60)	13 (7.69)	0 (0.00)	
Valvulopathie sévère	19 (9.64)	11 (6.51)	8 (28.6)	

Illustration adaptée de l'article d'Aghezaf et al, under review (39). CJP = critère de jugement principal. IMC = indice de masse corporelle. PAS = pression artérielle systolique. PAD = pression artérielle diastolique. VNI = ventilation non invasive. PPC = pression positive continue. NSTEMI = non ST-elevation myocardial infarction. STEMI = ST-elevation myocardial infarction.

B. Caractéristiques échocardiographiques de la population

Concernant les paramètres échocardiographiques, ils sont présentés en détail dans le **tableau 4**. La FEVG moyenne était à $49,1 \pm 16,2$ %, le volume télédiastolique moyen du ventricule gauche à $137 \pm 62,2$ mL, l'épaisseur du septum interventriculaire en parasternal grand axe à $9,77 \pm 3,88$ mm en moyenne, le volume moyen de l'atrium gauche en télésystole à $84 \pm 38,7$ mL, le rapport E/A à $1,2 \pm 0,6$ en moyenne, et le rapport E/E' moyen à $11,1 \pm 4,74$ en moyenne. On retrouvait une valvulopathie de grade supérieur ou égal à 2 sur 4 chez 38 patients soit 19,1% de la population. Le ventricule droit était dilaté dans 42 cas (21,2%). L'ITV sous-aortique était en moyenne à $14,7 \pm 4,72$ cm, le TAPSE à $19 \pm 5,01$ mm, la vitesse maximale de l'insuffisance tricuspide à $2,82 \pm 0,56$ m/s. La PAPs était en moyenne à $41,7 \pm 15,6$ mmHg.

Tableau 4 : Données échocardiographiques de la cohorte lilloise

	Total (n=200)	CJP - (n=172)	CJP + (n=28)	p-value
FEVG, %	49.1 ±16.2	51.3 ±14.1	35.6 ±21.0	0.001
VTDBG, mL	137 ±62.2	131 ±59.7	173 ±66.4	0.004
Épaisseur du SIV, mm	9.77 ±3.88	9.82 ±4.11	9.50 ±1.70	0.496
Volume atrial gauche, mL	84.0 ±38.7	82.5 ±39.3	93.4 ±33.6	0.128
Rapport E/A	1.20 ±0.60	1.18 ±0.58	1.36 ±0.74	0.337
Rapport E/e'	11.1 ±4.74	10.8 ±4.54	13.3 ±5.40	0.030
Valvulopathie ≥ 2, n (%)	38 (19.1)	28 (16.4)	10 (35.7)	0.031
Dilatation VD, n (%)	42 (21.2)	33 (19.4)	9 (32.1)	0.201
ITV sous-aortique, cm	14.7 ±4.72	15.2 ±4.40	11.6 ±5.51	0.003
TAPSE, mm	19.0 ±5.01	19.3 ±4.86	17.0 ±5.48	0.041
Vmax d'IT, m/s	2.82 ±0.56	2.79 ±0.55	3.02 ±0.61	0.142
PAPs, mmHg	41.7 ±15.6	41.0 ±16.0	46.3 ±12.3	0.113

Illustration adaptée de l'article d'Aghezzaf et al, under review (39). CJP = critère de jugement principal. FEVG = fraction d'éjection ventriculaire gauche. VTDBG = volume télédiastolique du ventricule gauche. SIV = septum interventriculaire. VD = ventricule droit. ITV = intégrale temps-vitesse. TAPSE = tricuspid annular plane systolic excursion. Vmax = vitesse maximale. IT = insuffisance tricuspide. PAPs = pression artérielle pulmonaire systolique.

C. Occurrence du critère de jugement principal

La **tableau 5** consigne le nombre d'évènements survenus dans notre cohorte. 28 patients de notre cohorte (soit 14%) ont présenté un évènement cardiovasculaire grave à 28 jours, repris dans le critère de jugement composite. O note 10 décès, soit 5% de l'effectif. Un choc cardiogénique est survenu dans 18 cas (9%). Enfin 9 personnes ont présenté un arrêt cardiaque récupéré, soit 4,5%.

Chez les patients ayant présenté un évènement du critère de jugement, la durée d'hospitalisation était significativement plus longue ($9,04 \pm 7,36$ jours, p 0,002).

Tableau 5 : Occurrence du critère de jugement principal

	Total (n=200)	CJP - (n=172)	CJP + (n=28)	p-value
Événements indésirables majeurs, n (%)	28 (14)	0 (0)	28 (100)	<0.001
Mort, n (%)	10 (5.00)	0 (0.00)	10 (35.7)	<0.001
Arrêt cardiaque récupéré, n (%)	9 (4.50)	0 (0.00)	9 (32.1)	<0.001
Choc cardiogénique, n (%)	18 (9.00)	0 (0.00)	18 (64.3)	<0.001
Durée d'hospitalisation, jours	4.80 (4.01)	4.10 (2.60)	9.04 (7.36)	0.002

Illustration adaptée de l'article d'Aghezzaf et al, under review (39). CJP = critère de jugement principal.

D. Modèle de Machine Learning

Le modèle ML XG-Boost, provenant de la cohorte ADDICT-ICCU, a été testé sur notre cohorte. Ce modèle présentait une aire sous la courbe ROC de 0,74, contre 0,84 dans la cohorte de dérivation.

Nous avons également calculé les aires sous la courbe ROC des autres scores existants : le TIMI était à 0,57, le score de GRACE à 0,59, l'ACUTE-HF à 0,69, et le qSOFA à 0,65. En comparaison avec ces scores (**figure 11**), le ML XG-Boost est systématiquement supérieur ($p < 0,001$). Cela est également vrai en *precision-recall*, avec une AUC à 0,44, contre 0,32 pour le score qSOFA, qui est le deuxième meilleur score testé sur cette métrique d'évaluation.

Figure 11 : Performances des scores évalués dans la cohorte de validation externe

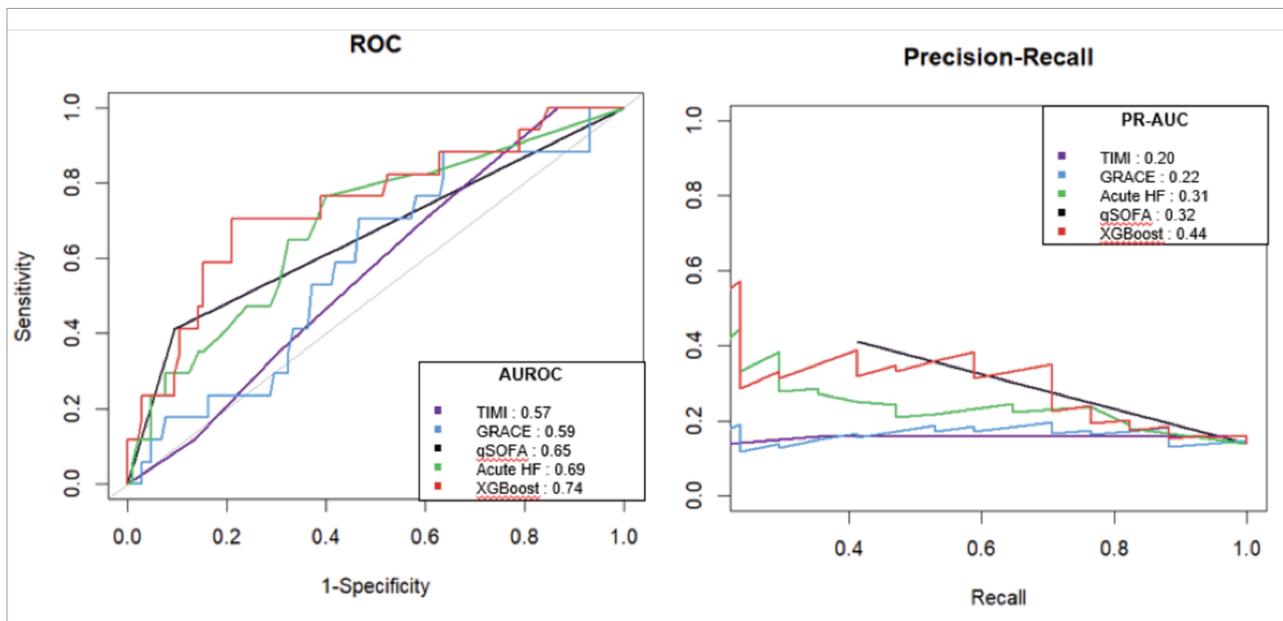


Illustration adaptée de l'article d'Aghezzaf et al, under review (39). PR-AUC = precision-recall area under curve. AUROC = area under ROC curve.

IV. Discussion

Notre travail a montré que chez 200 patients inclus à l'USIC du CHU de Lille sur une période de 6 mois, 28 (soit 14%) ont présenté un évènement cardiovasculaire grave et qu'un modèle prédictif obtenu par Machine Learning et n'utilisant que des paramètres d'échocardiographie permettait de prédire ces évènements. Cette étude a donc permis la validation externe de ce modèle.

Les données issues de notre cohorte permettent de confirmer les bonnes performances du modèle ML XG-Boost, avec des aires sous la courbe proches de celles de la cohorte de dérivation initiale. Cette cohorte de validation externe est une étape fondamentale dans l'établissement de ce score pronostique. ML XG-Boost discrimine mieux les patients qu'un score basé sur la régression logistique et fait également mieux que les scores existants auxquels il a été comparé. A notre connaissance, il s'agit du premier modèle pronostique de Machine Learning exclusivement construit à l'aide de données d'échocardiographie pour prédire le risque des patients hospitalisés en USIC. Stratifier le risque des patients admis en USIC est primordial, puisqu'il permet *in fine* de diminuer la mortalité notamment via une allocation personnalisée des ressources hospitalières.

Le modèle de *Random Survival Forest* a sélectionné 5 variables reflétant les fonctions systolique (FEVG, ITV sous-aortique, PAPs) et diastolique (PAPs, E/E') du ventricule gauche, ainsi que la fonction systolique du ventricule droit (TAPSE, PAPs) et les pressions pulmonaires (PAPs). Ces cinq paramètres sont intuitivement associés à un pronostic péjoratif, et le fait qu'ils aient été sélectionnés comme variables d'intérêt par le *Random Survival Forest* illustre les bonnes capacités du Machine Learning en matière de choix des variables.

Reconnaissons toutefois que pour un faible nombre de variables, l'utilisation d'une régression logistique était également possible.

Ce modèle prédictif présente deux originalités : la première est qu'il n'utilise aucune donnée démographique, clinique ou biologique, contrairement à la plupart des scores utilisés en routine, et la seconde est qu'il est applicable à tout patient admis enUSIC quel que soit le diagnostic final.

Concernant le premier point, Jentzer *et al* avaient déjà étudié les paramètres échographiques d'intérêt à l'admission enUSIC, et ils avaient conclu que l'ITV sous-aortique et le E/E' étaient les deux paramètres les plus fortement associés au devenir. Leur étude était de plus grande ampleur (n=6957), ils retrouvaient 8,8% de décès et n'utilisaient que la régression logistique. Il est intéressant de souligner que parmi nos 5 paramètres d'intérêt, 2 avaient été relevés par Jentzer. Notre modèle ne nécessite pas d'accumuler des données pendant plusieurs heures ou d'attendre des résultats de laboratoire, ce qui est un atout indéniable. Les paramètres utilisés sont des paramètres échographiques de routine dont l'obtention ne présente pas de difficulté technique particulière, ce qui rend cet outil utilisable y compris par de jeunes cardiologues en formation.

Concernant le second point, l'application du modèle ML XG-Boost dans une population de patients hospitalisés enUSIC, quel que soit le diagnostic final, est un avantage majeur par rapport aux autres scores existants. Avoir un outil applicable à tous les patients est très intéressant pour avoir une idée générale du niveau de risque du patient dès l'admission.

Concernant l'épidémiologie des patients admis enUSIC au CHU de Lille, on note que les diagnostics les plus fréquents correspondent aux SCA (STEMI et NSTEMI, 29% des admissions) suivis des décompensations cardiaques (21,8%) puis des valvulopathies sévères (presque 10%). Nous avons déjà précédemment expliqué qu'historiquement lesUSIC avaient été créés pour prendre en charge les patients victimes de SCA. La modification de

l'épidémiologie, liée en partie au vieillissement de la population et à l'amélioration du pronostic des SCA, entraîne une augmentation des admissions pour décompensation cardiaque et valvulopathie sévère. Ces dernières sont des pathologies pour lesquelles les données de l'ETT à l'admission sont importantes, ce qui explique peut-être l'efficacité de notre modèle échographique.

Il existe dans notre travail certaines limites qu'il convient de mentionner. L'un des biais potentiels de cette cohorte provient de son caractère monocentrique, celle-ci n'ayant inclus que des patients de l'USIC du CHU de Lille. Seuls les patients admis entre 8h30 et 18h ont été inclus. Nous ne pouvons exclure une épidémiologie différente chez les patients admis sur les périodes de week-ends ou de nuit. Par ailleurs, seuls 200 patients ont été inclus, avec un potentiel biais de sélection. Après consultation du journal des entrées de l'USIC du CHU de Lille sur la période d'inclusion (novembre 2022 à avril 2023 inclus) et vérification des critères d'inclusion et d'exclusion, 334 patients étaient incluables au total. Notre taux d'inclusion était donc de 59,9% (N = 200 /334). Les patients non inclus (N = 134) n'ont cependant pas fait l'objet d'erreurs systématiques et cela était parfaitement aléatoire. De plus, les caractéristiques de la population sont proches de celles retrouvées dans ADDICT-ICCU qui est, elle, multicentrique, et qui n'a pas été sujette à ce biais de sélection (recrutement de 1904 patients pour un nombre de théorique de 1755). Ce qui est également à relever, c'est que l'intensité des soins procurés en USIC en France est variable selon les centres. L'USIC du CHU de Lille présente des patients avec des tableaux en moyenne plus sévères que d'autres centres inclus dans la cohorte ADDICT-ICCU, ce qui aurait pu rendre les résultats de l'étude princeps difficilement extrapolables mais ce n'est pas le cas.

Notons que les patients sous assistance circulatoire n'ont pas été inclus dans la cohorte de dérivation ni dans la cohorte de validation. Ainsi le score n'est pas applicable dans cette population.

Enfin, les performances des scores comparés ont pu être sous-estimées : en effet, l'exclusion des patients présentant un arrêt cardiaque extrahospitalier, critère à prendre en compte dans le calcul du score de GRACE, a probablement participé à diminuer ses performances. De la même manière les scores ont pu être calculés en dehors de leur population d'intérêt initiale (calcul du score ACUTE-HF chez des patients admis pour syndrome coronarien aigu...). Cependant, cela illustre l'intérêt de ce nouveau modèle prédictif, applicable chez un grand nombre de patients.

V. Conclusion

Dans notre cohorte de patients de l'USIC du CHU de Lille, la performance du modèle ML XG-Boost est inférieure à celle de la publication princeps mais reste supérieure à celle des autres scores existants. Cela permet la validation externe de ML XG-Boost, modèle échographique seul, pour la prédiction des événements cardiovasculaires graves des patients d'USIC. Il a pour but de mieux stratifier le risque des patients, afin d'améliorer leur prise en charge. Ce modèle présente des avantages par rapport aux scores préexistants et est applicable à la quasi-totalité des patients admis en USIC. Sa méthode de création est basée sur le Machine Learning, branche en plein essor de l'intelligence artificielle permettant de répondre à des questions de régression ou de classification, et ce travail illustre son intérêt complémentaire par rapport aux statistiques conventionnelles. De plus amples travaux sont à réaliser pour corroborer nos résultats.

VI. Bibliographie

1. Beck CS. Ventricular fibrillation of long duration abolished by electric shock. JAMA. 13 déc 1947;135(15):985.
2. Kouwenhoven WB. Closed-chest cardiac massage. JAMA. 9 juill 1960;173(10):1064.
3. Killip T, Kimball JT. A survey of the coronary care unit: Concept and results. Progress in Cardiovascular Diseases. juill 1968;11(1):45-52.
4. Honey GE, Truelove SC. Prognostic factors in myocardial infarction. The Lancet. juin 1957;269(6980):1155-61.
5. Meltzer LE, Kitchell JB. The incidence of arrhythmias associated with acute myocardial infarction. Progress in Cardiovascular Diseases. juill 1966;9(1):50-63.
6. Julian DG. The history of coronary care units. Heart. 1 juin 1987;57(6):497-502.
7. Morrow DA, Fang JC, Fintel DJ, Granger CB, Katz JN, Kushner FG, et al. Evolution of Critical Care Cardiology: Transformation of the Cardiovascular Intensive Care Unit and the Emerging Need for New Medical Staffing and Training Models: A Scientific Statement From the American Heart Association. Circulation. 11 sept 2012;126(11):1408-28.
8. Killip T, Kimball JT. Treatment of myocardial infarction in a coronary care unit. The American Journal of Cardiology. oct 1967;20(4):457-64.
9. Katz JN, Minder M, Olenchock B, Price S, Goldfarb M, Washam JB, et al. The Genesis, Maturation, and Future of Critical Care Cardiology. Journal of the American College of Cardiology. juill 2016;68(1):67-79.
10. Holland EM, Moss TJ. Acute Noncardiovascular Illness in the Cardiac Intensive Care Unit. Journal of the American College of Cardiology. avr 2017;69(16):1999-2007.
11. Mercier G, Duflos C, Riondel A, Delmas C, Manzo-Silberman S, Leurent G, et al. Admissions to intensive cardiac care units in France in 2014: A cross-sectional, nationwide population-based study. Medicine. oct 2018;97(40):e12677.
12. Granger CB. Predictors of Hospital Mortality in the Global Registry of Acute Coronary Events. Arch Intern Med. 27 oct 2003;163(19):2345.
13. Byrne RA, Rossello X, Coughlan JJ, Barbato E, Berry C, Chieffo A, et al. 2023 ESC Guidelines for the management of acute coronary syndromes. European Heart Journal. 12 oct 2023;44(38):3720-826.

14. Antman EM, Cohen M, Bernink PJLM, 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*. 16 août 2000;284(7):835.
15. Cameli M, Pastore MC, De Carli G, Henein MY, Mandoli GE, Lisi E, et al. ACUTE HF score, a multiparametric prognostic tool for acute heart failure: A real-life study. *International Journal of Cardiology*. déc 2019;296:103-8.
16. Pastore MC, Mandoli GE, Campora A, Renzelli A, Olivoni G, Toscano M, et al. ACUTE HF score predicts in-hospital mortality in patients with acute heart failure. *European Heart Journal*. 9 nov 2023;44(Supplement_2):ehad655.1100.
17. Aujesky D, Obrosky DS, Stone RA, Auble TE, Perrier A, Cornuz J, et al. Derivation and Validation of a Prognostic Model for Pulmonary Embolism. *Am J Respir Crit Care Med*. 15 oct 2005;172(8):1041-6.
18. Jiménez D. Simplification of the Pulmonary Embolism Severity Index for Prognostication in Patients With Acute Symptomatic Pulmonary Embolism. *Arch Intern Med*. 9 août 2010;170(15):1383.
19. Konstantinides SV, Meyer G, Becattini C, Bueno H, Geersing GJ, Harjola VP, et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). *European Heart Journal*. 21 janv 2020;41(4):543-603.
20. Bennett CE, Wright RS, Jentzer J, Gajic O, Murphree DH, Murphy JG, et al. Severity of illness assessment with application of the APACHE IV predicted mortality and outcome trends analysis in an academic cardiac intensive care unit. *Journal of Critical Care*. avr 2019;50:242-6.
21. Jentzer JC, Bennett C, Wiley BM, Murphree DH, Keegan MT, Gajic O, et al. Predictive Value of the Sequential Organ Failure Assessment Score for Mortality in a Contemporary Cardiac Intensive Care Unit Population. *JAHA*. 20 mars 2018;7(6):e008169.
22. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure: On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine (see contributors to the project in the appendix). *Intensive Care Med*. juill 1996;22(7):707-10.
23. Jentzer JC, Anavekar NS, Bennett C, Murphree DH, Keegan MT, Wiley B, et al. Derivation and Validation of a Novel Cardiac Intensive Care Unit Admission Risk Score for Mortality.

JAHA. 3 sept 2019;8(17):e013675.

24. Jentzer JC, Tabi M, Wiley BM, Lanspa MJ, Anavekar NS, Oh JK. Doppler-derived haemodynamics performed during admission echocardiography predict in-hospital mortality in cardiac intensive care unit patients. *European Heart Journal Acute Cardiovascular Care*. 9 août 2022;11(8):640-50.
25. Jentzer JC, Wiley BM, Gersh BJ, Borlaug BA, Oh JK, Anavekar NS. Myocardial contraction fraction by echocardiography and mortality in cardiac intensive care unit patients. *International Journal of Cardiology*. déc 2021;344:230-9.
26. Ramesh A, Kambhampati C, Monson J, Drew P. Artificial intelligence in medicine. *Ann R Coll Surg Engl*. 1 sept 2004;86(5):334-8.
27. Turing AM. I.—Computing machinery and intelligence. *Mind*. 1 oct 1950;LIX(236):433-60.
28. Samuel AL. Some Studies in Machine Learning Using the Game of Checkers. *IBM J Res & Dev*. juill 1959;3(3):210-29.
29. Al'Aref SJ, Anchouche K, Singh G, Slomka PJ, Kolli KK, Kumar A, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *European Heart Journal*. 21 juin 2019;40(24):1975-86.
30. Deo RC. Machine Learning in Medicine. *Circulation*. 17 nov 2015;132(20):1920-30.
31. Pitt B, Pfeffer MA, Assmann SF, Boineau R, Anand IS, Claggett B, et al. Spironolactone for Heart Failure with Preserved Ejection Fraction. *N Engl J Med*. 10 avr 2014;370(15):1383-92.
32. Cohen JB, Schrauben SJ, Zhao L, Basso MD, Cvijic ME, Li Z, et al. Clinical Phenogroups in Heart Failure With Preserved Ejection Fraction. *JACC: Heart Failure*. mars 2020;8(3):172-84.
33. Johnson KW, Torres Soto J, Glicksberg BS, Shameer K, Miotto R, Ali M, et al. Artificial Intelligence in Cardiology. *Journal of the American College of Cardiology*. juin 2018;71(23):2668-79.
34. Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med*. janv 2019;25(1):65-9.
35. Knackstedt C, Bekkers SCAM, Schummers G, Schreckenber M, Muraru D, Badano LP, et al. Fully Automated Versus Standard Tracking of Left Ventricular Ejection Fraction and Longitudinal Strain: The FAST-EFs Multicenter Study. *J Am Coll Cardiol*. 29 sept

2015;66(13):1456-66.

36. Pezel T, Dillinger JG, Trimaille A, Delmas C, Piliero N, Bouleti C, et al. Prevalence and impact of recreational drug use in patients with acute cardiovascular events. *Heart*. nov 2023;109(21):1608-16.
37. Pezel T, Dillinger JG, Trimaille A, Delmas C, Piliero N, Bouleti C, et al. Prevalence of illicit drugs use and association with in-hospital major adverse events in patients hospitalised for acute cardiac events: the ADDICT-ICCU Trial. *European Heart Journal*. 3 oct 2022;43(Supplement_2):ehac544.1482.
38. Dillinger JG, Pezel T, Fauvel C, Delmas C, Schurtz G, Trimaille A, et al. Prevalence of psychoactive drug use in patients hospitalized for acute cardiac events: Rationale and design of the ADDICT-ICCU trial, from the Emergency and Acute Cardiovascular Care Working Group and the National College of Cardiologists in Training of the French Society of Cardiology. *Archives of Cardiovascular Diseases*. oct 2022;115(10):514-20.
39. Aghezzaf S, Coisne A, Hamzi K, Toupin S, Bouleti C, Fauvel C, et al. Machine Learning Model using only Echocardiographic data For Prediction of in-hospital outcomes in ICCU patients. Under review.
40. Fogelson S. datacamp.com. Extreme gradient boosting with XGBoost. Disponible sur: <https://app.datacamp.com/learn/courses/extreme-gradient-boosting-with-xgboost>
41. Lundberg S, Lee SI. A Unified Approach to Interpreting Model Predictions. 2017; Disponible sur: <https://arxiv.org/abs/1705.07874>
42. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, et al. Recommendations for Cardiac Chamber Quantification by Echocardiography in Adults: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Journal of the American Society of Echocardiography*. janv 2015;28(1):1-39.e14.
43. Nagueh SF, Smiseth OA, Appleton CP, Byrd BF, Dokainish H, Edvardsen T, et al. Recommendations for the Evaluation of Left Ventricular Diastolic Function by Echocardiography: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Journal of the American Society of Echocardiography*. avr 2016;29(4):277-314.

VII. Annexes

- **Tableau supplémentaire 1** : ce tableau détaille les différentes métriques utilisées pour comparer les quatre modèles testés. L'*accuracy*, ou exactitude, correspond au taux de prédictions correctes (vrais positifs, et vrais négatifs) sur l'ensemble des prédictions. Elle peut être ajustée en cas de dataset déséquilibré, ce qui est le cas ici puisque seuls 28 patients ont présenté le CJP. Le F score (ou F1 score) est une métrique permettant d'évaluer l'équilibre entre précision et rappel puisqu'il s'agit de la moyenne harmonique entre ces deux paramètres. Il se calcule de la manière suivante : $F1 = 2 \cdot \frac{(\text{Précision} \cdot \text{Rappel})}{(\text{Précision} + \text{Rappel})}$. Le score de Mc Nemar est une mesure statistique permettant d'évaluer si les différences de performances observées entre deux modèles de classification sont statistiquement significatives. Son résultat est exprimé en p-value. Le score de Brier évalue la précision des prédictions en calculant la moyenne des carrés des résidus (différence entre prédiction et observation).
- **Figure supplémentaire 1** : capture d'écran de la feuille d'inclusion, imprimée et disponible à l'USIC du CHU de Lille, entre novembre 2022 et avril 2023 inclus. Cette feuille permettait de limiter les valeurs manquantes sur les ETT d'admission et de surveiller la survenue ou non d'un évènement du critère de jugement composite.

Tableau supplémentaire 1 : Comparaison des différents scores testés

	Logistic regression	LASSO	Random Forest	XG Boost
Balanced accuracy	0.52	0.53	0.67	0.68
F Score	0.10	0.11	0.44	0.46
ROC AUC	0.76 [0.65 - 0.78]	0.79 [0.70 - 0.81]	0.81[0.74 - 0.84]	0.83 [0.75 - 0.86]
Precision-recall AUC	0.39	0.40	0.44	0.48
Specificity	0.98	0.99	0.97	0.98
Sensitivity	0.06	0.07	0.38	0.38
NPV	0.93	0.93	0.95	0.95
PPV	0.33	0.33	0.55	0.60
Mc Nemar test p value	0.003	0.003	0.301	0.181
Brier Score	0.92	0.92	0.88	0.83

Illustration adaptée de l'article d'Aghezzaf et al, under review (39). AUC = area under curve. NPV = negative predictive value. PPV = positive predictive value.

Figure supplémentaire 2 : feuille d'inclusion



Machine Learning Score ETT - USIC

Nom :

Prénom :

Date de naissance :

Date admission :

Date ETT :

Poids : kgs

Taille : cm

surface corporelle m²

ETT

1. VTDVG mL

2. FEVG %

3. ITV sous-aortique cm

4. valvulopathie > grade 2/4 oui non

5. E/A : E/e' moy :

6. Volume d'oreillette gauche mL mL/m²

7. Ventricule droit dilaté oui non

8. TAPSE mm

9. PAPs + mmHg

Synthèse des évènements

1. Décès oui non date : cause :

décès de cause cardiovasculaire : oui non

2. choc cardiogénique nécessitant amines ou assistance oui non

3. Arrêt cardio-respiratoire ressuscité oui non

Remarques

Illustration adaptée de la feuille d'inclusion disponible dans le service d'USIC du CHU de Lille entre novembre 2022 et avril 2023 inclus. VTDVG = volume télédiastolique du ventricule gauche. FEVG = fraction d'éjection ventriculaire gauche. ITV = intégrale temps-vitesse. TAPSE = tricuspide annular plane systolic excursion. PAPs = pression artérielle pulmonaire systolique.

AUTEUR : Nom : Grenier-Boley

Prénom : Jean

Date de soutenance : 27 juin 2024

Titre de la thèse : Validation externe d'un modèle obtenu par Machine Learning sur des données échographiques pour évaluer le pronostic des patients admis enUSIC : cohorte lilloise prospective.

Thèse - Médecine - Lille « 2024 »

Cadre de classement : Cardiologie

DES + FST/option : Médecine cardiovasculaire

Mots-clés :USIC, modèle pronostique, intelligence artificielle, Machine Learning

Résumé :

Introduction. Stratifier le risque d'évènement péjoratif majeur chez les patients admis enUSIC reste difficile. Des scores pronostiques existent mais ceux-ci ont été développés pour des populations d'unités de soins intensifs non cardiologiques ou pour des pathologies cardiaques spécifiques. Par ailleurs ils n'intègrent que peu ou pas les paramètres échocardiographiques. Un modèle pronostique (ML XG-Boost) réalisable dès l'admission a été créé à l'aide de paramètres échocardiographiques seuls et construit via Machine Learning. Ce modèle a montré de bonnes performances dans la cohorte multicentrique française de dérivation initiale (ADDICT-ICCU). L'objectif de ce travail était d'en réaliser sa validation externe.

Matériel et méthodes. Il s'agit d'une étude prospective observationnelle monocentrique menée à l'USIC du CHU de Lille. Les patients adultes admis à l'USIC étaient incluables et l'ETT était réalisée à l'admission. Le critère de jugement principal était la survenue d'un évènement clinique intra-hospitalier majeur parmi le décès toute cause, l'arrêt cardiaque ressuscité ou la survenue d'un choc cardiogénique dans les 28 jours suivant l'admission.

Résultats. Dans notre cohorte de 200 patients inclus à l'USIC du CHU de Lille entre novembre 2022 et avril 2023 inclus, 135 (67,5%) étaient des hommes, l'âge moyen était de $62,6 \pm 16,6$ ans, 122 patients étaient hypertendus (61,3%), 60 (30,2%) étaient diabétiques. La FEVG moyenne était à $49,1 \pm 16,2$ %, le rapport E/E' moyen à $11,1 \pm 4,74$. L'ITV sous-aortique était en moyenne à $14,7 \pm 4,72$ cm, le TAPSE à $19 \pm 5,01$ mm, la PAPs était en moyenne à $41,7 \pm 15,6$ mmHg. 28 patients (14%) ont présenté un évènement cardiovasculaire grave à 28 jours. On note 10 décès (5%), 18 chocs cardiogéniques (9%) et 9 arrêts cardiaques récupérés (4,5%). Dans cette population, les performances du modèle ML XG-Boost restaient supérieures aux scores préexistants (AUROC 0,74, $p < 0,001$).

Conclusion. Les performances du modèle ML XG-Boost dans une cohorte externe lilloise sont satisfaisantes et permettent sa validation externe. De plus amples études doivent permettre de corroborer ces résultats.

Composition du Jury :

Président :

Monsieur le Professeur Christophe BAUTERS

Assesseurs :

Monsieur le Professeur Gilles LEMESLE

Monsieur le Professeur Augustin COISNE

Directeur de thèse :

Monsieur le Docteur Samy AGHEZZAF