



**HAL**  
open science

# Modèle pronostique de machine learning supervisé, incluant des facteurs environnementaux, pour prédire la sévérité intra-hospitalière chez les patients admis pour insuffisance cardiaque aiguë

Benjamin Sibia

## ► To cite this version:

Benjamin Sibia. Modèle pronostique de machine learning supervisé, incluant des facteurs environnementaux, pour prédire la sévérité intra-hospitalière chez les patients admis pour insuffisance cardiaque aiguë. Médecine humaine et pathologie. 2024. dumas-04741070

**HAL Id: dumas-04741070**

**<https://dumas.ccsd.cnrs.fr/dumas-04741070v1>**

Submitted on 17 Oct 2024

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

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



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License

## **AVERTISSEMENT**

Cette thèse d'exercice est le fruit d'un travail approuvé par le jury de soutenance et réalisé dans le but d'obtenir le diplôme d'Etat de docteur en médecine. Ce document est mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt toute poursuite pénale.

Code de la Propriété Intellectuelle. Articles L 122.4

Code de la Propriété Intellectuelle. Articles L 335.2 - L 335.1

## UNIVERSITÉ PARIS CITÉ

Faculté de Santé

UFR de Médecine

Année 2024

N°

### THÈSE POUR LE DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE

Présentée et soutenue publiquement le : 17/05/2024

Par

**Benjamin SIBILIA**

**Modèle pronostique de machine learning supervisé, incluant des facteurs environnementaux, pour prédire la sévérité intra-hospitalière chez les patients admis pour insuffisance cardiaque aiguë.**

Dirigée par M. le Docteur Théo Pezel

#### JURY

M. le Professeur Patrick Henry

Président

M. le Docteur Théo Pezel

Directeur

M. le Professeur Damien Logeart

Membre du jury

Mme la Docteure Marie-Christine Iliou

Membre du jury

Mme la Docteure Florence Beauvais

Membre du jury

## Le Serment d'hippocrate

---

Je promets et je jure d'être fidèle aux lois de l'honneur et de la probité dans l'exercice de la Médecine.

Je respecterai toutes les personnes, leur autonomie et leur volonté, sans discrimination.

J'interviendrai pour les protéger si elles sont vulnérables ou menacées dans leur intégrité ou leur dignité. Même sous la contrainte, je ne ferai pas usage de mes connaissances contre les lois de l'humanité.

J'informerai les patients des décisions envisagées, de leurs raisons et de leurs conséquences. Je ne tromperai jamais leur confiance.

Je donnerai mes soins à l'indigent et je n'exigerai pas un salaire au-dessus de mon travail.

Admis dans l'intimité des personnes, je tairai les secrets qui me seront confiés et ma conduite ne servira pas à corrompre les mœurs.

Je ferai tout pour soulager les souffrances. Je ne prolongerai pas abusivement la vie ni ne provoquerai délibérément la mort.

Je préserverai l'indépendance nécessaire et je n'entreprendrai rien qui dépasse mes compétences. Je perfectionnerai mes connaissances pour assurer au mieux ma mission.

Que les hommes m'accordent leur estime si je suis fidèle à mes promesses. Que je sois couvert d'opprobre et méprisé si j'y manque.

## Remerciements

---

*60 ans après la thèse de mon Grand-père, Gaston SIBILIA, « Contribution à l'étude des tumeurs primitives de la paroi thoracique. (À propos de 17 observations personnelles) ».*

*33 ans après la thèse de mon oncle, Jean SIBILIA, « Contribution à l'étude des anticorps anti Ro/SS-A et anti La/SS-B. Etude clinique et biologique de 50 patients ».*

*À M. le Docteur Théo PEZEL. On m'a souvent dit qu'au cours de son internat, on fait à un moment donné une rencontre importante. Je pense que ta rencontre a fait basculer mon internat. Merci pour cette aventure que représente la thèse et que de travailler avec toi. Merci pour ta confiance, ton encadrement, ta disponibilité et ton souci de l'excellence qui nous pousse à donner le meilleur de nous même. Amitiés.*

*À M. le Professeur Patrick HENRY. Pour la confiance que vous m'avez porté pendant ce semestre chez vous qui m'a permis à la fois de m'épanouir dans mon rôle d'interne mais aussi de prendre mes responsabilités. Merci pour votre accessibilité, pour vos paroles toujours bienveillantes à mon égard et vos conseils avisés.*

*À M. le Professeur Damien LOGEART. Pour votre enseignement lors de mon passage dans votre service et votre investissement pour les sociétés savantes de cardiologie en dehors de votre activité de service. Comment ne pas vous associer à ce travail qui s'intéresse à votre sujet d'expertise.*

*À Mme la Docteure Florence BEAUVAIS. Pour votre gentillesse, votre ouverture d'esprit et nos échanges toujours aussi instructifs. Pour les beaux projets à construire ensemble au service des patients.*

*À Mme la Docteure Marie-Christine ILIOU. Je n'oublierai jamais mon stage dans votre service où j'ai fait mes premiers pas d'interne. Grâce à vous, j'ai pu apprendre à m'occuper des patients dans leur globalité et à découvrir aussi pour la première fois le monde de la recherche portée sur vos patients sous assistance cardiaque.*

*A Madame la Docteure en sciences Solenn TOUPIN et Madame Kenza HAMZI. Un grand merci pour votre temps et votre aide si importante pour la réalisation des statistiques de ce*

*travail de thèse. J'ai beaucoup apprécié nos échanges et nos discussions. Une belle collaboration entre biostatisticiens/ingénieurs et médecin.*

**À M. le Professeur Olivier Varenne et M. le Professeur Karim Wahbi.** *Je n'oublierai jamais l'équipe qui m'a fait découvrir et donné le goût à la cardiologie lors de mon externat. Terminer mon internat dans votre équipe représentait un moment important pour moi. Merci pour votre enseignement, l'accueil dans votre équipe et votre intelligence humaine.*

**À Toute l'Equipe Médicale de Cardiologie de l'hôpital Lariboisière-Saint Louis : Pr Alain COHEN-SOLAL, Dr Jean Guillaume DILLINGER, Dr Mathilde BAUDET, Dr Antoine LEQUIPAR, Dr Trécy GONCALVES, ainsi que tout l'équipe paramédicale de cardiologie de l'hôpital Lariboisière, Mariette, Tran et Christian.**

*Pour leur accueil, l'atmosphère familiale et leur encadrement depuis mon passage comme interne à Lariboisière. Merci infiniment pour la confiance dont vous m'avez fait preuve.*

**À Toute l'Equipe Médicale de Cardiologie de l'hôpital Cochin : Dr Fabien PICARD, Dr Laure Cabane, Dr Leslie MOTTÉ, Dr Quentin FISCHER et Dr Pierre BRAMI, ainsi que tout l'équipe paramédicale de cardiologie de l'hôpital Cochin, Jean-Louis et Hélène LE-LONG.** *Encore merci pour ces six mois passés chez vous dont je garderai un souvenir impérissable.*

**À M. le Docteur Emmanuel MAHÉ.** *Je n'oublierai jamais mon passage dans votre service de dermatologie où j'ai pu découvrir toute la beauté de votre spécialité et où j'ai pu apprendre à m'affirmer et exprimer mes éventuels désaccords. Vos bons mots continueront à m'accompagner. Merci pour votre enseignement !*

**À M. le Professeur Francis SCHNEIDER et M. le Professeur Charles-Edouard LUYT.** *Pour m'avoir ouvert les portes de leur réanimation de pointe dans le domaine de la transplantation hépatique et cardiaque.*

**À M. le Professeur Gilles MONTALESCOT et M. le Professeur Albert HAGÈGE.** *Pour avoir eu l'opportunité de travailler dans deux centres où tout ce qui est techniquement possible est réalisable.*

*À mes parents, sans qui rien de tout ce qui m'entoure ne serait possible.*

*À ma Maman. Pour sa beauté, sa douceur, sa patience, son rire et son amour mais aussi pour les difficultés que l'on a traversées. Merci pour m'avoir aidé à devenir l'homme que je suis. Je t'aime tant.*

*À mon Papa. Pour l'histoire de sa vie et des choix qu'elle nous fait tous prendre. Pour ta vision sur notre place dans le Monde et ton avis toujours éclairé. Je t'aime tant.*

*À mes sœurs chéries Laetitia, Clémentine et Pauline. Pour toutes ces épreuves traversées ensembles et ce lien indéfectible qui nous unit à vie. Merci de votre soutien malgré parfois mon caractère orageux. Je suis si fier de vous trois.*

*À mon beau-père Christophe. Pour ton amour, ta droiture et ton soutien sans faille depuis maintenant presque 25 ans. Je t'admire tant.*

*À ma belle-mère Karine. Pour ta douceur et ta gentillesse mais aussi pour les misères que j'ai pu te faire quand j'étais petit... Le temps ne saurait effacer tout l'affection que nous avons l'un envers l'autre.*

*À Audrey, Mathilde, Giorgia, Robin, Louis, Fabien et Elliott. On forme une si belle et grande famille.*

*À mes grands-parents de sang et de cœur. Pour tout leur amour, leur gentillesse, leur dévotion et leur présence ô combien précieuse. Vous avez encore tant à m'apprendre. Je ne peux m'empêcher de penser à Dady et à Grand-père qui nous ont déjà quitté mais qui ont une place au fond de mon cœur.*

*À Tonton Jean, mon oncle chéri, avec qui j'ai la chance de partager ce lien si particulier de neveu, filleul et maintenant de confrère. Tu sais bien le modèle que tu représentes pour moi. C'est une vraie chance d'avoir pu créer un lien si unique ! Ton neveu chéri.*

*À Dominique ma marraine, ma tante Anne, mon oncle Gilles et à mes cousins Arthur, Solène, Antoine, Aurélien, Marine, Achille, Théodore, Margot, Gaspard, Romain et Nicolas. Pour tous ces moments heureux vécus ensembles à Paris, en Bretagne, en Alsace ou à la montagne. De belles cousinades en perspective !*

**À mon ami Jean, mon meilleur ami, et à Hélène sa femme.** Merci pour ces 12 années d'amitié si riches et intenses. Quand on y repense c'est fou... Un mélange de chance, de fierté et de reconnaissance. A tous nos beaux projets réalisés et surtout aux suivants ! Rock on !

**À Augustin, mon plus fidèle copain.** De la classe de 4<sup>ème</sup> à la sixième année de médecine on aura décidément toujours été dans la même salle de classe, le même amphithéâtre, le même pôle. Pour tous nos si beaux souvenirs (de Rome en 2008 aux révisions de l'ECN en passant par les vendanges), l'accueil de ta famille et pour les futurs moments à partager. Merci Gucht !

**À Julien, mon plus vieux copain.** Merci pour ta fidélité dans l'amitié depuis maintenant 25 ans malgré la distance et nos emplois du temps surchargés. Pour tous nos souvenirs d'enfance à la campagne, les matchs de foot du dimanche matin, le séjour à Munich, la Bretagne et sans oublier les vacances mythiques à Sainte-Foy. Sans oublier tes parents Jane et Pierre qui ont toujours été comme une famille pour moi.

**À Amaury, Arnaud, FH, Sébastien, Alexandre, Côme, François et Maxime.** Quelle chance de vous avoir comme amis depuis maintenant presque 15 ans ! En particulier, à Maxime toi qui m'apprends jour après jour que rien n'est impossible. A François pour ton calme à toute épreuve et ton humilité. A Côme pour l'ami fidèle que tu es et le beau projet de vie que tu construis avec Julienne. A Alexandre pour ton eclectisme à l'heure du rap et du conformisme social et bien sûr à l'amour que l'on partage pour la musique et le Cinéma. A FH et Sébastien, pour votre sens de la fête et toutes les conneries que vous arrivez encore à faire. A Amaury et Arnaud avec qui j'ai eu la chance de partager plus d'un an de vie et le confinement à la colocation. Que serais-je devenu sans vous et votre soutien dans mon début d'internat difficile. Merci à vous huit !

**À Mathilde, Noémie, Alexander, Hubert, Marjolaine et Rébecca.** Pour l'amitié que l'on partage depuis maintenant plus de dix ans. Merci pour votre soutien, votre fidélité, nos fous-rires, nos souvenirs et nos discussions si importants pour moi.

**À mes amis de l'externat Benjamin, Valentin, Louis, Julien, Pierre et Constantin.** Pour tous ces souvenirs pendant ces années d'étude en passant par les soirées AMPC, les matchs de foot, le WEI, le ski-fac, les filles.. Je ne vous oublie pas.

**À Charles, Iyad et Manveer.** Pour toutes ces heures, week-ends et semaines passées ensemble en stage à l'hôpital et surtout cette solidarité et l'esprit d'entraide qui en sont sortis. Comment vous oublier mes « copains de galère » ?!

*À **Jeremy et Alexandre.** Pour ces derniers mois passés ensemble dans le cadre du Master et la belle amitié qui naît !*

*À tous mes cointernes de ma promo de cardiologie et à ceux de dermatologie et de réanimation : **Quentin, Déborah, Théo, Pierre et Nicolas.** Pour tous les bons moments passés ensemble pendant et en dehors des stages.*

*À **Céline LUC et Gilles ZAMARIA.** Céline, pour ta gentillesse et pour m'avoir fait découvrir le monde de la cardiologie de ville. Gilles, merci pour votre accueil et nos discussions. C'est une chance de pouvoir échanger avec vous.*

*À **mon Amoureuse.** Pour notre rencontre et la complicité que l'on partage depuis. Merci pour ta douceur et toute l'affection que tu me portes ! Tu as été un vrai soutien au cours de ces derniers mois. Je nous souhaite de poursuivre ce petit bout de chemin ensemble et de réaliser de beaux projets ! « Je t'apprécie beaucoup tu sais ».*

# Résumé et indexation en français

---

**Modèle pronostique de machine learning supervisé, incluant des facteurs environnementaux, pour prédire la sévérité intra-hospitalière chez les patients admis pour insuffisance cardiaque aiguë.**

## Résumé :

**Objectif:** Étudier la faisabilité et la robustesse d'un modèle de machine learning (ML) supervisé incluant des facteurs environnementaux pour prédire les événements indésirables majeurs intra-hospitaliers chez les patients hospitalisés pour insuffisance cardiaque aiguë aux soins intensifs.

**Méthode:** Nous avons sélectionné les patients admis pour insuffisance cardiaque aiguë issus du registre ADDICTO-USIC, une étude multicentrique prospective nationale française qui a inclus tous les patients consécutifs hospitalisés aux soins intensifs de cardiologie. Le critère de jugement principal était un critère composite de survenu d'événements majeurs intra-hospitaliers définis comme le décès, l'arrêt cardiaque récupéré ou le choc cardiogénique nécessitant une assistance médicamenteuse ou circulatoire. Plusieurs algorithmes de ML ont été entraînés puis comparés à un outil de statistique traditionnelle pour sélectionner les variables prédictives et construire notre modèle dans un échantillon d'entraînement. Les performances ont ensuite été validées dans un échantillon de validation et comparés à l'ACUTE-HF score.

**Résultats:** Parmi les 459 patients inclus (âge  $68 \pm 14$  ans, 68 % d'hommes), 47 soit 10,2% de la cohorte ont présenté un événement majeur intra-hospitalier. Sept variables ont été sélectionnées par la régression de LASSO (*Least absolute shrinkage and selection operator*) pour prédire l'événement dans le *train-set* (N=322) : la pression artérielle moyenne, l'étiologie ischémique, l'ITV sous-aortique, le rapport E/e', le TAPSE, l'utilisation de drogues récréatives et le taux de CO expiré. Le Random forest a présenté les meilleures performances (AUROC=0,82, PR-AUC=0,48, F1-score=0,56). Enfin, notre modèle a été comparé à l'ACUTE-HF score et a montré une meilleure ROC-AUC (random forest : 0,82 vs Acute HF-score : 0,57 ;  $p < 0,001$ ).

**Conclusion:** Notre modèle de ML, incluant sept variables dont deux variables environnementales, a montré les meilleures performances pour prédire les événements majeurs intra-hospitaliers chez les patients admis pour insuffisance cardiaque aiguë.

## Spécialité :

Médecine cardiovasculaire

## Mots clés français :

Insuffisance cardiaque aiguë ; facteurs environnementaux ; monoxyde de carbone; drogues récréatives; machine learning; soins intensifs de cardiologie échocardiographie.

## Forme ou Genre :

fMeSH : Dissertation universitaire

Rameau : Thèses et écrits académiques

## Résumé et indexation en anglais

---

### **Supervised Machine Learning using Environmental Factors to Predict In-Hospital Outcomes in Acute Heart Failure Patients.**

**Abstract:**

**Background:** While few traditional scores are available for risk stratification of patients hospitalized for acute heart failure (AHF), none of these scores include in particular environmental factors and the potential benefit of machine-learning (ML) is not well-established.

**Objectives:** To investigate the feasibility and accuracy of a supervised ML-model including environmental factors to predict in-hospital major adverse events (MAE) in patients hospitalized for AHF.

**Methods:** In April 2021, a French national prospective multicenter study included all consecutive patients hospitalized in ICCU. Patients admitted for AHF were included in the analyses. A ML-model involving automated feature selection by Least absolute shrinkage and selection operator (LASSO) and model building with a Random Forest (RF) algorithm was developed. The primary composite outcome was in-hospital MAE defined by death, resuscitated cardiac arrest, or cardiogenic shock requiring assistance.

**Results:** Among 459 patients included (age  $68 \pm 14$  years, 68% male), 47 experienced in-hospital MAE (10.2%). Seven variables were selected by LASSO for predicting MAE in the training dataset (N=322): mean arterial pressure, ischemic etiology, sub-aortic velocity time integral,  $E/e'$ , tricuspid annular plane systolic excursion, recreational drug use and exhaled carbon monoxide level. The RF model showed the best performance compared with other evaluated models (AUROC=0.82, PR-AUC=0.48, F1-score=0.56). Our ML-model exhibited a higher AUROC compared with an existing score for prediction of MAE (AUROC for our ML-model: 0.82 vs Acute HF-score: 0.57;  $p < 0.001$ ).

**Conclusion:** Our ML-model including in particular environmental variables exhibited a better performance than traditional statistical methods to predict in-hospital outcomes in patients admitted for AHF.

**English keywords:**

MeSH: acute heart failure; environmental factors; carbon monoxide; recreational drugs; machine learning; intensive cardiac care unit; echocardiography.

**Publication type:**

MeSH : Academic Dissertation

## Liste des abréviations

---

- AIT : accident ischémique transitoire
- AMPcyclique : adénosine-monophosphate cyclique
- ARA2 : antagoniste aux récepteurs de l'angiotensine 2
- ATCD : antécédent
- AUC : *area under curve*
- AVC : accident vasculaire cérébral
- BNP: *Brain natriuretic peptide*
- BPCO : bronchopneumopathie chronique obstructive
- CBD: cannabidiol
- CO: monoxyde de carbone
- ECG: électrocardiogramme
- ESC : *European Society of Cardiology*
- FA : fibrillation atriale
- FC: fréquence cardiaque
- FEVG : fraction d'éjection du ventricule gauche
- HbCO : carboxyhémoglobine
- HR : *Hazard ratio*
- HVG : hypertrophie ventriculaire gauche
- IA : intelligence artificielle
- IC: insuffisance cardiaque
- IEC : inhibiteurs de l'enzyme de conversion
- IMC : indice de masse corporelle
- LASSO : *Least Absolute Shrinkage and Selection Operator*
- LVAD : *left ventricular assist devices*
- MAE : *major adverse events*
- ML : *machine learning*
- NT-ProBNP: *N-terminal pro-B natriuretic peptide*
- NYHA: *New York heart association*
- PR-curve: precision recall curves

- PR-AUC: area under the PR curve
- RF: *Random forest*
- ROC: *receiver operating characteristics*
- ROC-AUC: area under ROC curve
- SpO2: saturation en oxygène
- SRA: système rénine-angiotensine
- THC: tétrahydrocannabinol
- USIC: unite de soins intensifs de cardiologie
- VD: ventricule droit
- VES: volume d'éjection du ventricule gauche
- VG : ventricule gauche
- VNI: ventilation non-invasive
- XGBoost: *extreme gradient boosting*
- 95%CI: intervalle de confiance à 95%

# Table des matières

---

<b>LE SERMENT D'HIPPOCRATE.....</b>	<b>1</b>
<b>REMERCIEMENTS.....</b>	<b>2</b>
<b>RESUME ET INDEXATION EN FRANÇAIS .....</b>	<b>7</b>
<b>RESUME ET INDEXATION EN ANGLAIS.....</b>	<b>8</b>
<b>LISTE DES ABREVIATIONS.....</b>	<b>9</b>
<b>PARTIE 1: INTRODUCTION.....</b>	<b>13</b>
<b>1 : INSUFFISANCE CARDIAQUE .....</b>	<b>14</b>
1.1 ÉPIDEMIOLOGIE .....	14
1.2 DEFINITION.....	14
1.3 PHYSIOPATHOLOGIE .....	15
1.3.1 Généralités .....	15
1.3.2 Mécanismes adaptatifs .....	16
1.3.3 Remodelage pathologique .....	17
1.4 DEFINITION DE L'ESC .....	17
1.5 DIAGNOSTIC CLINIQUE.....	18
1.6 DIAGNOSTIC PARACLINIQUE .....	19
1.7 TRAITEMENTS.....	21
1.7.1 Généralités .....	21
1.7.2 Insuffisance cardiaque aigüe.....	22
1.7.3 Insuffisance cardiaque chronique .....	23
<b>2 : FACTEURS ENVIRONNEMENTAUX ISSUS DE LA COHORTE ADDICTO-USIC .....</b>	<b>25</b>
2.1 POSITIONNEMENT DU PROBLEME .....	25
2.2 DESCRIPTION DE LA BASE DE DONNEES .....	25
2.2.1 Méthode de détection des drogues récréatives .....	25
2.2.2 Méthode de détection du CO expiré .....	26
2.3 AMPHETAMINES ET INSUFFISANCE CARDIAQUE .....	28
2.4 COCAÏNE ET INSUFFISANCE CARDIAQUE .....	28
2.5 CANNABIS ET INSUFFISANCE CARDIAQUE .....	30
2.6 MONOXIDE DE CARBONE ET INSUFFISANCE CARDIAQUE .....	31
<b>3 : INTELLIGENCE ARTIFICIELLE, MACHINE LEARNING SUPERVISE .....</b>	<b>32</b>

3.1 DEFINITION.....	32
3.2 PRINCIPES GENERAUX.....	33
3.3 FONCTIONNEMENT.....	34
3.3.1 Dataset : gestion des données manquantes.....	34
3.3.2 Gestion des données déséquilibrées.....	35
3.3.3 Sélection de variables.....	36
3.3.4 Évaluation d'un modèle pronostique.....	37
3.4 APPLICATIONS EN CARDIOLOGIE ET SCORES EXISTANTS.....	39
3.4.1 Méthodologie des scores.....	39
3.4.2 Intérêt du score en pratique clinique.....	39
3.4.3 Scores traditionnels.....	40
3.4.4 Score traditionnel mixte.....	44
3.4.5 Scores de ML.....	45
<b>4 : OBJECTIF ET RATIONNEL DE L'ETUDE.....</b>	<b>48</b>
<b>PARTIE 2 : ARTICLE MÉDICAL SCIENTIFIQUE.....</b>	<b>49</b>
<b>PARTIE 3 : DISCUSSION.....</b>	<b>94</b>
1. VARIABLES ENVIRONNEMENTALES.....	94
2. VARIABLES CLINIQUES.....	95
3. VARIABLES ECHOGRAPHIQUES.....	95
4. UTILISATION DU MACHINE LEARNING.....	96
5. LIMITES.....	96
<b>CONCLUSION.....</b>	<b>98</b>
<b>BIBLIOGRAPHIE.....</b>	<b>99</b>
<b>TABLE DES FIGURES.....</b>	<b>104</b>
<b>TABLE DES TABLEAUX.....</b>	<b>105</b>
<b>ANNEXES.....</b>	<b>106</b>

## **Partie 1: Introduction**

---

# 1 : Insuffisance cardiaque

---

## 1.1 Épidémiologie

L'**insuffisance cardiaque (IC)** représente actuellement un **défi majeur en santé publique**. En effet, la **mortalité cardiovasculaire** est la **première cause de mortalité chez la femme** et la **deuxième cause de mortalité chez l'homme**. (1)

En 2022, selon les rapports du CépiDC-INSERM-Santé Publique France, on dénombre près d'un million de patients souffrant d'IC en France avec 184 976 hospitalisations en 2019 et 69 551 décès en 2016.(2) Ces chiffres traduisent une **véritable progression de l'incidence de l'IC en France**. En effet, en comparaison, en 2006, l'INSEE rapportait 131 660 hospitalisations et 70 589 décès dus à l'IC.(2) Il ne s'agit pas d'une particularité française puisque **à l'échelle mondiale** selon l'Organisation mondiale de la Santé, les maladies cardiovasculaires sont responsables de **17,7 millions** de décès chaque année et représentent 31% des décès.(3)

Les principales étiologies d'IC en France sont : (4)

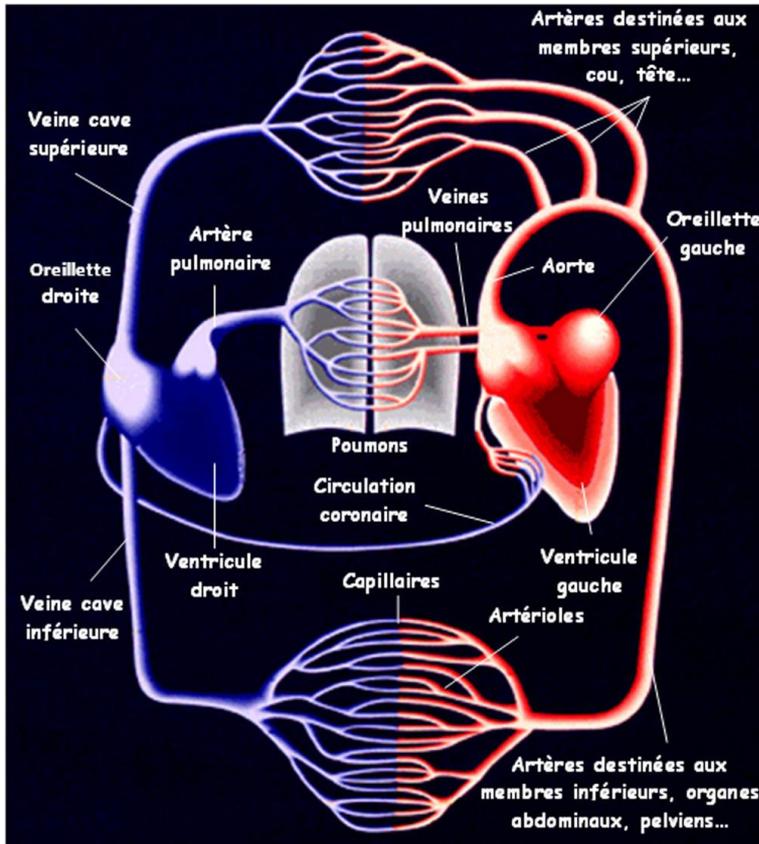
- l'infarctus du myocarde
- les cardiopathies hypertensives
- les troubles du rythme cardiaque avec en premier lieu la fibrillation atriale (FA)
- les valvulopathies

## 1.2 Définition

L'**IC** est une pathologie complexe comme en témoignent les différentes évolutions des recommandations de la société européenne de cardiologie (**ESC, European Society of Cardiology**) ces dernières années.(5–7)

L'ESC la définit comme étant « un état pathologique indiquant que la pompe cardiaque n'est plus capable d'assurer un débit sanguin suffisant pour satisfaire les besoins de l'organisme ».(8) Il peut s'agir alors soit d'une IC de novo, soit d'une aggravation de ses symptômes à l'issu d'un facteur déclenchant.

L'IC correspond donc à un état où la masse sanguine n'est plus délivrée de manière suffisante à l'organisme via le système cardiovasculaire (**Figure 1**). Il en résulte alors un bas-débit qui se manifeste par une **congestion** et/ou des signes d'**hypovolémie**.



**Figure 1** : Représentation de l'appareil cardiovasculaire. D'après *Chung, MK., and Rich, M.W.*  
*Introduction au système cardiovasculaire.*

## 1.3 Physiopathologie

### 1.3.1 Généralités

La physiopathologie de l'IC réside en un **déséquilibre entre les apports et les besoins en oxygène** de l'organisme en lien avec une atteinte du débit cardiaque, classiquement un bas débit, et la mise en jeu de mécanismes adaptatifs localement et à l'échelle de l'organisme par l'intermédiaire du système adrénergique et du système rénine-angiotensine (SRA)-aldostérone.(9)

Le bas débit peut s'évaluer en raisonnant à partir de l'équation du débit cardiaque noté Q. Elle s'écrit sous la forme suivante :  $Q = FC \times VES$ , où FC correspond à la fréquence cardiaque et VES au volume d'éjection systolique. Le volume d'éjection systolique correspond à la différence entre les volumes télédiastolique et télésystolique du ventricule gauche (VG).

On comprend alors que ces deux composantes jouent un rôle essentiel dans l'adaptation du débit cardiaque.

### 1.3.2 Mécanismes adaptatifs

#### a) la fréquence cardiaque

La **tachycardie** est le premier réflexe pour adapter le débit mais peut être la cause de la dysfonction cardiaque. En cas d'élévation importante de la fréquence cardiaque comme au cours d'une tachycardie ventriculaire, la diastole se raccourcit ce qui entraîne un remplissage du VG insuffisant et par conséquent une réduction du VES. De plus, cette réduction du temps diastolique participe à une hypoperfusion myocardique du fait que la perfusion coronaire intervient uniquement lors de la diastole. La FC en cas de bradycardie lente, comme lors des troubles de conduction, est à l'origine du bas débit, d'autant plus s'il prédomine une cardiopathie sous-jacente ou que le patient reçoit des inotropes négatifs au long cours.

#### b) le volume d'éjection systolique

**Le VES possède trois principaux déterminants** dont un seulement (l'inotropisme) est indépendant des conditions de charge(10) :

**La capacité intrinsèque du myocyte à se contracter**, ou inotropisme, intervient directement sur le volume télésystolique. Les deux déterminants majeurs sont la qualité du myocarde (et donc la présence de nécrose myocardique) et le système sympathique par son action inotrope et chronotrope sur le myocarde.

La **précharge** correspondant à la volémie et à la pression télédiastolique du VG qui est régie par la loi de Franck-Starling. Elle intervient directement sur le volume télédiastolique du VG et est liée à la qualité de relaxation du muscle cardiaque. Plus la précharge augmente, plus les fibres du myocyte sont mises en tension et plus la force de contraction est importante jusqu'à atteindre un plateau. La dilatation du ventricule va entraîner une nécrose myocardique. La

dilatation du ventricule droit (VD) en cas d'élévation des pressions de remplissage gauche va diminuer le VES ventriculaire droit et va gêner la diastole ventriculaire gauche. On parle alors d'interdépendance VG-VD puisque les débits pulmonaire et aortique sont égaux en l'absence de shunt.

La **postcharge** correspondant aux forces de pression s'opposant à l'éjection du sang du ventricule. Il s'agit de la circulation systémique pour le **VG** et de la circulation pulmonaire pour le **VD**. Le système sympathique par son action vasoconstrictrice permet certes une redistribution du volume vasculaire au profit de la perfusion cérébrale et coronaire, mais augmente les résistances périphériques et par conséquent la post charge du VG. A cela s'ajoute le **SRA** qui augmente la postcharge par rétention hydrosodée au niveau du rein et augmente les résistances périphériques par vasoconstriction sous l'action de l'angiotensine 2.

### 1.3.3 Remodelage pathologique

En cas d'IC, les conditions de charge volumiques et barométriques entraînent donc un **remodelage** répondant à la loi de Franck-Starling par la dilatation du VG mais aussi de la Loi de Laplace entraînant un épaissement de la paroi pour contrebalancer la tension pariétale. La **libération de catécholamines** tels que la noradrénaline et l'adrénaline, permet certes initialement de maintenir une hémodynamique efficace en jouant sur la fréquence cardiaque et le VES, mais est connue pour entraîner une désensibilisation des récepteur bêta-adrénergiques ainsi qu'un remodelage la matrice cellulaire (hypertrophie) et extracellulaire (production de fibrose).(11)

**Le SRA altère le remodelage myocardique en augmentant la postcharge par rétention hydrosodée au niveau du rein et augmente les résistances périphériques.(11)**

## 1.4 Définition de l'ESC

Depuis la publication les recommandations de l'ESC en 2021 on distingue trois grands types d'IC qui sont représentés dans le **tableau 1** : (6)

- l'IC avec **altération de la fraction d'éjection du ventricule gauche (FEVG)** qui correspond schématiquement à une force de contraction cardiaque diminuée.
- l'IC à **FEVG modérément réduite** laquelle rend compte d'une distinction plus récente située en zone intermédiaire entre les deux types d'IC précédents.

- l'IC à FEVG préservée avec une capacité au remplissage altérée mais une force de contraction cardiaque conservée.

Type d'IC	IC-FEr	IC-FEmr	IC-FEp
Critère 1	Symptômes +/- signes <sup>a</sup>	Symptômes +/- signes <sup>a</sup>	Symptômes +/- signes <sup>a</sup>
Critère 2	FEVG ≤ 40%	FEVG 41-49% <sup>b</sup>	FEVG ≥ 50%
Critère 3	-	-	Mise en évidence d'une anomalie morphologique et/ou fonctionnelle cardiaque par la présence d'une dysfonction diastolique, d'une élévation des pressions de remplissage ou d'une élévation des peptides natriurétiques <sup>c</sup>

**Tableau 1:** Définition de l'insuffisance cardiaque à FEVG préservée, à FEVG modérément réduite et à FEVG réduite adapté selon les recommandations ESC 2021.(6)

*FEVG : fraction d'éjection ventriculaire gauche ; IC : Insuffisance cardiaque ; IC-Fer : Insuffisance cardiaque à FEVG réduite ; IC-FEmr : Insuffisance cardiaque à FEVG modérément altérée ; IC-FEp : Insuffisance cardiaque à FEVG préservée*

<sup>a</sup> *Signes pouvant ne pas être présents au stade précoce de la maladie ou sous traitement optimal*

<sup>b</sup> *Pour le diagnostic de l'IC-FEmr, la présence de plusieurs signes d'atteinte cardiaque structurale (oreillette gauche dilatée, hypertrophie du ventricule gauche ou élévation des pressions de remplissage du ventricule gauche) rend le diagnostic plus probable.*

<sup>c</sup> *Pour le diagnostic d'IC-FEp, plus le nombre d'anomalie est important, plus le diagnostic d'IC-FEp est probable.*

## 1.5 Diagnostic clinique

De même qu'il existe un cœur droit et un cœur gauche, il existe des signes d'IC droite et gauche tout en gardant à l'esprit que la première cause d'IC droite est l'IC gauche :

- **Signes droits** : dyspnée selon la classification NYHA, galop auscultatoire B4 et souffle d'insuffisance tricuspide, signe de Harzer : perception pulsatile du cœur à la palpation sous xiphoïdienne qui témoigne d'une dilatation du VD, une turgescence jugulaire droite, hépatalgie pulsatile et reflux hépato-jugulaire qui témoignent de la congestion en amont de l'oreillette droite, prise de poids avec œdèmes des membres inférieurs bilatéraux, symétriques, prenant le godet et ceux des zones déclives (ascite et hydrocèle à un stade plus tardifs)
- **Signes gauches** : orthopnée, lipothymies, galop B3 et B4, crépitations bilatéraux, choc de pointe et signes d'insuffisance cardiaque droite.

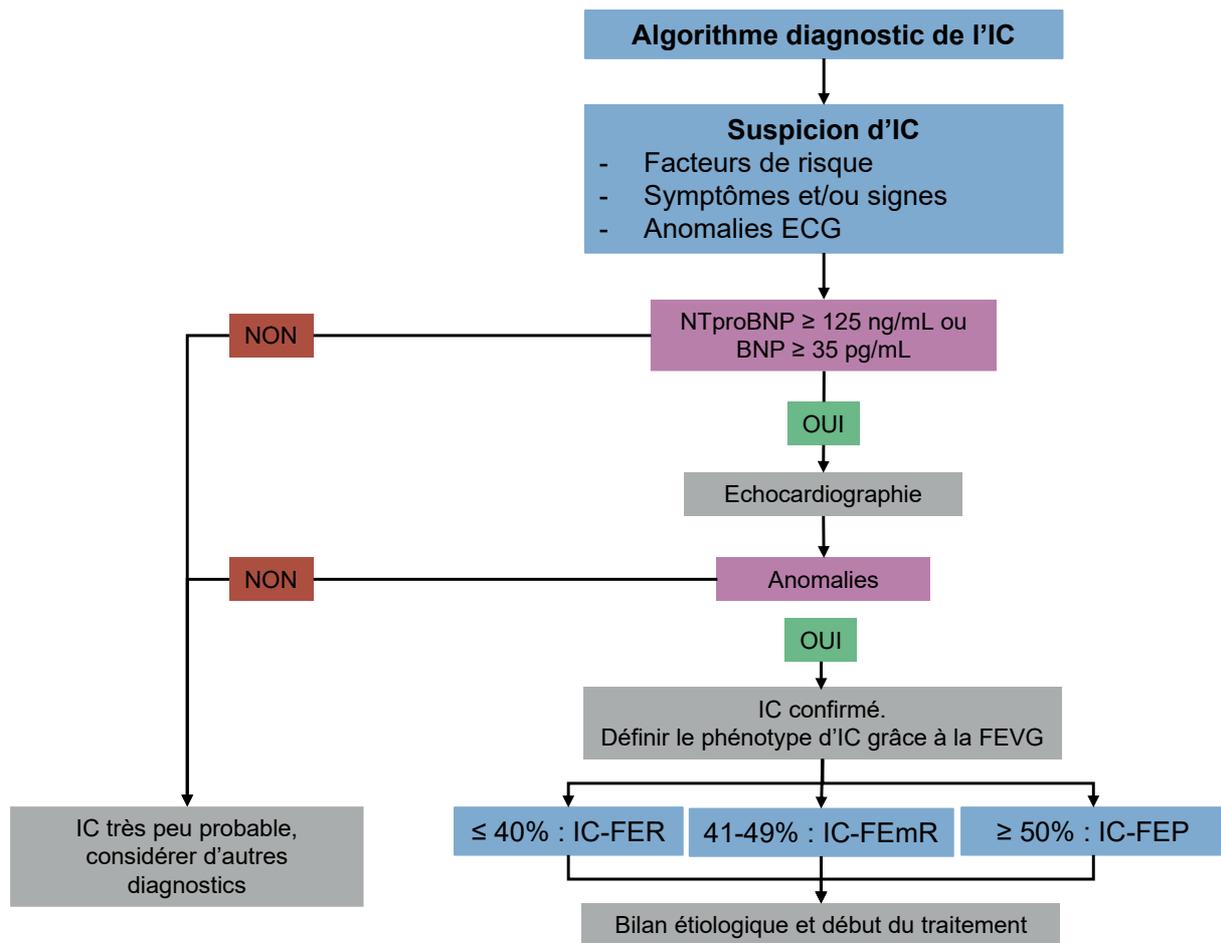
Ces signes cliniques peuvent évoluer à terme vers un état de **choc cardiogénique** qui correspond à l'IC dans son état le plus grave (sans considérer l'arrêt cardiaque qui par définition est un arrêt du débit cardiaque par défaillance électromécanique de la pompe cardiaque) avec une hypotension artérielle, des signes d'hypoperfusion périphériques (extrémités froides, temps de recoloration cutanée >3 secondes, marbrures, altération de la vigilance, douleurs abdominales...) et des signes d'IC droites et/ou gauches.

## 1.6 Diagnostic paraclinique

Actuellement, les examens biologiques recommandés par l'ESC pour le diagnostic de l'IC sont le **BNP** et le **NT-ProBNP**.<sup>(7)</sup>

Il s'agit de deux biomarqueurs synthétisés à partir du Pro-BNP par les ventricules et les oreillettes en réponse à un stress des myocytes, ici l'étirement. Le BNP est connu pour avoir un **effet natriurétique** et **vasodilatateur** mais a, de plus, une **action sur le remodelage** ventriculaire via l'inhibition du **SRA** et de la sécrétion d'**aldostérone**. Le NT-ProBNP qui est la partie N-terminale est la forme inactive.<sup>(12)</sup>

Il faut garder à l'esprit que l'homéostasie de ces deux hormones dépend de plusieurs facteurs : elles augmentent avec l'âge, ont des taux plus élevés chez la femme, leurs concentrations sont plus faibles chez l'obèse et s'accumulent en cas d'insuffisance rénale.



**Figure 2 :** Arbre diagnostique de l'IC chronique (traduit des recommandations ESC 2021)

*BNP : Brain natriuretic peptide ; FEVG : fraction d'éjection ventriculaire gauche ; IC : Insuffisance cardiaque ; IC-FeR : Insuffisance cardiaque à FEVG réduite ; IC-FEmR : Insuffisance cardiaque à FEVG modérément altérée ; IC-FEP : Insuffisance cardiaque à FEVG préservée ; NT-proBNP : N-terminal pro-B natriuretic peptide*

Comme tout marqueur, le BNP et le NT-ProBNP ont leurs **limites** :

- Un patient admis aux urgences pour une dyspnée aiguë fébrile sans signe droit ni gauche avec un dosage de BNP élevé ne doit pas faire conclure à tort à un IC aiguë. La valeur du marqueur **dépend de la probabilité pré-test**.
- Le marqueur peut être **pris à défaut en cas d'obésité** avec un faux négatif ou en cas d'**insuffisance rénale** avec un faux positif.

Enfin, l'**échocardiographie transthoracique** occupe une place centrale dans l'algorithme de diagnostic de l'IC (**Figure 2**) et son étiologie comme il a été mentionné dans la partie précédente.

## 1.7 Traitements

En 2023, l'ESC a mis à jour les dernières recommandations sur le traitement de l'IC aiguë et chronique.(7)

### 1.7.1 Généralités

Tout d'abord il convient de revenir sur plusieurs points :

- Les traitements de l'IC aiguë ont pour but de restaurer une hémodynamique efficace et le retour à l'euvolémie.
- Les traitements de l'IC chronique ciblent le remodelage cardiaque et la prévention de nouvelle décompensation.
- Les diurétiques de l'anse sont toujours recommandés quelle que soit la FEVG en cas de symptôme congestif et un thiazidique peut être ajouté en cas de résistance.
- Un distinguo peut être fait entre les traitements médicamenteux et les traitements par resynchronisation ou assistances de courte/longue durée ou encore la transplantation cardiaque.
- Les traitements non médicamenteux tels que la réadaptation cardiaque et la prise en charge globale des facteurs de risque notamment en cas de FEVG préservées ne doivent pas être oubliés tout comme la correction de la carence martiale définie comme ferritinémie inférieure à 100 ou 300µg/L en présence d'un coefficient de saturation inférieur à 20%.

Enfin, seuls les traitements de l'IC à FEVG altérée ont montré un gain significatif sur la morbi-mortalité.(13) Cependant, ces traitements sont étendus aux deux autres types d'IC en raison de leur analogie de profil des patients à FEVG modérément réduite, mais avec un grade de recommandation plus faible IIb. A noter que l'**empagliflozine** en 2021 a montré un bénéfice significatif chez les patients insuffisants cardiaques à FEVG préservées (EMPEROR-PRESERVED).(14) Plus récemment, l'étude DELIVER présentée à l'ESC de 2022, a montré que l'ajout de la **dapagliflozine** au traitement standard de l'IC entraînait une diminution de la

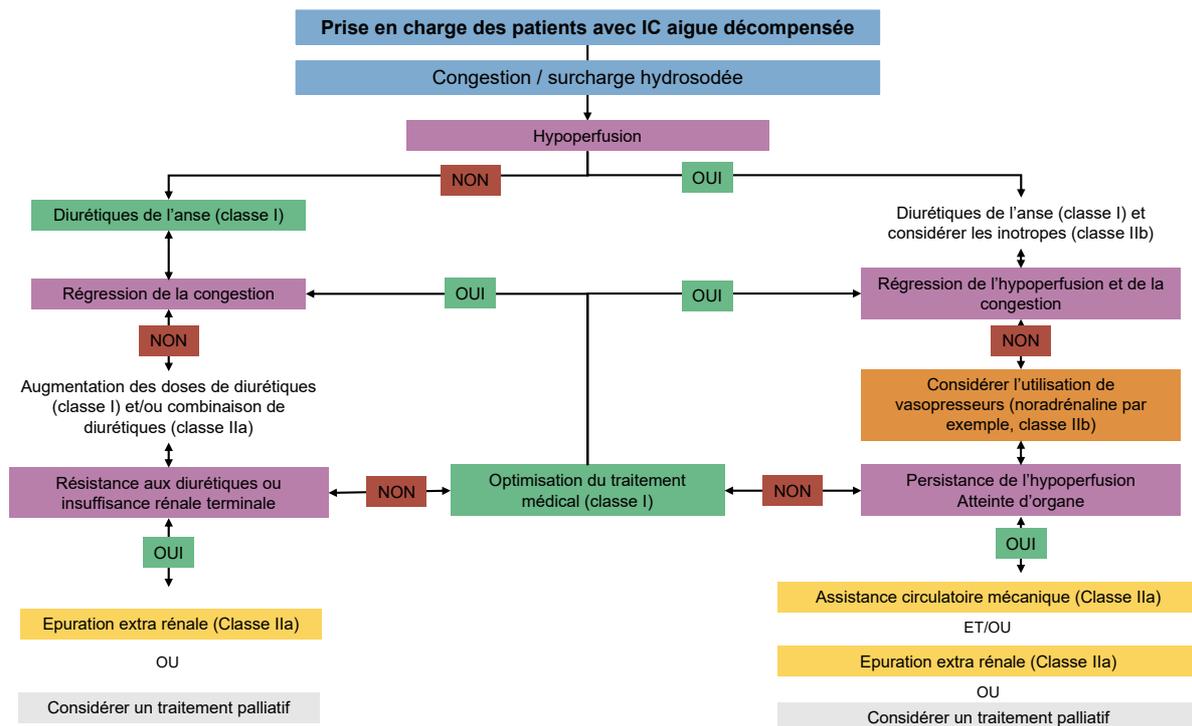
mortalité cardiovasculaire et des hospitalisations pour IC chez des patients à FEVG préservée et modérément réduite. (15)

### 1.7.2 Insuffisance cardiaque aigüe

La prise en charge de l'IC aigüe repose évidemment sur la **prise en charge étiologique** qui ne sera pas détaillée ici, et sur la **prise en charge symptomatique** dont les spécificités sont rappelées dans la **Figure 3**.

Les trois questions principales devant tout patient avec IC sont :

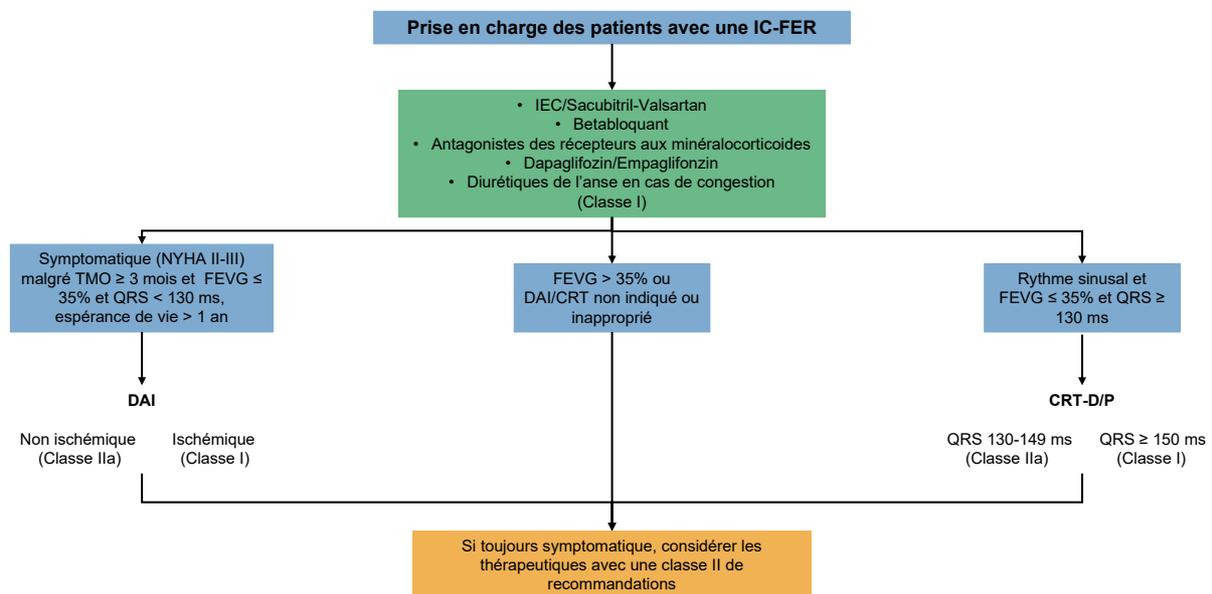
- la présence de signes de **surcharge hydrosodée** : diurétiques de l'anse plus ou moins associés à un thiazidique
- la présence de signes d'**hypoperfusion périphérique** : inotrope dobutamine, chronotrope isoprénaline, support vasopresseur par noradrénaline
- la nécessité d'une **assistance d'organes** : assistances cardiaque, respiratoire et/ou rénale.



**Figure 3** : Prise en charge des patients en insuffisance cardiaque aigüe (traduit des recommandations ESC 2021)

### 1.7.3 Insuffisance cardiaque chronique

La principale nouveauté apportée aux précédentes recommandations (**Figure 4**) est qu'il convient d'introduire de manière plus précoce, pour ne pas dire d'emblée, le bêta-bloquant à faible dose, l'inhibiteur de l'enzyme de conversion (IEC) ou l'antagoniste des récepteurs à l'angiotensine (ARA2) ou les inhibiteurs de la néprysiline, les antagonistes aux récepteurs aux minéralocorticoïdes, et les inhibiteurs du cotransporteur du sodium /glucose II (gliflozines). Il relève donc du clinicien de décider de commencer d'emblée par le SACUBITRIL-VALSARTAN ou l'IEC/ARA2. Les médicaments s'introduisent à faible dose et se titrent progressivement jusqu'à dose maximale tolérée.



**Figure 4 :** Prise en charge de l'insuffisance cardiaque chronique à FEVG réduite (traduit des recommandations ESC 2021)

*IC-FER : Insuffisance cardiaque à FEVG réduite ; IEC : inhibiteurs de l'enzyme de conversion ; NYHA : New York Heart Association ; TMO : traitement médical optimal ; FEVG : fraction d'éjection du ventricule gauche ; DAI : défibrillateurs automatique interne ; CRT : traitement par resynchronisation cardiaque.*

En cas de persistance à trois mois des symptômes (dyspnée NYHA II) et/ou de la persistance d'une FEVG < 35% sous traitement optimal, l'**ivabradine** peut être considérée si le patient est

en rythme sinusal avec une fréquence cardiaque supérieure à 70 bpm ; la **digoxine** ou l'ablation percutanée est indiquée en cas de FA.

La **resynchronisation** doit se proposer en cas de bloc de branche gauche associée >150 ms (grade IA) ou >130ms (grade IIa). En l'absence de bloc de branche mais avec une espérance de vie estimée à plus d'un an, un **défibrillateur automatique implantable** est recommandé en grade IA en cas de cardiopathie ischémique et grade IIa en cas de cardiopathie non ischémique.

L'algorithme de la gestion des thérapies en cas d'IC avancée ne sera pas détaillé dans ce travail. Il sera juste précisé que le traitement de référence reste la transplantation cardiaque et pour ceux ne pouvant en bénéficier, les assistances ventriculaires type LVAD (*left ventricular assist device*) dont le bénéfice par rapport au traitement médicamenteux a été montré notamment par l'étude ROADMAP.(16)

## 2 : Facteurs environnementaux issus de la cohorte ADDICTO-USIC

---

### 2.1 Positionnement du problème

Alors que la consommation de substances récréatives ou illicites est en constante progression, il existe peu de données sur l'impact clinique de la consommation de telles substances sur la sévérité clinique initiale chez un patient hospitalisé pour IC aiguë.

La cohorte ADDICTO-USIC (ou *ADDICT-ICCU study*) est la **première étude prospective multicentrique** évaluant l'impact pronostique de ces substances récréatives lors d'une décompensation cardiaque aiguë.

### 2.2 Description de la base de données

Le projet ADDICTO-USIC est à l'initiative d'une collaboration conjointe entre le Groupe des Unités de Soins Intensifs de Cardiologie (USIC) et le Collège des Cardiologues en Formation (CCF) de la Société Française de Cardiologie. Il s'agit d'une étude prospective multicentrique incluant 39 centres à travers la France.(17)

L'objectif de cette étude était d'évaluer l'association entre la consommation de substances récréatives et la survenue d'événements majeurs (MAE) intra-hospitaliers incluant la mortalité toute cause, le choc cardiogénique nécessitant un support par amines ou assistance mécanique, et un arrêt cardiaque récupéré, chez tous les patients consécutifs hospitalisés en USIC.

Le recrutement a duré 15 jours au mois d'**avril 2021** avec un recrutement final de **1576 patients**.

#### 2.2.1 Méthode de détection des drogues récréatives

Il s'agit de détecter de manière systématique chez tout patient admis aux soins intensifs de cardiologie la **présence de substances récréatives** dans les urines comprenant : tétrahydrocannabinol ou THC (cannabis, marijuana, haschisch), cocaïne/crack, héroïne, morphine, opiacés, amphetamine, ecstasy. La fiabilité du test (NarcoCheck<sup>®</sup>, Kappa City

Biotech SAS, MontlEtats-UnisFrance) est excellente puisqu'elle est validée à 97-99% et les résultats se lisent en cinq minutes (**Figure 5**).



**Figure 5** : Kit de dosage urinaire utilisé pour l'étude ADDICTO-USIC (NarcoCheck®, Kappa City Biotech SAS, Montluçon, France)

De plus, afin d'évaluer la fiabilité du test NarcoCheck® une analyse comparative entre NarcoCheck® et les résultats du laboratoire régional de référence en toxicologie biologique a été réalisée sur un échantillon aléatoire de 60 patients de l'étude. Ainsi, le test urinaire de dépistage de drogues NarcoCheck® avait une sensibilité de 91,7% et une spécificité de 97,9% (résultats présentés dans le papier princeps de l'étude ADDICT-ICCU).(53).

### 2.2.2 Méthode de détection du CO expiré

La seconde particularité de ce projet est de mesurer le **monoxyde de carbone (CO)** à l'admission des patients. La méthode de dosage utilisée dans notre registre CO-Check Pro device (Bedfont Scientific Ltd, Kent, UK).

Pour rappel, le CO est un gaz naturellement présent dans l'atmosphère. Ses propriétés sont d'être incolore, inodore et non irritant lors de son inhalation.

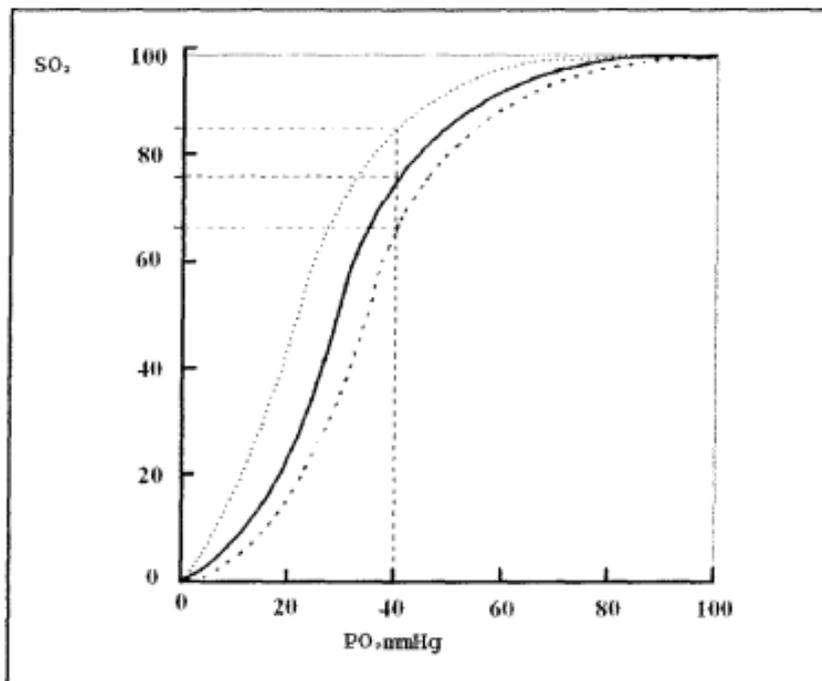
Il est issu de la combustion incomplète et ses principales sources d'émission sont le trafic automobile, la fumée de tabac et les appareils ménagers défectueux. Son taux dans l'atmosphère se mesure en partie par million de volume, ou ppm, 1 ppm correspondant à 0,873 mg/m<sup>3</sup>.

Sa mesure dans la circulation sanguine se traduit par le taux de Carboxyhémoglobine (HbCO). En effet, après avoir été inhalé, 80% du monoxyde de carbone se lie à l'hémoglobine de manière stable mais réversible jusqu'à atteindre un plateau en 8 à 12h. La carboxyhémoglobine

s'élimine avec un temps de demi-vie de quatre heures ce qui signifie qu'elle est éliminée au bout de 24h.

On parle d'intoxication aiguë au CO à partir d'un taux de 5% chez le non-fumeur et de 10% chez le fumeur. Il est décrit que les concentrations peuvent atteindre 15 à 20% après avoir fumé une cigarette.(18)

Le retentissement le plus connu est celui qui s'apparente à l'effet Bohr : l'affinité de l'hémoglobine pour le monoxyde de carbone est deux cents fois supérieure à celle de l'oxygène (Figure 6). Il en résulte alors une hypoxie. Le traitement est l'oxygénothérapie pure à haut débit et le caisson hyperbare.



**Figure 6** : courbe de dissociation de l'hémoglobine en fonction de son affinité.

*SO<sub>2</sub> : saturation en oxygène ; PO<sub>2</sub> : pression partielle en oxygène*

## 2.3 Amphétamines et insuffisance cardiaque

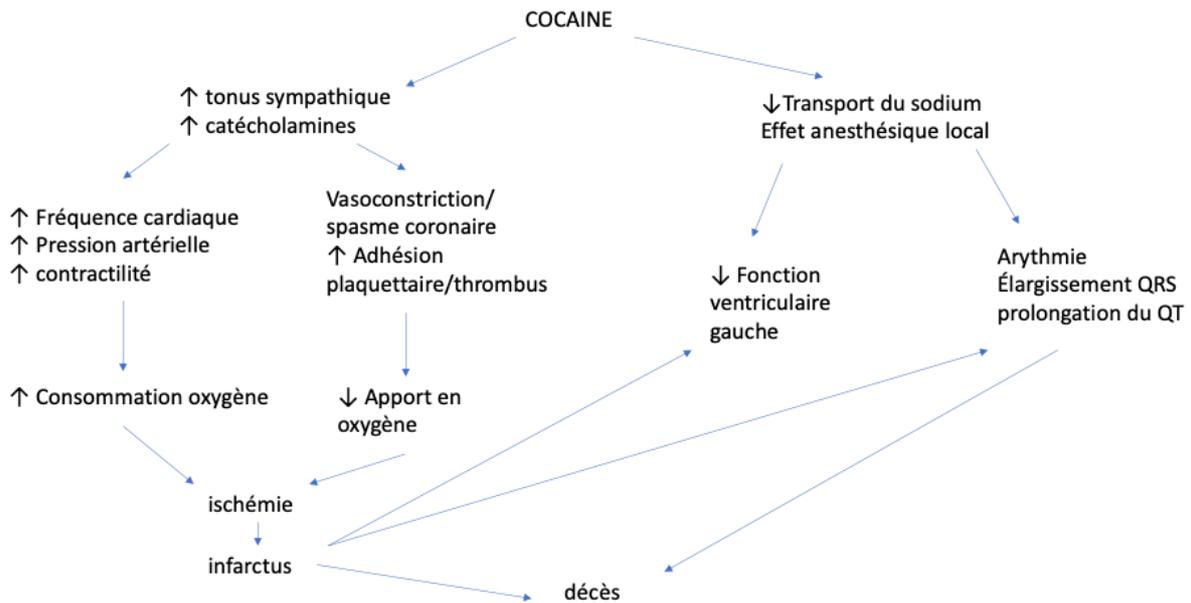
L'usage de **méthamphétamines** est estimé à 27 millions de personnes dans le monde et représente la deuxième cause de morbidité aux États-Unis après les opiacés.(19) Certaines études suggèrent qu'il existe un lien entre l'usage de cette drogue et la survenue d'IC et d'hypertension artérielle pulmonaire.(20) Les mécanismes mis en causes sont multifactoriels et intéressent en premier lieu le **système adrénergique** : les catécholamines entraînent une tachycardie, une hypertension artérielle et un vasospasme coronaire. Le second mécanisme fait intervenir les radicaux libres qui entraîne une lyse des cardiomyocytes et altèrent l'ADN mitochondrial.(21) Enfin, la prise de méthamphétamine associée à la cocaïne augmenterait le risque d'infarctus d'origine athéromateuse par rupture de plaque ou à coronaires saines.

L'étude suivante s'est intéressée à **l'impact de la prise d'amphétamine chez le patient insuffisant cardiaque**.(22) Il s'agit d'une étude rétrospective sur une cohorte de 978 217 patients insuffisants cardiaques, qui a montré que la consommation de stupéfiants dont les méthamphétamines entraînait un taux de réhospitalisation pour IC aiguë à 30 jours supérieur avec un *hazard ratio* ajusté à 1.16 [1.12-1.21]; ( $P < 0.01$ ) comparé à des patients non consommateurs.

## 2.4 Cocaïne et insuffisance cardiaque

La **cocaïne** à faible dose est un puissant stimulant du système nerveux **en inhibant la recapture des catécholamines** au niveau de la synapse et en **sensibilisant les récepteurs aux catécholamines**. A forte dose, elle a une action antalgique en bloquant les canaux sodiques et potassiques ce qui équivaut au niveau du myocyte à une action antiarythmique de type I.(23) A cela s'ajoutent des **effets vasoconstricteurs** en relarguant de l'endothéline et en inhibant la production de monoxyde d'azote, ainsi qu'une activation plaquettaire et des facteurs de la coagulation (**Figure 7**).(22)

La **consommation de cocaïne** entraîne donc une augmentation de la **fréquence cardiaque**, une augmentation de la **pression artérielle**, une diminution du calibre et de la **perfusion coronaire**, la survenue d'**arythmie** et l'allongement de **l'intervalle QT**. Les conséquences cliniques sont multiples : thrombose coronaire, dissection aortique, arythmies, IC, mort subite et endocardite.



**Figure 7 :** Figure adaptée d'une figure tirée de l'étude : effets aigus de la cocaïne.(22)

*La cocaïne affecte le système cardiovasculaire par 2 voies principales : la stimulation sympathique et le transport du sodium. En augmentant le tonus sympathique et la libération de catécholamine, la cocaïne augmente la fréquence cardiaque, la tension artérielle et la contractilité myocardique, ce qui augmente la demande d'oxygène myocardique. L'apport d'oxygène myocardique est réduit par une vasoconstriction coronarienne et une thrombogénicité accrue. La demande en oxygène du myocarde peut dépasser l'apport en oxygène du myocarde, ce qui entraîne une ischémie ou un infarctus. La cocaïne affecte directement les myocytes cardiaques en bloquant les canaux sodiques, ce qui diminue la contractilité ventriculaire gauche et est pro-arythmogène.*

L'impact pronostique de la cocaïne a pu être évalué à travers une étude rétrospective parue en 2022 qui s'était intéressée à une population de 738 patients insuffisants cardiaques consommant de la cocaïne entre 2001 et 2019. Le critère de jugement principal était la **mortalité toute cause** et le **taux d'hospitalisation non planifiée toute cause** à 90 jours. L'usage de la cocaïne augmentait la mortalité avec un *hazard ratio* ajusté à 1.21 [1.00-1.48] et les réadmissions

hospitalières avec un *hazard ratio* ajusté 1.49[1.20-1.85] pour les hospitalisations toute cause, et un *hazard ratio* ajusté 1.49 [1.10-2.01] pour celles pour IC.(21)

## 2.5 Cannabis et insuffisance cardiaque

Le cannabis contient en proportion variable deux substances : le **tétrahydrocannabinol (THC)** et le **cannabidiol (CBD)**. Le THC se fixe sur les récepteurs endogènes CB1 et CB2 dont le tropisme est respectivement le système nerveux et les cellules de l'immunité ; leur activation emprunte la voie qui active l'AMPcyclique. La pharmacodynamie et pharmacocinétique dépend notamment du mode d'absorption et la quantité utilisée.(24)

Sur le plan cardiovasculaire, le THC, en stimulant le **système sympathique**, augmente la fréquence cardiaque, le travail respiratoire du cœur et favorise l'activation plaquettaire. De plus, le THC augmente aussi la carboxyhémoglobine jusqu'à trois fois, ce qui a un impact sur la **dysfonction endothéliale** et **diminue la fixation de l'oxygène à l'hémoglobine** par effet Bohr. Le **CBD** quant à lui stimule le système nerveux parasympathique et diminue la fréquence cardiaque, la pression artérielle, favorise la fonction endothéliale et diminue l'inflammation.(24)

L'association de la consommation de cannabis avec la survenue d'événement cardiovasculaire n'est pas encore démontrée. L'étude prospective CARDIA(25) s'est intéressée à la survenue d'évènements cardiovasculaires chez une population jeune de 18 à 30 ans dont 84% consommaient du cannabis ; et elle n'a pas montré d'association avec la survenue de cardiopathie. Une autre étude menée en 2019 par Chami et Kim(26) sur près de 300 000 patients a montré que l'incidence cumulée sur trois ans d'infarctus du myocarde était plus élevée en cas de consommation de cannabis (RR 2,53 [2,45-2.61]. Le cannabis est aussi décrit comme un facteur de risque de survenue de **FA** et **d'accident vasculaire cérébral (AVC)**.(27)

Une autre étude rétrospective sur une cohorte de 978 217 patients insuffisants cardiaque a montré que la consommation de stupéfiant, dont **le cannabis**, **entraînait un taux supérieur de réhospitalisation pour IC aiguë** à 30 jours avec un *hazard ratio* ajusté à 1.16 [1.12-1.21] ( $P < 0.01$ ) comparé à des patients non consommateurs.(21)

## 2.6 Monoxyde de carbone et insuffisance cardiaque

Au-delà de la consommation de substances récréatives, l'étude ADDICTO-USIC a également évalué **l'impact clinique du taux de CO expiré** chez les patients hospitalisés pour IC aiguë. En effet, bien qu'il existe quelques données montrant un lien possible entre exposition au CO et survenue d'IC(28)(29), il existe **très peu de données** évaluant la question de la sévérité et du pronostic. Il est observé qu'en cas de pic de pollution, **la mortalité et les hospitalisations pour IC aiguë augmentent**. Cette constatation a fait l'objet d'une méta-analyse parue dans le *Lancet* en 2013.(28)

Sur le plan fondamental, une étude animale(30) a été réalisée sur des chiens et des rats insuffisants cardiaques (ligature de la coronaire gauche) exposés pendant quatre semaines consécutives au monoxyde carbone avec concentrations comparables à celles des pics de pollution. Les taux sanguins de carboxyhémoglobine étaient autour de 6%. Les résultats ont montré une **aggravation de la fonction systolique cardiaque** et la survenue de **troubles du rythme ventriculaire**.

Une étude de cas(31) montre un authentique **état de choc cardiogénique** réversible chez un jeune adulte sans antécédent à l'issue d'une intoxication au monoxyde de carbone (taux de HbCO 29,5%). L'atteinte de la fonction systolique était sévère avec une FEVG à moins de 25%. La troponine I était élevée à 124 ng/l (<29 ng/l), tandis que les autres biomarqueurs étaient normaux. Après une semaine de prise en charge, la fonction systolique était évaluée à plus de 55%.

## 3 : Intelligence artificielle, machine learning supervisé

---

### 3.1 Définition

L'intelligence artificielle ou IA, est définie dans le dictionnaire Larousse comme un « ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine ». Son intérêt réside dans **la capacité à pouvoir traiter un grand nombre de données en un temps réduit** à l'heure où il existe une multitude de données, le *Big Data*.

Le *machine learning* (ML) ou « apprentissage automatique » est une sous-catégorie de l'intelligence artificielle, permettant à l'aide d'algorithmes de mettre en évidence des relations complexes parmi un grand jeu de variables. La notion de statistique et de probabilité tout comme d'auto-apprentissage sont au cœur du ML. Il s'agit donc d'un mécanisme toujours en mouvement : **les données de sortie sont mises à jour au fur et à mesure que de nouvelles données sont disponibles.**

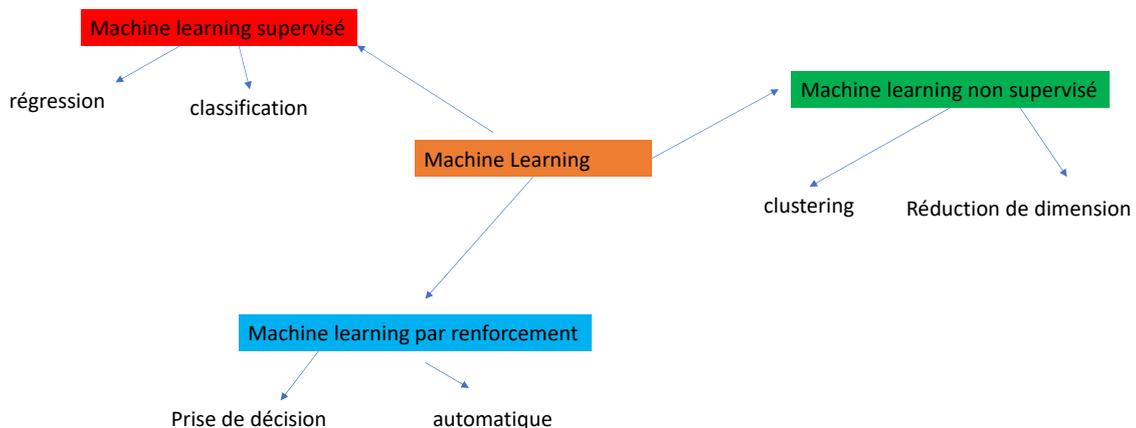
On distingue classiquement trois types de ML (**Figure 8**) :

le ML **supervisé**

le ML **non supervisé**

le ML **par renforcement**

Seul le ML supervisé sera développé par la suite dans le cadre de ce travail



**Figure 8** : Schéma synthétiques des trois principales formes de ML

## 3.2 Principes généraux

Le ML supervisé s'entraîne sur des **données déjà étiquetées** par le programmeur et ce dernier ajuste les résultats pour que la machine comprenne comment s'améliorer d'où le terme de « supervisé ». Une fois entraîné, l'algorithme peut être utilisé sur des nouvelles données dont il en déterminera la sortie.

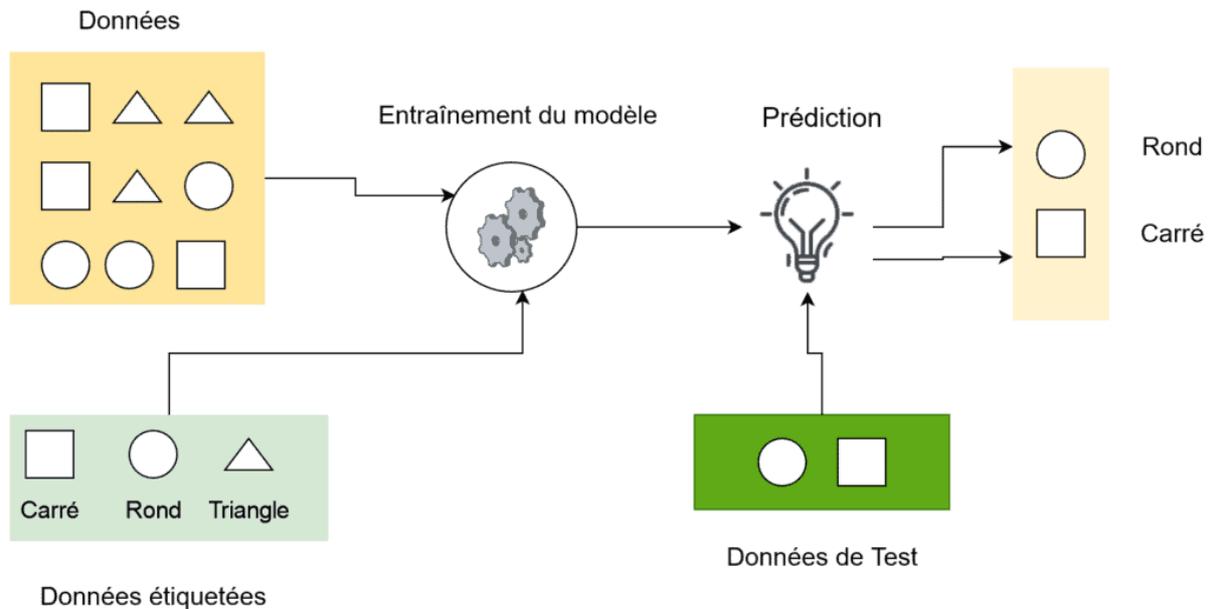


Figure 9 : Fonctionnement et étapes du ML supervisé (32)

La première étape est **la constitution du problème** que le ML doit résoudre à partir des données étiquetées qui forment le « *dataset* ». Cet ensemble de données regroupe d'un côté la variable binaire que l'on souhaite prédire la « *target variable (y)* » et de l'autre côté les « *features (x)* » ou variables explicatives.

Selon la variable étudiée, la tâche de l'algorithme consistera soit en **une classification** pour les variables discrètes soit **une régression** en cas de variable continue pour décrire  $y=f(x)$ .

La seconde étape est la constitution du modèle dont il faut **optimiser les paramètres** au sein d'une population d'entraînement ou *train-set*.

**Dans le ML supervisé**, c'est l'investigateur qui choisit le modèle et c'est la machine qui va déterminer les paramètres afin d'être le plus performant.

Par exemple, si le nuage de points dessiné sur le graphique  $y=f(x)$  s'apparente à une droite, le modèle proposé sera une fonction affine de type  $y=ax+b$  où  $a$  et  $b$  seront les paramètres que la machine devra trouver.

La troisième étape est **la gestion des erreurs**. Plus le modèle est performant moins les erreurs de prédiction seront importantes. Cette étape capitale est au cœur du ML : la machine va apprendre par elle-même à minimiser ses erreurs. On parle alors d'**algorithmes de minimisation de la fonction coût**.

La dernière étape est **l'évaluation du modèle** dans une population différente de la population d'entraînement, appelée *test-set*. Cette étape est cruciale car elle permet de **valider le modèle** et de vérifier qu'il ne soit pas orienté par les seules caractéristiques de la population du *train-set*.

Enfin, bien qu'elle ne soit pas obligatoire, une étape additionnelle peut consister à appliquer le modèle final ML dans **une cohorte indépendante dite de validation externe**.

### 3.3 Fonctionnement

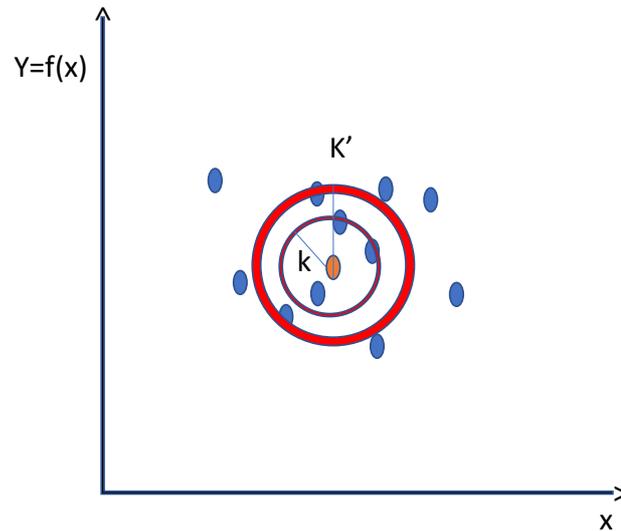
#### 3.3.1 Dataset : gestion des données manquantes

Afin de permettre un traitement performant des données issues du *Dataset*, il convient d'avoir un recueil exhaustif de ces dernières. Dans le cas où il manquerait des données, la question de la **stratégie de gestion des données manquantes** s'impose afin de contrebalancer la perte de puissance engendrée.

Les techniques d'imputation sont multiples :

- attribuer la donnée manquante des patients à la **médiane** ou la **moyenne**.
- attribuer la **dernière valeur connue**.
- **l'imputation multiple** par équations de chaîne : on attribue à la valeur manquante, **la valeur la plus probable** selon les caractéristiques du patient.

- la méthode **KNN des k plus proches voisins** qui consiste à attribuer à la valeur manquante la moyenne des valeurs d'un groupe de k patients les plus proches (**Figure 10**).



**Figure 10** : méthode des k plus proches voisins

Valeur connue ● valeur à déterminer ●

$K$  : nombre des plus proches voisins dans le cercle violet : ici  $k = 3$

$K'$  : nombre des plus proches voisins dans le cercle rouge : ici  $K' = 5$

### 3.3.2 Gestion des données déséquilibrées

Les **données déséquilibrées** ou *unbalanced data* sont des situations fréquentes d'autant plus que les événements sont rares. On parle de données déséquilibrées lorsque le rapport entre deux d'entre elles est différent de 50%. On parle de **classe majoritaire** pour celle la plus représentée et de **classe minoritaire** pour celle la moins représentée. Le problème posé en cas d'*unbalanced data* sera une **forte disparité** entre les classes à prédire ce qui se traduira par une **chute de la performance** et une augmentation de la variance ; l'échantillon constituant la classe minoritaire étant plus difficile à détecter par l'algorithme du ML.

Les techniques permettant de rééquilibrer l'échantillon sont le **sur-échantillonnage** ou *oversampling* et le **sous-échantillonnage** ou *undersampling*.

L'*undersampling* s'attache à diminuer la classe majoritaire afin de donner plus d'importance à la classe minoritaire. Inversement, l'*oversampling* va dupliquer les données issues de la classe minoritaire. Une des techniques utilisées est le **SMOTE** (*Synthetic Minority Oversampling Technique*) qui ne génère non pas des individus de manière strictement aléatoire mais des **individus similaires**. Pour se faire, la méthode des plus proches voisins est utilisée et il convient de définir le nombre d'individus synthétiques souhaité.

### 3.3.3 Sélection de variables

Plusieurs modèles peuvent être utilisés :

- a) Analyses uni- et multivariée

La **régression logistique** est un modèle statistique qui permet de prédire l'association entre plusieurs variables. Si on cherche un lien statistique entre deux variables, on parle de **régression univariée**. Si on prend en compte plusieurs variables, on parle de **régression multivariée**. L'intérêt du modèle multivarié est qu'il permet d'**ajuster la mesure sur les autres covariables** incluses du modèle et de s'affranchir des potentiels facteurs de confusion.

- b) Méthodes par arbres de décision : *random forest* et *XGBOOST*

Les **forêts aléatoires** ou *random forest* correspondent à une méthode basée sur **des arbres de décision** où chaque branche se sépare selon la réponse à un critère donné ou une question posée parmi les *features*. L'algorithme va donc déterminer les meilleurs *features* pour aboutir au résultat final (la *target variable*) selon le gain d'information obtenu.

Le nombre d'arbres est fixé par l'investigateur et chaque arbre est **entraîné sur un échantillon aléatoire issu du dataset**. L'ensemble des résultats est alors combiné pour former l'algorithme final.

*XGBOOST* (*extreme gradient boosting*) repose sur l'idée que **l'utilisation de plusieurs algorithmes décisionnaires même faibles peuvent être plus performants qu'un seul**. Chaque arbre de décision est constitué de plusieurs branches successives qui ajustent successivement l'erreur de la précédente branche par rapport à la valeur réelle. Cette différence est appelée « résidu ». A ce mécanisme, s'ajoute celui de **l'élagage** : les branches qui ne sont

pas suffisamment bonnes sont élaguées et les arbres les moins bons sont supprimés. Ainsi le *XGBOOST* permet de ne **garder que les algorithmes les plus performants**.

c) Méthode du LASSO (Least Absolute Shrinkage and Selection Operator)

La méthode du **LASSO** est une régression pénalisée s'apparentant à une régression logistique dont les coefficients à déterminer sont pondérés par une fonction coût. Ainsi, les variables ayant une valeur proche de la valeur nulle sont supprimées et seules les variables les plus pertinentes sont sélectionnées.

### 3.3.4 Évaluation d'un modèle pronostique

a) Matrice de confusion

Dans le cadre d'un **modèle pronostique**, la matrice de confusion est un outil pour évaluer la performance du modèle où sont rapportées les **prédictions par rapport aux valeurs réelles**. De ce tableau sont tirés notamment les vrais positifs/négatifs et les faux positifs/négatifs. Elle fournit donc une compréhension détaillée des performances du modèle et permet d'identifier les types d'erreurs que le modèle commet.

Sont alors calculées la **sensibilité** qui mesure la capacité du modèle à identifier correctement les vrais positifs, la **spécificité** qui est relative aux vrais négatifs, la **précision** qui mesure la proportion de vrais positifs parmi les individus prédits comme positifs (vrais et faux positifs) et enfin l'**exactitude** qui mesure la proportion totale de prédictions correctes par le modèle.

b) Courbes ROC et ROC-AUC

Dans le cadre d'un **évènement à prédire binaire**, la courbe ROC est une représentation graphique de la sensibilité (taux de vrais positifs) par rapport à 1-spécificité (taux de faux positifs) pour différents seuils de classification. Ainsi sa valeur va être notamment déterminée par le taux de non-évènement qui représente souvent la classe majoritaire en médecine.

En plus d'être connue dans la littérature scientifique, elle offre l'avantage d'une **interprétation visuelle immédiate** : plus la courbe se rapproche du coin supérieur gauche du graphique, plus le modèle est performant.

L'aire sous la courbe ROC, ou **ROC-AUC**, fournit une mesure quantitative de la capacité du modèle à classifier les événements. Plus la ROC-AUC est proche de 1 plus la performance est proche de la perfection. En revanche, une valeur proche de 0,5 indique une performance comparable au lancer d'une pièce de monnaie.

#### c) Courbe precision-recall et PR-AUC

Différemment de la courbe ROC qui se base sur la sensibilité et la spécificité, la courbe PR est une **représentation graphique de la précision** (vrais positifs parmi les prédictions positives) **par rapport au rappel** (vrais positifs parmi les vrais positifs et les faux négatifs). Elle est particulièrement utile lorsque les **vrais positifs à prédire est en classe minoritaire**. On dit alors que les données sont déséquilibrées.

Son interprétation est d'une part graphique : une courbe PR idéale qui maximise la précision pour chaque niveau de rappel, forme ainsi une courbe qui atteint le **coin supérieur droit du graphique**. Son interprétation numérique est plus fine et dépend notamment de la prévalence de l'évènement à prédire dans le jeu de données. Plus le chiffre de PR-AUC est élevé, plus les performances du modèle sont bonnes.

#### d) F1-score

Le F1-score est une métrique pour les modèles prédictifs qui intègre à la fois la **précision** et le **rappel**, ce qui en fait une mesure précieuse en cas de données déséquilibrées.

Sa formule mathématique est la suivante :

$$F1 - score = \frac{2 \times (precision + rappel)}{precision + rappel}$$

Plus le F1-score est proche de 1, plus le modèle est performant pour discriminer les événements positifs.

## 3.4 Applications en cardiologie et scores existants

### 3.4.1 Méthodologie des scores

Comme il a été rappelé dans la première partie de ce travail, l'IC est un vrai défi diagnostique et thérapeutique pour la prise en charge de nos patients, car repose sur l'évaluation de dizaines de variables cliniques, biologiques et échocardiographiques. Ainsi, l'utilisation de scores pronostiques permettant d'intégrer de façon optimisée ces différents paramètres pronostiques pourraient constituer un outil décisif dans l'évaluation précoce des patients pris en charge pour IC aiguë.(33)

L'article publié par l'équipe de Christian Mueller en 2020 dans la revue Intelligence médicale,(34) recense quelques grands principes pour l'élaboration d'un score pronostique robuste :

- Obtenir des données fiables dans un registre de qualité si possible multicentrique
- Avoir un effectif adapté au nombre de variables
- Doit être validé dans une population externe d'une autre population
- Le score doit être appliqué dans un contexte particulier : par exemple enUSIC
- Les variables doivent être accessibles facilement rapidement en routine clinique
- Le score doit être comparé à un score validé existant
- Le score doit être appliqué dans des études pragmatiques
- Le score doit être incorporé avec une dimension économique
- Le score doit pouvoir s'adapter aux avancés thérapeutiques

### 3.4.2 Intérêt du score en pratique clinique

**L'utilisation en pratique clinique d'un score validé ne peut trouver d'intérêt que s'il permet de guider la stratégie thérapeutique pour améliorer le pronostic des patients.**

Pour se faire, l'étude interventionnelle MTCP(35) (*multidisciplinary team-based care process*) a évalué l'impact d'une prise en charge thérapeutique selon la gravité d'un score pronostique portant sur le taux de réhospitalisation et de mortalité toute cause à 30 jours chez 6182 patients IC . Le risque de chaque patient était calculé dans les 24 heures suivant

l'admission. Le score pour évaluer la mortalité était le IMRS (*Intermountain Risk Score*)(36) et le score pour évaluer le risque de réadmission à l'hôpital était l'IHF (*inpatient heart failure*)(37) qui intègre cinq variables : **l'utilisation de diurétique, un taux de BNP>200pg/mL, une FEVG<40%, l'éligibilité à l'accès au système de santé et un antécédent d'hospitalisation pour IC.**

Les **patients à haut risque** recevaient **soit le protocole standard** soit le **protocole MTCP** tandis que les patients à faible risque étaient traités selon les recommandations standards. Le protocole MTCP consistait à **une prise en charge standard** associée à **un temps supplémentaire au lit du patient de l'équipe soignante** et à **un suivi plus rapproché à la sortie** de l'hospitalisation.

Le groupe MTCP à haut-risque avait un taux de réhospitalisation à 30 jours inférieur au groupe contrôle à haut risque de 21% (p ajusté=0.013, HR=0.79 (95%CI[0.66- 0.95]), un taux de mortalité à 30 jours inférieur de 52% (p ajusté<0,001, HR 0,48 (95%CI [0,33-0,69])). Les patients à faible risque n'avaient pas plus de réhospitalisation ni de surmortalité (respectivement HR 0,88 p=0,19 et HR 0,88 p=0,61) par rapport au groupe contrôle.

Cette étude illustre tout l'intérêt d'un tel score en éveillant l'attention des soignants sur les **patients à risque et en adaptant le niveau de soins** afin de réduire la mortalité et le taux de réhospitalisation pour IC aiguë des patients les plus graves.

### 3.4.3 Scores traditionnels

Plusieurs études utilisant des méthodes statistiques traditionnelles se sont déjà intéressées à l'établissement de scores pronostiques pour prédire la sévérité de l'IC à court terme. L'ensemble des résultats des études présentées sont synthétisés dans le **Tableau 2**.

	EHRMG	OHFRS	HF-TR	Optimize HF	STRATIFY	MEESSI-AHF	GWTG-HF	ACUTE-HF
Age	X		X	X	X	X	X	X
ethnicité							X	
SAMU	X							
PAS	X		X	X		X	X	
PAD					X			
Fréquence cardiaque	X	X		X				

Fréquence cardiaque après 3 min marche		X						
Fréquence respiratoire					X	X		
Saturation en oxygène	X	X			X	X		
Indice de masse corporelle					X			
NYHA						X		
Hypoperfusion périphérique						X		
SCA						X		
Oxygénothérapie					X			
Index de barthel						X		
ATCD Insuffisance cardiaque								X
ATCD AVC		X						X
ATCD Cancer	X							
Dialyse			X		X			
ATCD BPCO							X	
ATCD intubation		X						
Longueur du séjour			X					
Sous IEC/ARA2					X			
Metolazone	X							
Hémoglobine			X					
Inotrope			X					
Ventilation non invasive								X
Créatinine	X			X		X		X
Kaliémie	X					X		
Natrémie				X	X		X	
Urée		X			X		X	
Capnie artérielle		X						
Troponine	X	X			X	X		
BNP/NTProBNP		X			X	X		
ECG : signes ischémies nouveaux		X						
QRS durée					X			
HVG						X		
FEVG				X				X
IM au moins moyenne								X

**Tableau 2 : Récapitulatif des variables des scores.**

*ARA2 : antagoniste des récepteurs à l'angiotensine 2 ; AVC : accident vasculaire cérébral ; BNP : peptide B natriuretic ; BPCO : bronchopneumopathie chronique obstructive ; FEVG : fraction d'éjection du ventricule gauche ; ECG : électrocardiogramme ; HVG : Hypertrophie ventriculaire gauche ; IEC : inhibiteurs enzyme de conversion ; IM : insuffisance mitrale ; NT-proBNP : N-terminal pro-B natriuretic peptide ; NYHA : New York Heart Association ; PAS : pression artérielle systolique. PAD : Pression artérielle diastolique ; SCA : syndrome coronarien aiguë.*

- Score **EHRMG** (*Emergency Heart Failure Mortality Risk Grade*): étude qui a évalué la mortalité à 7 jours chez 12591 patients (âge moyen 75 ans et 52% d'hommes) admis pour IC aiguë avec un taux de mortalité à 7 jours de 2%.(38) La méthode statistique a utilisé une régression logistique utilisant comme covariables : **l'âge, l'utilisation d'un transport d'urgence, la pression artérielle systémique, la fréquence cardiaque, la SpO2, la créatinémie, la kaliémie, la Troponine, la présence d'un cancer actif et la consommation de Metolazone**. En termes de performance, le score avait une aire sous la courbe (ROC-AUC) de 0,808.
- Score **ACUTE-HF** (*Acute Heart Failure score*) : étude rétrospective qui a évalué la mortalité à 30 jours, 6 mois et 5 ans chez 771 patients (âge moyen 72,3 ans et 65% d'hommes) admis pour insuffisance cardiaque en soins intensifs.(39) La mortalité observée n'était pas communiquée. Après une analyse univariée puis multivariées pour chaque date limite, les auteurs ont déterminé un score qui prédit la mortalité à 6 mois. Les variables du score sont : **âge >76 ans, créatinémie, FEVG≤30%, ATCD d'hospitalisation pour IC, un ATCD d'AVC ou AIT, une insuffisance mitrale au moins moyenne et le recours à la VNI**. Une courbe ROC a été utilisée pour déterminer la mortalité à 30 jours avec une ROC-AUC de 0,78 (0,79 à 6 mois et 0,76 à 5 ans). On peut cependant noter que les variables issues de l'analyse multivariée spécifique de la mortalité à 30 jours étaient : **la présence d'une valvulopathie, un infarctus du myocarde non évalué par coronarographie, un ATCD d'hospitalisation pour IC et la créatinémie**. Seule l'usage de la VNI ressortait dans l'analyse univariée à 30 jours.

- Score **GWTG-HF** (*Get with the Guidelines-Heart Failure*) : étude de cohorte observationnelle menée dans 198 hôpitaux qui a cherché à déterminer la mortalité intrahospitalière chez 39 783 patients (âge moyen 72,5 ans et 49% d'hommes) hospitalisés pour IC. (40) Le taux d'évènements était de 2,86%. L'analyse multivariée a permis de déterminer les variables suivantes : **l'âge, la pression artérielle systolique, l'urée, la fréquence cardiaque, la natrémie, la BPCO et l'ethnicité (« non black race »)**. Le score de risque avait une bonne discrimination avec une C-statistique de 0,75.
- Score **OHFRS** (*Ottawa Heart Failure Score*): étude observationnelle prospective qui évaluait sur six centres canadiens la survenue d'évènements cardiovasculaires chez 559 patients (âge moyen 76 ans et 56% d'hommes) admis aux urgences pour IC aiguë. (41) Le taux d'évènements cardiovasculaires était de 11,6%. L'analyse statistique s'est appuyée sur une régression logistique utilisant comme covariables : **NT-ProBNP>5000ng/L, l'antécédent d'AVC/AIT, ATCD d'intubation, signes ECG d'ischémie aiguë, FC>110bpm, SpO2<90%, Troponine I ou T positive, Urée>12mM, CO2 plasmatique >35mM et une FC >110 bpm après un test de marche de 3 minutes**. En termes de performance, le score avait une aire sous la courbe (ROC-AUC) de 0,77.
- Score **STRATIFY** (*Improving Heart Failure Risk stratification in the Emergency Department*) : étude observationnelle prospective multicentrique portant sur la survenue d'évènements graves à 30 jours chez 1033 patients (âge médian 64 ans et 57% d'hommes) admis pour IC aux urgences.(42) La méthode pour évaluer les variables du score s'est appuyée sur une régression logistique utilisant comme covariables : **l'âge, l'IMC, le BNP, la pression artérielle diastolique, l'urée, la natrémie, la fréquence respiratoire, la SpO2, la Troponine I, la dialyse, durée des QRS et la prescription d'IEC**. En termes de performance, le score avait une C-statistique de 0,68.
- Score **MEESSI-AHF**(*Multiple Estimation of risk based on Emergency Department*) : étude observationnelle prospective qui s'est intéressée à prédire la mortalité à 30 jours chez 4867 patients (âge moyen 80 ans et 57% d'hommes) issus du registre *Acute Heart Failure in Emergency Departments (AHFE)* admis aux urgences pour IC aiguë.(43) La

mortalité à 30 jours était de 10,1%. L'analyse statistique par régression logistique a identifié 13 variables : **l'indice de Barthel, la pression artérielle systolique, l'âge, le taux de NT-ProBNP, la kaliémie, la Troponine, la présence d'une dyspnée NYHA IV, la fréquence respiratoire, les signes de bas-débit, la SpO2, la présence d'un SCA, des signes d'HVG à l'ECG et le taux de créatinémie.** En termes de performance, le score avait une aire sous la courbe (ROC-AUC) de 0,836.

- **Score Journey HF-TR** (*Patient Journey in Hospital with Heart Failure in Turkish Population*) : étude multicentrique observationnelle qui s'est intéressée à prédire la mortalité intra-hospitalière de 1045 patients (âge moyen 75 ans et 51% d'hommes) de plus de 65 ans.(44) La mortalité intra-hospitalière était de 7,4%. L'analyse multivariée a sélectionné 6 variables d'intérêt : **âge>77,5 ans, l'anémie, l'utilisation d'inotrope, l'hémodialyse/ultrafiltration, une pression artérielle non contrôlée et la longueur du séjour>6,5 jours.** En termes de performance, les auteurs ne communiquent pas sur la ROC-AUC du score mais rapportent les ROC-AUC des items du score séparément : âge : 0,62 ; anémie : 0,56 ; administration d'inotrope : 0,62 ; hémodialyse/ultrafiltration : 0,55 ; pression artérielle non contrôlée : 0,49 et la longueur du séjour : 0,58.
- **Score OPTIMIZE-HF** (*Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure*) : étude qui s'est intéressée à appliquer le modèle OPTIMIZE à la mortalité intra hospitalière chez 15219 patients asiatiques (âge moyen 70 ans et 56% d'hommes) admis pour IC aiguë, dont 41% avaient une cardiopathie ischémique.(45) La mortalité intra-hospitalière était de 1,9%. L'analyse multivariée a sélectionné 6 variables : **l'âge, la fréquence cardiaque, la pression artérielle systolique, la natrémie, la créatinémie et une FEVG<40%.** En termes de performance, le score avait une aire sous la courbe (ROC-AUC) de 0,741.

#### 3.4.4 Score traditionnel mixte

Très récemment en juillet 2022, une équipe canadienne s'est intéressé à comparer la performance des algorithmes de ML au score EHMRG issu de méthode traditionnelle sur la mortalité à 7 et 30 jours d'une population de 12 608 patients hospitalisés pour IC.(46)

Deux modèles distincts ont été utilisés pour étudier la mortalité à 7 et 30 jours. La population totale a été séparée en trois : une population de validation (N=1690), puis une population d'entraînement qui constitue le *train-set* (N=6576, 80%) et une population test qui constitue le *test-set* (N=4162, 20%).

Le **score EHRMG** comprend les 10 variables décrites plus haut. Les méthodes LASSO, XGBOOST, le réseau de neurone et le *Random Forrest* ont été utilisées pour sélectionner les variables parmi les 111 issues de leur *Dataset*.

**Ce qu'il faut retenir des résultats de cette étude est que les algorithmes de ML ne faisaient pas mieux que le modèle de base de l'EHRMG. Cependant, le ML a permis de déterminer de nouvelles variables permettant d'améliorer la ROC-AUC du score sans que la différence soit significative.** Pour la mortalité à 7 jours, la seule caractéristique qui augmentait la ROC-AUC de manière significative par rapport au modèle EHRMG était **le temps écoulé entre la première hospitalisation pour cause non cardiovasculaire du patient au cours des deux dernières années et la date d'arrivée index.**

Pour la mortalité à 30 jours, **la natrémie et le statut de résidence en soins de longue durée** ont légèrement amélioré la ROC-AUC lorsqu'ils ont été ajoutés au modèle original EHRMG de 30 jours

### 3.4.5 Scores de ML

**L'étude MIMIC-III** est une analyse de cohorte rétrospective sur 1177 patients (âge moyen 74 ans et 47% d'hommes) admis en soins intensifs pour IC entre 2001 et 2012, dont le but était d'établir un **score prédictif de la mortalité intra-hospitalière** toute cause.(47)

La mortalité observée était de 13,5% patients. La population a été scindée en deux groupes pour constituer le *train-set* (70%) et le *test-set* (30%). La sélection des variables d'intérêt parmi les vingt variables disponibles s'est faite grâce au modèle du LASSO et du XGBOOST. La significativité de chaque variable a été testée en analyse univariée, puis en analyse multivariée. Le XGBOOST a permis de déterminer six variables (**le trou anionique, le lactate, la calcémie, l'urée, la pression artérielle diastolique et l'insuffisance rénale chronique**) tandis que le LASSO a identifié ces six variables ainsi que le **volume glomérulaire moyen et l'anémie ferriprive.**

Les performances des modèles XGBOOST et LASSO étaient équivalentes. Elles ont été évaluées par leur courbe ROC avec une aire sous la courbe respectivement de 0,852 (95%CI [0,775-0,911]) et 0,865 (95%CI [0,797-0,920]) dans le *train-set* et de 0,803 (95%CI [0,685-0,903]) et 0,819 (95%CI [0,720-0,920]) dans le *test-set*. Il n'y avait **pas de différence significative** entre les deux modèles ( $p=0,19$ ).

Ces deux scores ont été ensuite **comparés séparément à un score traditionnel** existant et validé (le GWTG-HF score) et ont montré des **performances supérieures** avec des ROC-AUC respectifs de 0,842 (95%CI [0,786-0,897]), 0,8562 (95%CI [0,805-0,907]) et 0,775 (95%CI [0,716-0,833]) avec pour les deux une p-value inférieure à 0,05.

Le second article est une étude monocentrique rétrospective récente(48) publiée en mars 2022 qui s'est intéressé à **stratifier le risque de mortalité intra-hospitalière** de 5676 patients (âge moyen 70 ans, 56% d'hommes) admis en soins intensifs pour IC aiguë via la méthode de *XGBOOST*. La mortalité était de 10,5% (595 décès).

La performance du *XGBOOST* a été comparé trois autres modèles : GWTG-HF, SAPS-II, et la régression logistique.

**Le trou anionique, l'échelle de Glasgow, la diurèse, l'urée plasmatique, la pression artérielle maximale en O2, l'âge, la glycémie minimale, le calcium moyen, la fréquence respiratoire moyenne, l'excès de base artérielle moyen, la créatinine moyenne, la température moyenne, l'IMC, les plaquettes minimales et la température maximale** étaient les 15 caractéristiques les plus importantes des modèles prédictifs.

Les performances des modèles GWTG-HF, SAPS-II, de régression logistique et *XGBOOST* ont été évaluées à l'aide ROC-AUC qui étaient respectivement de 0,667 (95%CI [0,656-0,678]), 0,720 (95%CI [0,710-0,736]), 0,817 (95%CI [0,798–0,835]) et **0,831 (95%CI [0,820–0,843])**.

En plus de proposer des performances intéressantes par rapport à des scores validés traditionnels, ce score souffre peu du **surajustement**. En effet, lorsque le score proposé par le XGBOOST issu de données monocentriques a été appliqué dans la population de validité externe constituée de données multicentriques (eICU *electronic Intensive Care Unit*)(49), **les performances sont restées similaires avec une ROC-AUC de 0,809 (95%CI [0,805-0,814])**.

Ce qu'il faut retenir de ces deux études est que le ML a permis à l'aide d'algorithmes d'avoir des **performances intéressantes** avec des ROC-AUC > 0,80, **supérieures à celles de score traditionnel validé** et surtout des **résultats transposables** dans des populations différentes avec des métriques qui se maintiennent (peu de surajustement). Cependant, en raison de la nature déséquilibrée de leurs données, les ROC-AUC peuvent être **artificiellement augmentées**. Il aurait alors été souhaitable de renseigner les *PR-curves* et le F1-score.

## 4 : Objectif et rationnel de l'étude

---

A partir de l'ensemble des informations présentées, on comprend que l'IC et l'**impact des facteurs environnementaux** représentent un **enjeu important** tant à l'échelle du patient qu'en terme de santé publique. De plus, certains travaux décrivent l'impact du CO sur le système cardiovasculaire mais **aucune étude ne s'est encore intéressée à son impact pronostique** lors de la prise en charge des patients admis pour IC aiguë.

Ces questions d'ordre médical s'intègrent dans un cadre **d'innovations technologiques** en constante évolution où l'intelligence artificielle occupe **une place de plus en plus importante**.

Il apparaît donc pertinent d'utiliser ces nouveaux outils pour mieux comprendre les **facteurs prédictifs de sévérité** de nos patients insuffisants cardiaques. L'établissement, puis l'utilisation, d'un **score pronostique inédit** intégrant ces différents paramètres constituent l'enjeu de ce travail original.

## **Partie 2 : ARTICLE MÉDICAL SCIENTIFIQUE**

---

# **Supervised Machine Learning using Environmental Factors to Predict In-Hospital Outcomes in Acute Heart Failure Patients.**

Benjamin Sibia, MD <sup>1,2,3</sup>; Solenn Toupin, PhD <sup>1,2,3</sup>; Nabil Bouali, MD <sup>4,5</sup>; Jean-Baptiste Brette, MD <sup>6</sup>; Arthur Ramonatxo, MD <sup>4</sup>; Guillaume Schurtz, MD <sup>7</sup>; Kenza Hamzi, MSc <sup>1,2,3</sup>; Antonin Trimaille, MD <sup>8</sup>; Emmanuel Gall, MD <sup>1,2,3</sup>; Nicolas Piliero, MD <sup>9</sup>; Alexandre Unger, MD <sup>1,2,3,10</sup>; Stéphane Andrieu, MD <sup>11</sup>; Trecy Gonçalves, MD <sup>1,2,3</sup>; Fabien Picard, MD, PhD <sup>12</sup>; Vincent Roule, MD <sup>13</sup>; François Roubille, MD <sup>14</sup>; Sonia Houssany-Pissot, MD <sup>15</sup>; Océane Bouchot, MD<sup>16</sup>; Victor Aboyans, MD, PhD<sup>17</sup>; Reza Rossanaly Vasram, MD <sup>18</sup>; Thomas Bochaton, MD, PhD <sup>19</sup>; Damien Logeart, MD, PhD <sup>1,2</sup>; Alain Cohen Solal, MD, PhD <sup>1,2</sup>; Jérôme Cartailier, PhD <sup>2,3</sup>; Alexandre Mebazaa, MD, PhD <sup>2,20</sup>; Jean-Guillaume Dillinger, MD, PhD <sup>1,2,3</sup>; Patrick Henry, MD, PhD <sup>1,2,3</sup>; and Théo Pezel, MD, PhD <sup>1,2,3</sup>.

– on behalf of the Investigators of ADDICT-ICCU study

<sup>1</sup> Université Paris Cité, Service de Cardiologie, Hôpital Lariboisière, Assistance Publique-Hôpitaux de Paris, 75010, Paris, France.

<sup>2</sup> Inserm MASCOT - UMRS 942, University Hospital of Lariboisiere, 75010, Paris, France.

<sup>3</sup> MIRACL.ai laboratory, Multimodality Imaging for Research and Analysis Core Laboratory and Artificial Intelligence, University Hospital of Lariboisiere (AP-HP), 75010, Paris, France.

<sup>4</sup> Department of Cardiology, University Hospital of Poitiers, 86000 Poitiers, France.

<sup>5</sup> Service de Cardiologie, Centre hospitalier de Saintonge, 11, boulevard Ambroise-Paré, 17100 Saintes, France.

<sup>6</sup> Cardiology department, Rangueil University Hospital, Toulouse, France

<sup>7</sup> Department of Cardiology, University Hospital of Lille, France.

- <sup>8</sup> Department of Cardiovascular Medicine, Nouvel Hôpital Civil, Strasbourg University Hospital, 67000 Strasbourg, France
- <sup>9</sup> Univ. Grenoble Alpes, service de Cardiologie, CHU Grenoble Alpes, 38000 Grenoble, France
- <sup>10</sup> Department of Cardiology, Hôpital universitaire de Bruxelles, Hôpital Erasme, Université libre de Bruxelles, Brussels, Belgium
- <sup>11</sup> Service de cardiologie, Hôpital Henri Duffaut, 84902 AVIGNON, France
- <sup>12</sup> Université Paris Cité, Service de Cardiologie, Hôpital Cochin, Paris, France
- <sup>13</sup> Department of Cardiology, Caen University Hospital, Caen, France
- <sup>14</sup> Cardiology Department, University Hospital, PhyMedExp, University of Montpellier, INSERM, CNRS, CHRU, INI-CRT, Montpellier, France
- <sup>15</sup> Service de cardiologie et médecine aéronautique, Hôpital d'Instruction des Armées Percy, 92140 Clamart, France.
- <sup>16</sup> Service de Cardiologie, Centre hospitalier Annecy Genevois, 74370 Epagny Metz-Tessy.
- <sup>17</sup> Department of Cardiology, University Hospital of Limoges, France.
- <sup>18</sup> Department of Cardiology, Felix-Guyon University Hospital, Saint-Denis-de-La-Reunion, France
- <sup>19</sup> Intensive Cardiological Care Division, Louis Pradel Hospital, Hospices Civils de Lyon, Bron, France.
- <sup>20</sup> Department of Cardiology, Nîmes University Hospital, Montpellier University, Nîmes, France.
- <sup>21</sup> Department of Anesthesia, Burn and Critical Care Medicine, Hôpital Lariboisière, Assistance Publique-Hôpitaux de Paris, France.

Running Head: Machine learning for predicting in-hospital outcomes in acute heart failure

Total Word count: 2,847

Relationship With Industry Policy:

All the authors declare no conflict of interest and no financial disclosure in relation to this study.

Address for correspondence:

Dr. Theo Pezel, MD, PhD

Service de Cardiologie, Hôpital Lariboisière, Assistance Publique-Hôpitaux de Paris,

2 rue Ambroise Paré, 75010 Paris, France

Phone: +33 1 49 95 82 24, Fax: +33 1 49 95 62 99,

Email: [theo.pezel@aphp.fr](mailto:theo.pezel@aphp.fr)

1 **ABSTRACT**

2 **Background and aims:** While few traditional scores are available for risk stratification of  
3 patients hospitalized for acute heart failure (AHF), the potential benefit of machine-learning  
4 (ML) is not well-established. We aimed to assess the feasibility and accuracy of a supervised  
5 ML-model including environmental factors to predict in-hospital major adverse events (MAE)  
6 in patients hospitalized for AHF.

7 **Methods:** In April 2021, a French national prospective multicentre study included all  
8 consecutive patients hospitalized in ICCU. Patients admitted for AHF were included in the  
9 analyses. A ML-model involving automated feature selection by Least absolute shrinkage and  
10 selection operator (LASSO) and model building with a Random Forest (RF) algorithm was  
11 developed. The primary composite outcome was in-hospital MAE defined by death,  
12 resuscitated cardiac arrest, or cardiogenic shock requiring assistance.

13 **Results:** Among 459 patients included (age  $68\pm 14$  years, 68% male), 47 experienced in-  
14 hospital MAE (10.2%). Seven variables were selected by LASSO for predicting MAE in the  
15 training dataset (N=322): mean arterial pressure, ischemic aetiology, sub-aortic velocity time  
16 integral,  $E/e'$ , tricuspid annular plane systolic excursion, recreational drug use and exhaled  
17 carbon monoxide level. The RF model showed the best performance compared with other  
18 evaluated models (AUROC=0.82, PR-AUC=0.48, F1-score=0.56). Our ML-model exhibited a  
19 higher AUROC compared with an existing score for prediction of MAE (AUROC for our ML-  
20 model: 0.82 vs Acute HF-score: 0.57;  $p<0.001$ ).

21 **Conclusions:** Our ML-model including in particular environmental variables exhibited a better  
22 performance than traditional statistical methods to predict in-hospital outcomes in patients  
23 admitted for AHF.

24 **TRIAL REGISTRATION:** ClinicalTrials.gov Identifier: NCT05063097

- 1 **KEYWORDS:** acute heart failure; environmental factors; carbon monoxide; recreational
- 2 drugs; machine learning; intensive cardiac care unit.

- 1 **ABBREVIATIONS:**
- 2 AHF: Acute heart failure
- 3 CI: Confidence interval
- 4 CO: Carbon monoxide
- 5 ICCU: Intensive cardiac care unit
- 6 LASSO: Least absolute shrinkage and selection operator
- 7 MAE: Major adverse events
- 8 ML: Machine learning
- 9 PR: Precision-recall
- 10 TAPSE: Tricuspid annular plane systolic excursion
- 11 VTI: Velocity time integral

## **INTRODUCTION**

Acute heart failure (AHF) continues to be the main cause for hospital admissions among patients over the age of 65 years and is responsible for a mortality exceeding 25% (1), including an 8% mortality rate within hospitals at the 30-day mark.(2) It is therefore crucial to offer a high-performance prognostic stratification tool to identify these patients at risk of in-hospital poor outcomes. Beyond the clinical, biological and echocardiographic parameters well-established as prognostic factors,(3,4) several recent studies have shown a role for environmental factors in the in-hospital prognosis of these AHF patients. Indeed, recent studies reported that recreational drug use could be a strong prognosticator of in-hospital outcomes due to multifactorial acute cardiotoxicity.(5–7) Indeed, our working group has recently showed that recent consumption of recreational drugs in AHF patients was independently associated with a risk of in-hospital major adverse event (MAE) multiplied by more than seven compared to non-users.(6) Another recent subanalysis showed that the level of exhaled CO, as a biomarker of air pollution in the patient's lungs was independently associated with a higher risk of in-hospital mortality in AHF patients.(8) Furthermore, a large meta-analysis emphasized the impact of air pollution, including carbon monoxide (CO) due to active and passive smoking, as an aggravating factor of AHF.(9) However, until now, no tool is available to take into account all of these variables simultaneously.

Machine learning (ML) methods can recognize patterns in large datasets with a multiplicity of variables and can be used for improving risk stratification. Recently, ML techniques have emerged as highly effective methods for prognostic prediction and decision-making in several cardiovascular diseases.(10,11) Although of potential interest, supervised ML techniques using all this spectrum of variables have not been evaluated for risk stratification in AHF patients.

Therefore, the aim of this study was to investigate the feasibility and accuracy of a supervised ML-model using clinical, biological, environmental, and echocardiographic data to predict in-hospital MAE in a consecutive cohort of patients hospitalized for AHF.

## **METHODS**

### **Study population**

All details of the Addiction in Intensive Cardiac Care Units (ADDICT-ICCU) study design have been published previously.(12) Briefly, this is a multi-centre, prospective, observational study of all consecutive patients aged  $\geq 18$  years admitted to ICCUs over two weeks in April 2021 at 39 centres across France. The main exclusion criterion was hospitalization for either a planned interventional procedure or more than 24 hours at any hospital facility before admission to the ICCU. This was to prevent the risk of obtaining a negative urine drug assay in patients with recreational drug consumption more than 24 hours prior to admission. The methodology of the baseline characteristics collection is detailed in **eSupplement 1**. In this study, we decided to focus the analysis on patients admitted for AHF defined by symptoms and/or signs of heart failure with evidence of diastolic or systolic dysfunction by echocardiography and elevated levels of natriuretic peptide (BNP  $>35$  pg/ml and/or NT-proBNP  $>125$  pg/ml)(3) (**Figure 1**). The main admission diagnosis was adjudicated by two independent experts at the end of the hospitalization following the current guidelines (**eSupplement 2**). The investigation conforms with the principles outlined in the *Declaration of Helsinki*. All patients had a written informed consent obtained at enrolment and was informed of the urine drug assay and exhaled CO measurement. This study has been approved by the Ethics Committee (Committee for the Protection of Human Subjects, Ile de France-7, France) and registered on the Clinicaltrials.gov website under the number NCT05063097.

### **Primary outcome**

The primary outcome was in-hospital MAE defined as a composite of all-cause mortality, cardiogenic shock (requiring medical or mechanical hemodynamic support) and resuscitated cardiac arrest (severe ventricular arrhythmia requiring defibrillation or anti-arrhythmic agents). An independent committee of experts was in charge of reviewing anonymized medical documents according to standardized definitions.(13)

### **Baseline characteristics collection**

The database records of the AHF population comprise baseline data that include clinical, demographic, reason for hospitalization, list of medications especially cardiovascular medications at admission, history of cardiovascular disease. Biological data were collected systematically upon admission, including haemoglobin, serum potassium, creatinine, the peak of troponin (hsTNI), the N-terminal prohormone of B-type natriuretic peptide (NT-proBNP) or B-type natriuretic peptide (BNP). Standardized transthoracic echocardiography was performed systematically within the first 24 hours of admission for all patients by a cardiologist and reviewed by a senior.

Regarding environmental factors, the presence of recreational drugs was determined through urine analysis using a dedicated drug assay (NarcoCheck<sup>®</sup>, Kappa City Biotech SAS, Montluçon, France) within two hours of admission to the ICCU with excellent performance (the sensitivity and specificity of the urine drug assay were: 91.7% and 97.7%, respectively).(6)

The following recreational drugs will be screened for all patients: cannabinoids (tetrahydrocannabinol [THC]), including cannabis and hashish; cocaine and metabolites, including cocaine and crack; amphetamines; 3,4-methylenedioxy-methylamphetamine (MDMA or ecstasy); heroin and other opioids. Finally, a standardized exhaled CO

measurement was systematically performed with a CO-Check Pro device (Micro Direct Diagnostics Ltd, UK) immediately on arrival in ICCU.

## **MACHINE LEARNING**

Fourteen clinical, 4 biological, 3 environmental, and 7 echocardiographic parameters were available (detailed list provided in **eSupplement 3**). Machine learning involved feature selection by least absolute shrinkage and selection operator (LASSO) algorithm, and model building with Random Forest algorithm following guidelines for machine learning.(14)

### **Feature selection**

After imputation of missing values by the K-nearest-neighbours with a value of K=5, feature selection was performed using LASSO (**Figure 2**), which is a penalized logistic regression that reduces the number of variables (glmnet R package).(15) The regularization parameter  $\lambda$  controls the strength of the shrinkage to get the best compromise between prediction performance and interpretability. The best  $\lambda$  was found using 10-fold cross-validation and the “1 standard-error rule”, for which the area under receiver operating curve (AUROC) of the most parsimonious model is no more than one standard error above the AUROC of the best model.(16)

### **Model building**

To build the prediction model, we compared four modelling strategies with: logistic regression, Random Forest, extreme gradient boosting model (XGBoost), and LASSO. To assess the performance of these algorithms, we used several metrics: AUROC, precision-recall area under curve (PRAUC), F1-score, sensitivity, specificity, positive and negative prediction value, accuracy, balanced accuracy, Cohen’s Kappa, Brier score, and Mc-Nemar test p-value. Due to the low occurrence of events and the imbalanced nature of the data, the F1-score and the

precision-recall curve were preferred to determine the best model to minimize the overall misclassification cost and mitigate the problem of imbalanced data. For the prediction of in-hospital MAE, the best model was the ML-model using Random Forest (**Figure 3**). Random Forest corresponds to a method based on decision trees, where each tree is trained on a random sample from the dataset. All the results are then combined to form the final algorithm.

For model development, patients hospitalized for AHF in the ADDICT-ICCU study were randomly partitioned into training (70%) and validation (30%) cohorts (**eSupplement 4**). We used a 10-fold cross-validation process to tune the model hyperparameters using grid search. In addition to the internal validation dataset (30% of the ADDICT-ICCU study).

### **Statistical analysis**

Continuous data will be reported as mean  $\pm$  standard deviation for normally distributed data or as medians and interquartile range for non-normally distributed data. Categorical data will be reported as counts and percentages. Between-group comparisons will be performed using the Student's t-test or Mann-Whitney test for continuous variables and the chi-squared or Fisher's exact test for categorical variables, as appropriate.

Model discrimination was measured using the AUROC and its 95% CIs. In addition, the positive predictive value (or precision) and the sensitivity (recall) across all possible risk thresholds for predicting in-hospital MAE were plotted using precision-recall curves and their area-under-the-curve. The PRAUC, unlike the AUROC, is not affected by the number of true-negative results. In datasets with small event rates and therefore a large expected true-negative rate, such as the current study, the PRAUC is better for comparing different models. For both the AUROC and PRAUC, values closer to 1 correspond to more accurate models. DeLong's test was used to compare the AUROC of each model. We also calculated the F1-score,

sensitivity, specificity, positive predictive value, and negative predictive value. In addition, we calculated a Brier score for each model as a measure of model accuracy.

Model calibration was measured using: i) the Hosmer and Lemeshow goodness-of-fit test; ii) the reliability component of the Brier score; and iii) graphically with plots showing the observed and predicted proportion of events, grouped by level of risk.

A two-tailed p-value  $<0.05$  will be considered statistically significant. All data will be analysed using R software, version 4.0.3 (R Project for Statistical Computing, R Foundation, Vienna, Austria).

## **RESULTS**

### **Study population**

Detailed flowchart of the study is depicted in **Figure 1**. Among the whole population of the study, 459 patients were admitted for AHF (68% male, mean age  $68\pm 14$  years), 46% had ischemic cardiomyopathy and 53% with LVEF value  $<50\%$ . All characteristics of the population are described in **Table 1**. Concerning the main admission diagnosis associated with AHF, 135 (29%) patients had acute coronary syndromes, 39 (9%) had severe conduction/arrhythmia abnormalities, 29 (6%) had pulmonary embolism, 20 (4%) had myocarditis, 9 (2%) had Takotsubo syndrome, 6 (1%) had spontaneous coronary dissection, 24 (5%) had other cardiovascular diagnoses, and 197 (44%) had isolated AHF.

### **In-hospital outcomes**

During a median duration of hospitalization of 7 days (IQR 5-12 days) in the ICCU, 47 patients (10.2%) experienced in-hospital MAE, including 18 (3.9%) in-hospital deaths, 8 (1.7%) resuscitated cardiac arrests, and 21 (4.6%) cardiogenic shocks requiring medical and/or mechanical hemodynamic support.

### **Feature selection**

Using the LASSO feature selection method with the best regularization parameter  $\lambda$ , seven of the available variables were selected for the ML-model (2 clinical, 2 environmental, and 3 echocardiographic) in descending order of importance: left ventricular outflow track velocity time integral (LVOT VTI), peak E/e' ratio, exhaled CO level, mean arterial pressure, ischemic AHF aetiology, recreational drugs use, and tricuspid annular plane systolic excursion (TAPSE) (**Figure 2**).

### **Model building**

The ML-model using Random Forest exhibited the best global performance for prediction of in-hospital MAE compared with LASSO, stepwise regression, and XG-Boost using all the metrics available (**Table 2**). The performance of the ML-model in women and men is provided in **eSupplement 5**.

### **Model performance compared with traditional methods and existing score**

The ML-model showed a better accuracy to predict in-hospital MAE compared with a logistic regression model (AUROC: 0.82 vs. 0.78; PRAUC: 0.48 versus 0.35, **Figure 3**). The proposed ML-model exhibited the best performance for prediction of in-hospital MAE compared with the ACUTE-HF score,<sup>(17)</sup> with a higher AUROC (0.82 vs. 0.57) and PRAUC (0.48 versus 0.12, **Figure 4**).

### **Calibration of our machine learning model**

The Hosmer-Lemeshow goodness-of-fit test showed no significant difference between the observed and predicted proportions of in-hospital MAE ( $p=0.91$ ), in line with the Brier-score

of 0.07. All these findings indicated a good calibration of the proposed ML-model across all risk levels (**eSupplement 6**).

## **DISCUSSION**

In a prospective multicentre cohort study of consecutive patients hospitalized for AHF in ICCU, we built a ML model to predict in-hospital outcomes (**Graphical abstract**). This ML model was based on few simple clinical, environmental, and echocardiographic variables: mean arterial pressure, ischemic AHF aetiology, recreational drugs use, exhaled CO level, LVOT VTI, peak E/e' ratio, and TAPSE. The main findings are as follows: i) the ML-model exhibited a better performance than traditional statistical methods; ii) the ML-model outperformed existing scores as ACUTE HF-score; and iii) its calibration was excellent. The study presents the first large-scale investigation of ML for prognostic risk evaluation using a wide spectrum of data ranging from clinical to echocardiography through environmental data. These findings suggest that ML and environmental factors could have an important role for risk stratification and decision-making in patients with AHF.

In line with prior large cohorts with a rate of in-hospital outcomes between 5 and 10%, (2,18,19) we reported a rate of in-hospital MAE of 10.2% and in-hospital death of 1.7%. In our cohort, MAE was more driven by cardiogenic shock than all-cause in-hospital mortality and, resuscitated cardiac arrest events.

Regarding the environmental factors selected by ML, exhaled CO level was one of the highest ranked features for risk-prediction. It has been previously described as a measure of particular interest for the study of active and passive smoking exposure as an important marker of air pollution. An ancillary study of patients from the Framingham cohort showed an association

between exhaled CO level and the occurrence of cardiovascular outcomes regardless of the patient's smoking status.(20) This suggests that exhaled CO level could be an interesting marker for assessing the cardiovascular impact of air pollution. Indeed, an important meta-analysis showed a strong association between daily increases in gaseous air pollutants (particulate matter, carbon monoxide, Sulphur dioxide, nitrogen dioxide, ozone) and the occurrence of severe HF outcomes.(9) Although more studies from developing nations are required, air pollution is a pervasive public health issue with major cardiovascular and health economic burden, and it should remain a key target for global health policy. In line with these results, our working group has very recently shown that elevated CO level is independently associated with a 10-fold increase in in-hospital MAE and 6-fold 1-year mortality in patients hospitalized for acute cardiac events.(8) All these findings are also supported by Tun et al. who showed that increased levels of exhaled CO were associated with an adverse cardiovascular biomarker profile and a higher risk of AHF, especially with reduced LVEF.(21)

The second environmental factor selected by ML was recreational drugs use. This finding is in line with a prior analysis from ADDICT-ICCU study showing that recreational drugs use is strongly associated with the occurrence of in-hospital outcomes with an incremental prognostic value above traditional risk factors.(6) These findings can be explained by several types of sympathomimetic effects of recreational drugs which can increase heart rate, blood pressure, temperature and consequently myocardial oxygen demand.(22)

Two clinical variables were selected by ML with mean blood pressure and ischemic aetiology of HF. Blood pressure is a well-known feature present in most of AHF scores.(23,24) In the current study, we preferred the mean blood pressure than systolic blood pressure because it includes the effect of the diastolic value which can reflect different situations as the aortic

regurgitation or vasoplegia. Knowing that ischemic aetiology is also well described as prognosticator for HF outcomes, two scores already published have also selected this feature as a prognosis factor.(19,25)

Beyond the traditional clinical data, the current study emphasized the importance of echocardiographic parameters at baseline in ICCU for risk stratification in AHF patients. Indeed, among the seven parameters selected by ML, three were echocardiographic variables with: LVOT VTI, peak E/e' ratio, and TAPSE which are mean full in the day-to-day practice and can be performed by any cardiologist. Several studies have shown that left ventricular dysfunction with a drop in cardiac output,(3,4) assessed by the LVOT VTI, was associated with a poor prognosis in a wide spectrum of acute cardiovascular diseases. Right ventricular dysfunction defined using TAPSE,(26,27) and left ventricular filling pressure assessed by the peak E/e' ratio have also been described as two strong prognosticators in the acute setting of AHF.(2,3,28) Consistently, these three echocardiographic parameters have a central place in the last ESC guidelines.(3,4)

Finally, our ML-model showed the best accuracy to predict in-hospital outcomes compared to traditional methods and existing score assessed of AHF patients in ICCU. In particular, its performance outstripped ACUTE-HF score.(17) Up to date, no specific ML-model including clinical, biological, environmental and echocardiographic data has been built.

### **Study limitations**

First, the sample size of patients analysed was decreased due to the split between training and testing cohorts. In addition, the exclusion of 8 centres to build an external validation cohort also lowered the sample size. Nevertheless, as there are no other studies on the AHF population that

have measured exhaled CO and recreational drugs use, the results necessarily had to be validated on a part of the ADDICT-ICCU study. Second, even if recreational drugs use,(6) and high exhaled CO level have recently reported as increasing cardiovascular events,(8) and had pathophysiological background on cardiovascular system, their screenings are not available in daily practice, hence a limited translation in practice. Third, missing values always represent a challenge in medical research. In line with the literature,(29) we imputed missing data using with KNN algorithm for all data with a missing data rate <40%. Fourth, some ICCU scores could not be calculated due to missing biological collected data. The only score we could compare with was ACUTE HF-score,(17) and results may be interpreted with caution. Finally, this study assessed only the performance of ML, but not its practical implementation. Additional parameters would have been of interest, especially renal function (not only creatininemia, CRP, intake of some drugs...). Once validated in clinical routine, further randomized clinical trials based upon ML risk-stratification using environmental data should be considered.

## **CONCLUSION**

In a large multicentre registry of consecutive patients hospitalized for AHF in ICCU, a ML-model using seven clinical, environmental – including exhaled CO level and recreational drugs use – and echocardiographic variables was an accurate method for predicting in-hospital MAE. This ML-model exhibited a higher prognostic value to predict in-hospital MAE than all statistical methods or existing score.

## **ACKNOWLEDGEMENTS**

We thank the medical, paramedical and research staff of all investigating centres involved.

## **FUNDING**

None.

## **CONFLICT OF INTEREST**

All the authors declare no conflict of interest and no financial disclosure in relation to this study.

## **REFERENCES**

1. Crespo-Leiro MG, Anker SD, Maggioni AP, Coats AJ, Filippatos G, Ruschitzka F, et al. European Society of Cardiology Heart Failure Long-Term Registry (ESC-HF-LT): 1-year follow-up outcomes and differences across regions. *Eur J Heart Fail.* 2016 Jun;18(6):613–25.
2. Lombardi C, Peveri G, Cani D, Latta F, Bonelli A, Tomasoni D, et al. In-hospital and long-term mortality for acute heart failure: analysis at the time of admission to the emergency department. *ESC Heart Fail.* 2020 Oct;7(5):2650–61.
3. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J.* 2021 Sep 21;42(36):3599–726.
4. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2023 Focused Update of the 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J.* 2023 Oct 1;44(37):3627–39.
5. Nishimura M, Bhatia H, Ma J, Dickson SD, Alshawabkeh L, Adler E, et al. The Impact of Substance Abuse on Heart Failure Hospitalizations. *Am J Med.* 2020 Feb;133(2):207-213.e1.
6. 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 Br Card Soc.* 2023 Oct 12;109(21):1608–16.
7. Curran L, Nah G, Marcus GM, Tseng Z, Crawford MH, Parikh NI. Clinical Correlates and Outcomes of Methamphetamine-Associated Cardiovascular Diseases in Hospitalized Patients in California. *J Am Heart Assoc.* 2022 Aug 16;11(16):e023663.
8. Dillinger JG, Pezel T, Delmas C, Schurtz G, Trimaille A, Piliero N, et al. Carbon monoxide and prognosis in smokers hospitalised with acute cardiac events: a multicentre, prospective cohort study. *eClinicalMedicine [Internet].* 2024 Jan 1 [cited 2024 Jan 3];67. Available from: [https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370\(23\)00578-3/fulltext](https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(23)00578-3/fulltext)
9. Shah AS, Langrish JP, Nair H, McAllister DA, Hunter AL, Donaldson K, et al. Global association of air pollution and heart failure: a systematic review and meta-analysis. *The Lancet.* 2013 Sep 21;382(9897):1039–48.
10. Mark E, Goldsman D, Gurbaxani B, Keskinocak P, Sokol J. Using machine learning and an ensemble of methods to predict kidney transplant survival. *PLoS One.* 2019;14(1):e0209068.
11. Pezel T, Sanguineti F, Garot P, Untersee T, Champagne S, Toupin S, et al. Machine-Learning Score Using Stress CMR for Death Prediction in Patients With Suspected or Known CAD. *JACC Cardiovasc Imaging.* 2022 Jul;S1936878X22003412.

12. 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. *Arch Cardiovasc Dis*. 2022 Sep;S1875213622001553.
13. Hicks KA, Tchong JE, Bozkurt B, Chaitman BR, Cutlip DE, Farb A, et al. 2014 ACC/AHA Key Data Elements and Definitions for Cardiovascular Endpoint Events in Clinical Trials: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (Writing Committee to Develop Cardiovascular Endpoints Data Standards). *J Am Coll Cardiol*. 2015 Jul 28;66(4):403–69.
14. Sengupta PP, Shrestha S, Berthon B, Messas E, Donal E, Tison GH, et al. Proposed Requirements for Cardiovascular Imaging-Related Machine Learning Evaluation (PRIME): A Checklist: Reviewed by the American College of Cardiology Healthcare Innovation Council. *JACC Cardiovasc Imaging*. 2020 Sep;13(9):2017–35.
15. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Series B Stat Methodol* 1996;58: 267–288.
16. Krstajic D, Buturovic LJ, Leahy DE, Thomas S. Cross-validation pitfalls when selecting and assessing regression and classification models. *J Cheminformatics*. 2014 Mar 29;6(1):10.
17. Cameli M, Pastore MC, Carli GD, Henein MY, Mandoli GE, Lisi E, et al. ACUTE HF score, a multiparametric prognostic tool for acute heart failure: A real-life study. *Int J Cardiol*. 2019 Dec 1;296:103–8.
18. Abraham WT, Fonarow GC, Albert NM, Stough WG, Gheorghide M, Greenberg BH, et al. Predictors of in-hospital mortality in patients hospitalized for heart failure: insights from the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF). *J Am Coll Cardiol*. 2008 Jul 29;52(5):347–56.
19. Miró Ò, Rossello X, Gil V, Martín-Sánchez FJ, Llorens P, Herrero-Puente P, et al. Predicting 30-Day Mortality for Patients With Acute Heart Failure in the Emergency Department: A Cohort Study. *Ann Intern Med*. 2017 Nov 21;167(10):698–705.
20. Cheng S, Enserro D, Xanthakis V, Sullivan LM, Murabito JM, Benjamin EJ, et al. Association of exhaled carbon monoxide with subclinical cardiovascular disease and their conjoint impact on the incidence of cardiovascular outcomes. *Eur Heart J*. 2014 Nov 7;35(42):2980–7.
21. Tun B, Ehrbar R, Short M, Cheng S, Vasan RS, Xanthakis V. Association of Exhaled Carbon Monoxide With Ideal Cardiovascular Health, Circulating Biomarkers, and Incidence of Heart Failure in the Framingham Offspring Study. *J Am Heart Assoc*. 2020 Nov 3;9(21):e016762.

22. Ghuran A, Van Der Wieken LR, Nolan J. Cardiovascular complications of recreational drugs. *BMJ*. 2001 Sep 1;323(7311):464–6.
23. Lee DS, Stitt A, Austin PC, Stukel TA, Schull MJ, Chong A, et al. Prediction of Heart Failure Mortality in Emergent Care. *Ann Intern Med*. 2012 Jun 5;156(11):767–75.
24. Peterson PN, Rumsfeld JS, Liang L, Albert NM, Hernandez AF, Peterson ED, et al. A Validated Risk Score for In-Hospital Mortality in Patients With Heart Failure From the American Heart Association Get With the Guidelines Program. *Circ Cardiovasc Qual Outcomes*. 2010 Jan;3(1):25–32.
25. Stiell IG, Clement CM, Brison RJ, Rowe BH, Borgundvaag B, Aaron SD, et al. A risk scoring system to identify emergency department patients with heart failure at high risk for serious adverse events. *Acad Emerg Med Off J Soc Acad Emerg Med*. 2013 Jan;20(1):17–26.
26. Zehender M, Kasper W, Kauder E, Schönthaler M, Geibel A, Olschewski M, et al. Right ventricular infarction as an independent predictor of prognosis after acute inferior myocardial infarction. *N Engl J Med*. 1993 Apr 8;328(14):981–8.
27. Berrill M, Ashcroft E, Fluck D, John I, Beeton I, Sharma P, et al. Right Ventricular Dysfunction Predicts Outcome in Acute Heart Failure. *Front Cardiovasc Med*. 2022;9:911053.
28. 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. *J Am Heart Assoc*. 2019 Sep 3;8(17):e013675.
29. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Med Res Methodol*. 2017 Dec 6;17(1):162.

## **FIGURES AND LEGENDS:**

### **Figure 1. Flowchart of the study population.**

*ICCU: Intensive Cardiac Care Unit; ML: machine learning.*

### **Figure 2. Feature importance selected by LASSO.**

LASSO was used to evaluate the worth of each variable by measuring the log-rank-based variable importance with respect to the outcome and then to rank the attributes by their individual evaluations (**top to bottom**). Only attributes resulting in log-rank based variable importance >30 (above the **orange line**) were subsequently used for the model building.

*CO: carbon monoxide ; HF: Heart failure ; LASSO: Least absolute shrinkage and selection operator ; LVOT VTI: left ventricular outflow tract velocity-time integral; MAP: mean arterial ; TAPSE: Tricuspid annular plane systolic excursion.*

### **Figure 3. Performance of machine learning models for prediction of in-hospital MAE.**

Receiver-operating characteristic (AUROC, panel A) and precision recall curves (PRAUC, panel B) for prediction of in-hospital MAE were used to compare LASSO, Random Forest, Logistic Regression, and XGBoost, developed using the set of variables selected by LASSO.

*LASSO: Least absolute shrinkage and selection operator*

### **Figure 4. Comparison of our model with ACUTE-HF score**

Receiver-operating characteristic (AUROC, panel A) and precision recall curves (PRAUC, panel B) for prediction of in-hospital MAE were used to compare our ML model to the ACUTE-HF score.

### **Graphical abstract. Prediction models using machine learning including environmental factors to assess the risk of in-hospital major adverse events in patients hospitalized for acute heart failure.**

*ADDICT-ICCU: Addiction in intensive cardiac care unit ; AHF: acute heart failure ; CO: carbon monoxide ; LASSO: Least absolute shrinkage and selection operator ; MAE: major adverse events ; PR: Precision-recall ; TAPSE: Tricuspid annular plane systolic excursion ; TTE: transthoracic echocardiogram ; LVOT VTI: left ventricular outflow tract velocity-time integral ; XGBoost: extreme gradient boosting.*

**Figure 1.**

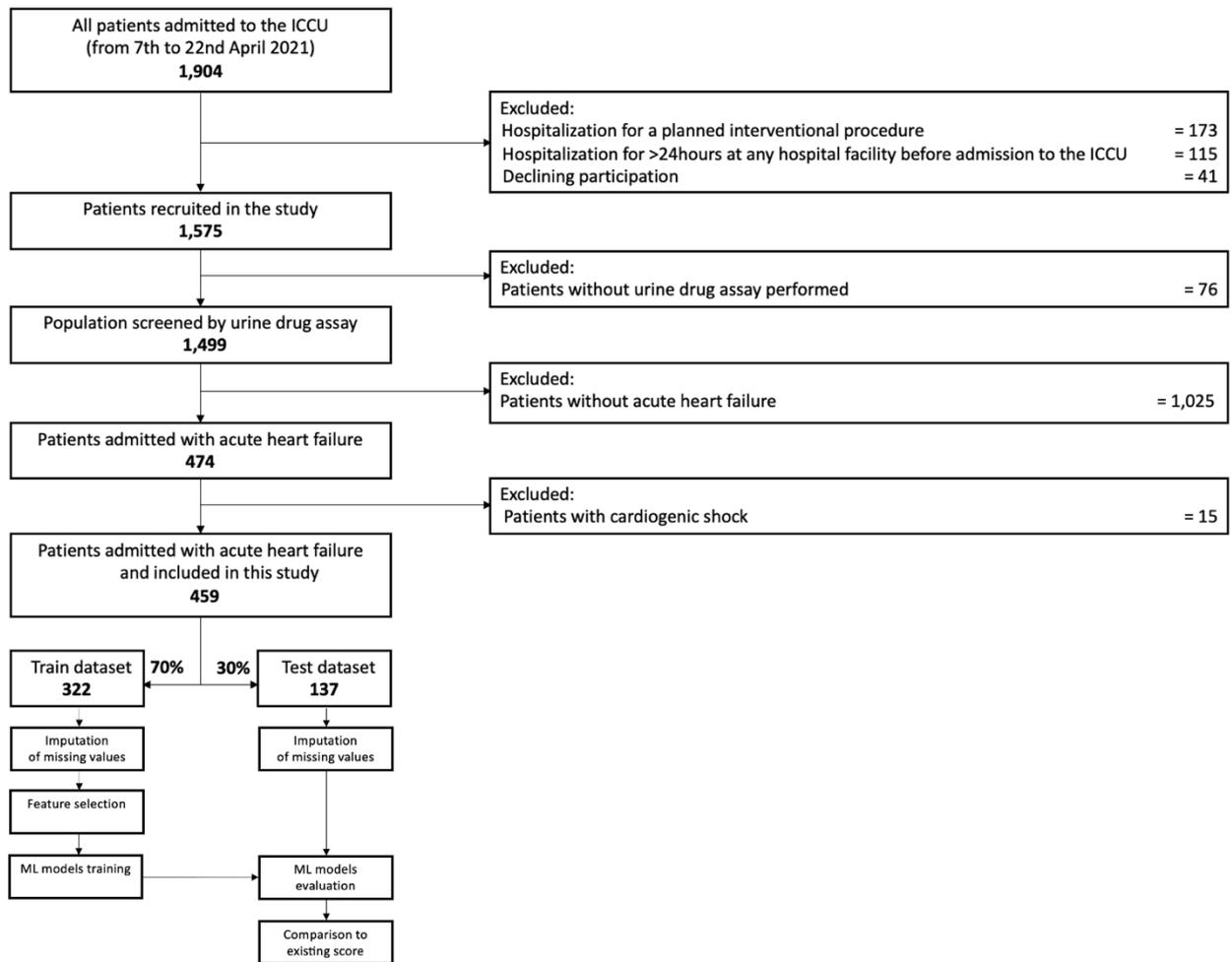


Figure 2.

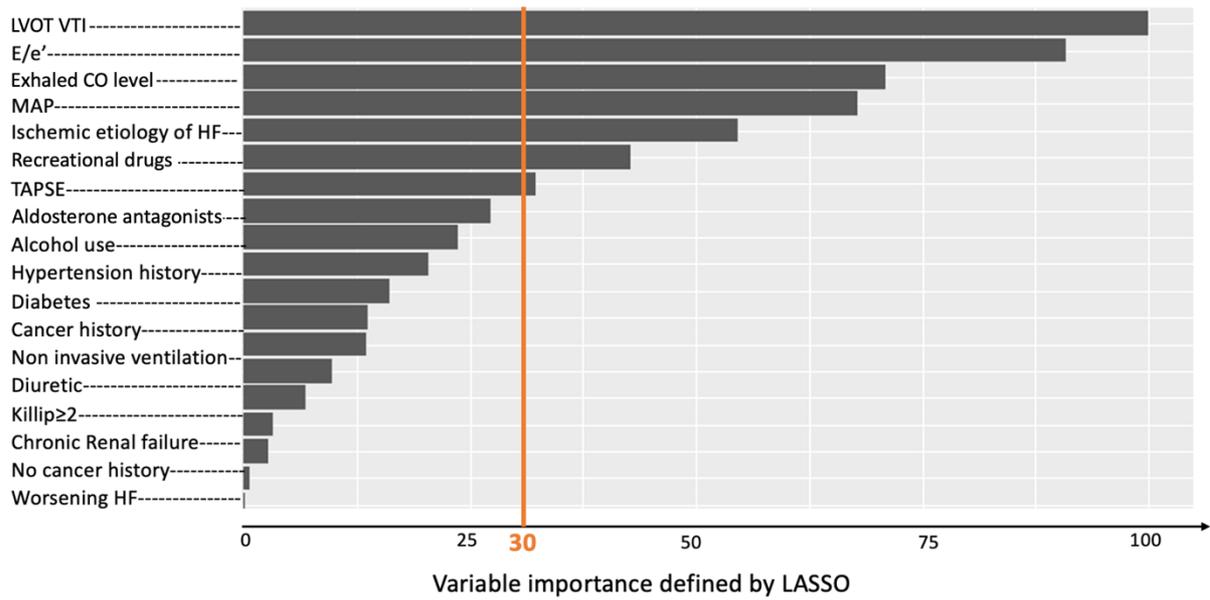


Figure 3.

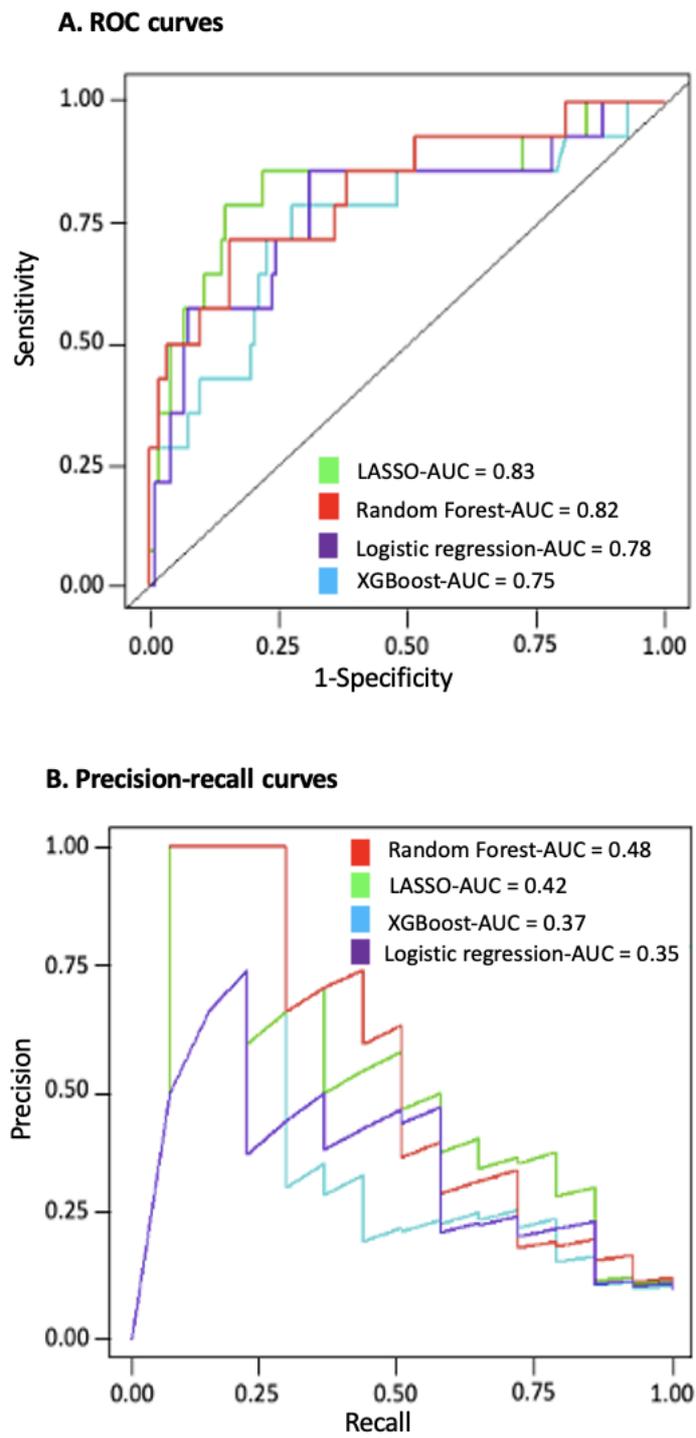
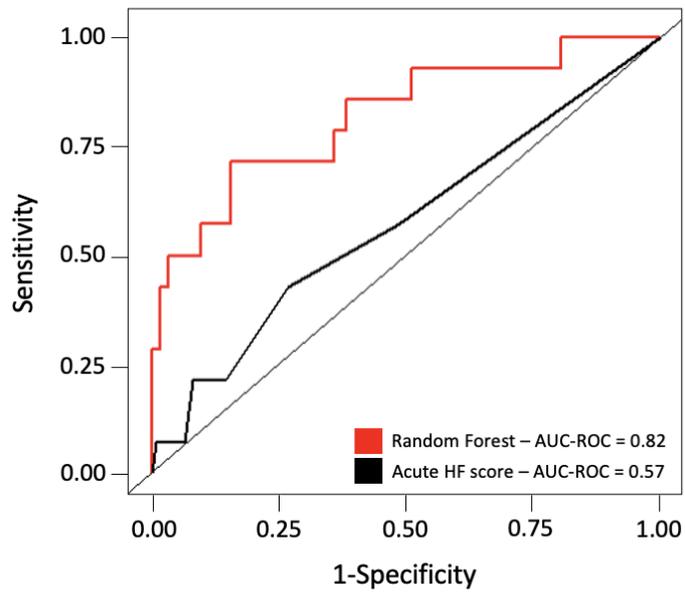
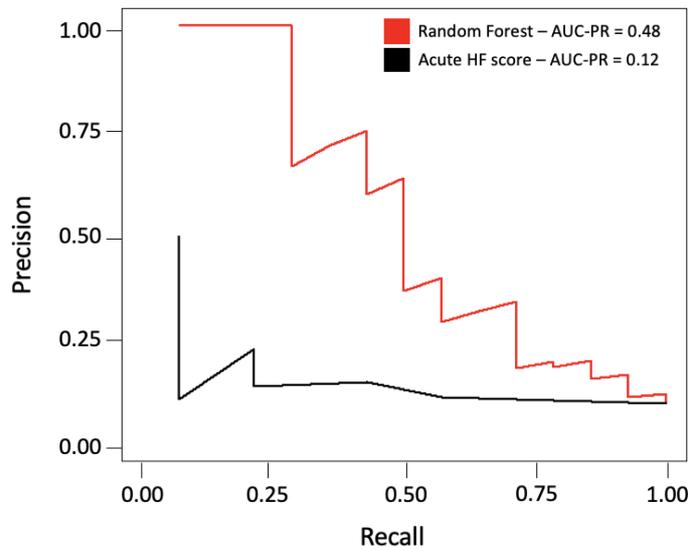


Figure 4.

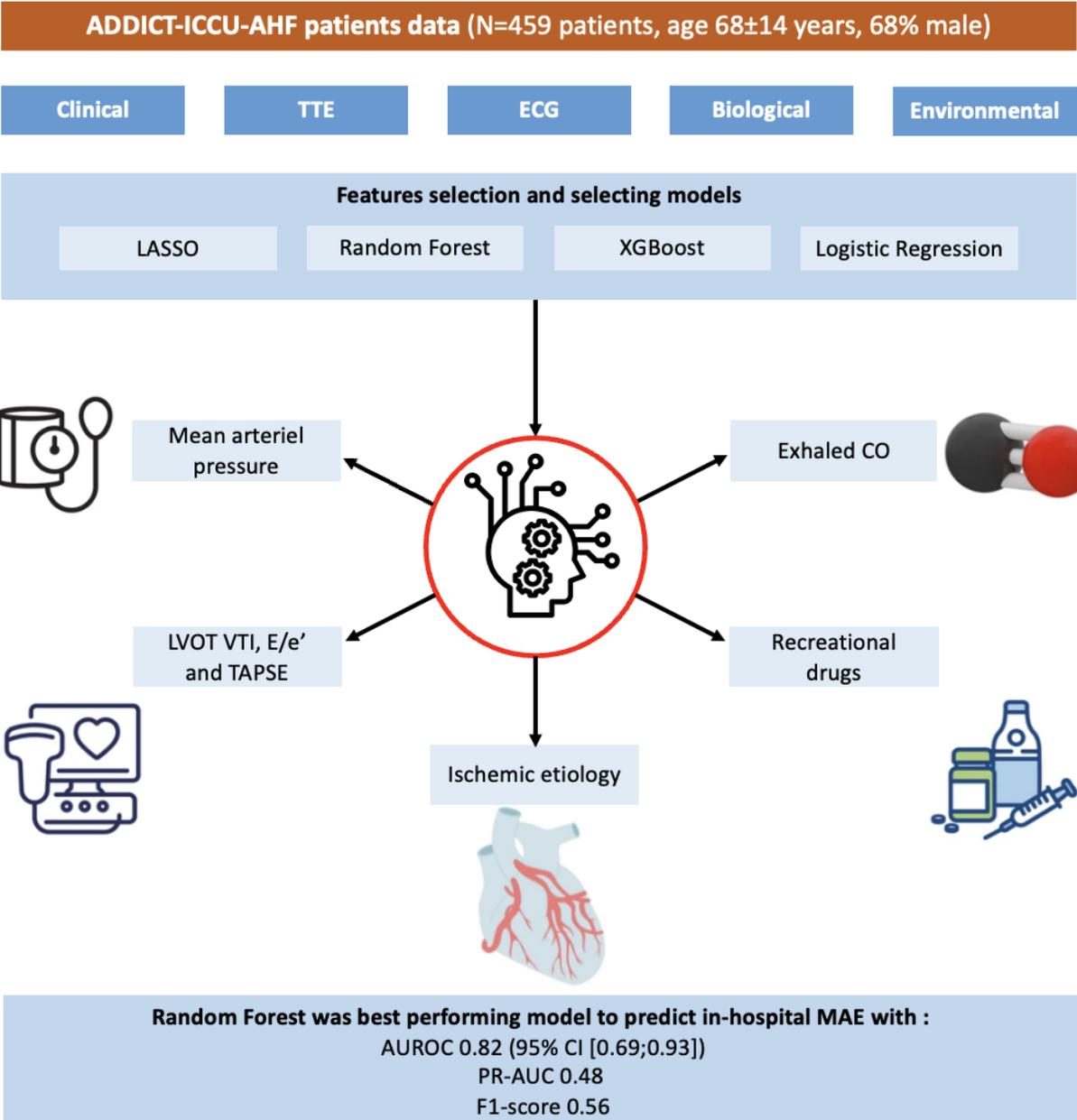
**A. ROC curves**



**B. PR curves**



Graphical abstract.



## **TABLES**

**Table 1. Baseline characteristics according to the occurrence of in-hospital major adverse events (N=459).**

	<b>All patients (N=459)</b>	<b>Patients without MAE (N=412)</b>	<b>Patients with MAE (N=47)</b>	<b>p-value</b>
<b>Demographic data</b>				
Age, years	68 ± 14	68 ± 14	70 ± 15	0.37
Males, n (%)	311 (67.8%)	276 (67.0%)	35 (74.5%)	0.30
Body mass index, kg/m <sup>2</sup>	28 ± 6	28 ± 7	27 ± 5	0.32
<b>Cardiovascular risk factors, n (%)</b>				
Diabetes mellitus	131 (28.5%)	112 (27.2%)	19 (40.4%)	0.06
Hypertension	287 (62.5%)	255 (61.9%)	32 (68.1%)	0.41
Dyslipidemia	204 (44.4%)	183 (44.4%)	21 (44.7%)	0.97
Current or previous smoking	285 (63.1%)	255 (63.0%)	30 (63.8%)	0.91
COPD	38 (8.3%)	32 (7.8%)	6 (12.8%)	0.26
<b>History of CV disease, n (%)</b>				
Known CAD	107 (23.3%)	93 (22.6%)	14 (29.8%)	0.27

Chronic renal failure	96 (20.9%)	85 (20.6%)	11 (23.4%)	0.66
History of HF hospitalization	65 (14.2%)	58 (14.1%)	7 (14.9%)	0.88

**Clinical parameters at admission**

Mean arterial pressure, mmHg	98 ± 21	100 ± 20	84 ± 19	<b>&lt;0.001</b>
Heart rate, bpm	91 ± 27	90 ± 27	96 ± 28	0.17
Oxygen saturation, %	96.3 ± 3.6	96.4 ± 3.6	95.9 ± 3.5	0.22
Killip Class				
I	216 (47.1%)	194 (47.1%)	22 (46.8%)	0.97
II	169 (36.8%)	151 (36.7%)	18 (38.3%)	0.82
III	74 (16.1%)	67 (16.3%)	7 (14.9%)	0.81

**Laboratory results**

Hemoglobin, g/dl	12.9 ± 2.2	12.9 ± 2.2	12.4 ± 2.4	0.18
Creatinemia, µmol/l	118 ± 91	116 ± 91	128 ± 83	0.15
High-sensitivity cardiac troponin peak, U/l	354 ± 1906	335 ± 1979	516 ± 1079	<b>&lt;0.001</b>
NTproBNP, pg/ml	13960 ± 23843	13815 ± 24401	15233 ± 18396	0.145

### Echocardiography data

LA dilatation $\geq 32$ ml/m <sup>2</sup> , n (%)	144 (31.4%)	123 (29.9%)	21 (44.7%)	<b>0.038</b>
Mitral E/A ratio	1.32 $\pm$ 0.71	1.30 $\pm$ 0.72	1.49 $\pm$ 0.55	<b>&lt;0.001</b>
Mitral E/e' ratio	10.7 $\pm$ 4.2	10.3 $\pm$ 3.8	13.9 $\pm$ 5.6	<b>&lt;0.001</b>
LVEF, %	45 $\pm$ 16	46 $\pm$ 16	37 $\pm$ 17	<b>&lt;0.001</b>
LVEDV, ml/m <sup>2</sup>	124 $\pm$ 57	122 $\pm$ 56	135 $\pm$ 57	0.089
LVOT VTI, cm	17.7 $\pm$ 5.9	18.1 $\pm$ 5.7	14.2 $\pm$ 5.8	<b>&lt;0.001</b>
sPAP, mmHg	31 $\pm$ 15	31 $\pm$ 15	34 $\pm$ 14	<b>0.048</b>
RV dilatation, n (%)	63 (13.7%)	58 (14.1%)	5 (10.6%)	0.52
TAPSE, mm	19.2 $\pm$ 4.8	19.4 $\pm$ 4.7	16.6 $\pm$ 4.8	<b>&lt;0.001</b>

### Heart failure data

Worsening heart failure	95 (20.7%)	86 (20.9%)	9 (19.1%)	0.78
Ischemic etiology of heart failure	212 (46.2%)	180 (43.7%)	32 (68.1%)	<b>0.001</b>
Type of heart failure				<b>0.011</b>
HFrEF	155 (33.8%)	131 (31.8%)	24 (51.1%)	
HFmEF	87 (19.0%)	77 (18.7%)	10 (21.3%)	
HFpEF	217 (47.3%)	204 (49.5%)	13 (27.7%)	

**Previous treatment**

Aldosterone antagonists	44 (9.6%)	35 (8.5%)	9 (19.1%)	<b>0.032</b>
ACEI/ARB, or Entresto	219 (47.7%)	195 (47.3%)	24 (51.1%)	0.63
Diuretics	136 (29.6%)	117 (28.4%)	19 (40.4%)	0.09

**Environmental factors**

Alcohol consumption	230 (50.1%)	210 (51.0%)	20 (42.6%)	0.27
CO, ppm	4.5 ± 5.1	4.2 ± 4.4	7.8 ± 8.5	<b>0.026</b>
Recreational drug use, n (%)	42 (9.2%)	30 (7.3%)	12 (25.5%)	<b>&lt;0.001</b>

---

Values are n (%), mean ± SD, or median (interquartile range).

*Abbreviations:* ACEI/ARB, angiotensin-converting enzyme inhibitor/angiotensin receptor blocker ; CAD: coronary artery disease; CO: Carbone monoxide; COPD: chronic obstructive pulmonary disease; CV: cardiovascular ; HF: Heart Failure; HFrEF: Heart Failure reduced ejection fraction; HFmEF: Heart Failure mild ejection fraction; HFpEF: Heart Failure preserved ejection fraction; LA: Left Atrium ; LVEF: Left ventricular ejection fraction ; LVEDV: Left Ventricular end diastolic volume; LVOT VTI: left ventricular outflow tract velocity-time integral; MAE: major adverse events; NT-ProBNP: N-terminal prohormone of brain natriuretic peptide ; RV: right ventricular ;

**Table 2. Performance of some evaluated machine learning models for predicting in-hospital MAE.**

<b>Evaluation metrics</b>	<b>Logistic regression</b>	<b>LASSO</b>	<b>Random Forest</b>	<b>XGBoost</b>
AUROC	0.78	0.83	0.82	0.75
PRAUC	0.35	0.42	0.48	0.37
Accuracy	0.88	0.90	0.92	0.87
Cohen's Kappa	0.21	0.36	0.52	0.28
Sensitivity	0.36	0.36	0.50	0.36
Specificity	0.93	0.96	0.97	0.93
Precision	0.38	0.50	0.64	0.37
F1 score	0.37	0.42	0.56	0.36
Brier score	0.08	0.07	0.07	0.08

Evaluation of the performance of logistic regression, LASSO, Random Forest and XGBoost.

*Abbreviations: AUC: area under curve; ROC: receiver operating characteristic; LASSO: least absolute shrinkage and selection operator; XGBoost: Extreme gradient boosting.*

## **SUPPLEMENTAL MATERIAL**

**eSupplement 1.** Baseline and in-hospital characteristics collection.

**eSupplement 2.** Definition of main admission diagnosis.

**eSupplement 3.** Variables used in the feature selection algorithm.

**eSupplement 4.** Baseline characteristics of the training and the testing cohorts (N=459).

**eSupplement 5.** Evaluation of the Random Forest performances according to the gender.

**eSupplement 6.** Calibration curve for the ML-model.

**eReferences.**

---

## **eSupplement 1. Baseline and in-hospital characteristics collection**

### ***Baseline characteristics***

- Baseline data included the date of birth, gender, height, weight, temperature, systolic and diastolic blood pressures, heart rate, Glasgow score, Killip class, oxygen saturation and ventilation mode, presence of symptoms, list of medications at admission, history of cardiovascular disease, psychiatric illness or other significant clinical histories and main admission diagnosis.
- Regarding the drug addiction surveys, we assessed: i) declaration of psychoactive drug use (cannabinoids, cocaine, amphetamines, MDMA, heroin or other opioids); ii) smoking history (daily or nondaily smoker, cigarettes smoked per day by daily smokers, age started smoking, years smoked and e-cigarette use); iii) declaration of alcohol use.
- Of note, history of cardiovascular disease (CVD) was defined by the presence of: known MI, previous PCI, previous CABG, peripheral atheroma with revascularization, stroke, history of heart failure, history of atrial fibrillation, history of surgery for valvular heart disease, pacemaker or ICD, and cardiomyopathies.

### ***Hospital clinical characteristics***

- Electrocardiogram (ECG) and transthoracic echocardiography (TTE) with left ventricular ejection fraction (LVEF) were performed systematically within the first 24 hours of admission for all patients. Need for revascularization [percutaneous coronary intervention (PCI) or coronary artery bypass surgery (CABG)], ventricular arrhythmia (sustained ventricular tachycardia or fibrillation), cardiogenic shock, need for hemodynamic support, and urgent repeat revascularization were recorded.
- Laboratory results were also collected systematically upon admission, including hemoglobin, potassium, creatinine, the maximum peak of troponin (hsTNI), the N-terminal prohormone of B-type natriuretic peptide (NT-proBNP) or B-type natriuretic peptide (BNP).
- All the diagnostic procedures of cardiovascular imaging or invasive angiography reports were collected. All treatment introduced during hospitalization and the procedures performed were collected.
- The COVID status of each patient was systematically assessed at ICCU admission using RT-PCR, following current World Health Organization guidelines.

## **eSupplement 2. Definition of main admission diagnosis**

- The medical reasons for admission and the main admission diagnosis were adjudicated by a committee of two experts at the end of the hospitalization in each center.
- Main admission diagnosis was categorized into different subgroups: i) ST elevation myocardial infarction (STEMI), including mechanical complication of acute coronary syndrome; ii) non-ST elevation myocardial infarction (NSTEMI); iii) acute heart failure; iv) myocarditis; v) pericarditis; vi) pulmonary embolism; vii) atrial arrhythmia; viii) ventricular arrhythmia; ix) cardiac conduction abnormalities; x) coronary spasm; xi) Takotsubo; xii) aortic dissection; xiii) spontaneous coronary dissection, xiv) chest pain without identified cardiac cause, and xv) other cardiovascular or non-cardiovascular diagnosis.
- Of note, other cardiovascular diagnoses included: infectious endocarditis, acute hypertensive crisis without heart failure, prosthetic valve dysfunction without heart failure, vagal discomfort or orthostatic hypotension without severe cardiac conduction abnormality detected, and monitoring after electrocution.
- Of note, other non-cardiovascular diagnoses included: gastric ulcer, pancreatitis, acute cholecystitis, anemia, exacerbation of COPD, severe asthma, lung or systemic infection, severe COVID-19, migraine crisis, palpitations, or dyspnea without a cardiovascular diagnosis.

<b>Main admission diagnosis</b>	<b>Definition</b>
<b>Acute coronary syndrome</b>	<p>Acute coronary syndrome will be defined by typical angina of <math>\geq 20</math> min duration, ECG changes, and a rise in troponin or creatine kinase level above the 99<sup>th</sup> percentile of the upper reference limit after elimination of the differential diagnosis (myopericarditis, Takotsubo syndrome, Tachyarrhythmias, acute heart failure...).<sup>1</sup></p> <p>Acute coronary syndrome will be classified as ST-segment elevation and non ST-segment elevation categories.<sup>2</sup></p>
<b>Acute heart failure</b>	<p>An hospitalization for heart failure (HF) will be defined by symptoms and/or signs of HF with evidence of diastolic or systolic dysfunction by echocardiography and elevated levels of natriuretic peptide (BNP &gt;35 pg/ml and/or NT-proBNP &gt;125 pg/ml).<sup>3</sup></p>
<b>Myocarditis, pericarditis, Takotsubo syndrome</b>	<p>Myocarditis will be defined by chest pain, a rise in troponin or creatine kinase level above the 99 percentile of the upper reference limit and confirmation by cardiovascular magnetic resonance (CMR) using the Lake Louis criteria.<sup>4</sup></p> <p>Pericarditis will be defined when two out of the four following criteria are fulfilled: a) chest pain b) pericardial rubs c) ECG changes d) pericardial effusion.</p> <p>Takotsubo syndrome will be defined using the clinical expert consensus statement on takotsubo syndrome.<sup>5</sup></p>
<b>Pulmonary Embolism</b>	<p>Symptoms of pulmonary embolism (dyspnea, chest pain...) confirmed by imaging tests.<sup>6</sup></p>
<b>Acute supraventricular arrhythmias</b>	<p>Symptoms of tachycardia (palpitations, dyspnea...) leading to hospitalization in ICCU and 12-lead ECG confirming the supraventricular arrhythmia</p>
<b>Ventricular arrhythmias</b>	<p>Symptoms of tachycardia (palpitations, dyspnea...) leading to hospitalization in ICCU and 12-lead ECG confirming sustained ventricular tachycardia.</p>
<b>Other</b>	<p>The subgroup “other diagnosis” includes all diagnoses not eligible for the above categories, including aortic dissection, coronary spasm, unstable angina, endocarditis, hypertensive emergency, acute chest pain without etiology</p>

**eSupplement 3. Variables used in the feature selection algorithm.**

<i>Type of variable</i>	<i>Detailed features</i>
<b>Demographics</b>	Age, gender, BMI.
<b>CV risk factors</b>	Hypertension, dyslipidemia, diabetes, known CAD, family history of CAD.
<b>Medical history of non-CV disease</b>	Active cancer, psychiatric history
<b>Environmental factors</b>	Recreational drug use, alcohol consumption, CO level.
<b>Clinical parameters at admission</b>	Mean arterial pressure, heart rate, oxygen saturation, Killip class
<b>Biological parameters at admission</b>	Hemoglobin, creatininemia, troponin peak, NTproBNP
<b>Echocardiographic data</b>	LVEF, LA dilatation, RV dilatation, LV dilation, LV dysfunction, peak E/e' ratio, TAPSE.

*Abbreviations: BMI stands for Body Mass Index, CAD stands for Coronary Artery Disease, LA: Left Atrium, LV: Left Ventricular, LVEF: Left Ventricular Ejection Fraction.*

**eSupplement 4. Baseline characteristics of the training and the testing cohorts (N=459).**

	<b>All patients (N=459)</b>	<b>Training cohort (N=322)</b>	<b>Testing cohort (N=137)</b>	<b>p-value</b>
<b>Demographic data</b>				
Age, years	68 ± 14	68 ± 14	69 ± 13	0.36
Males, n (%)	311 (67.8%)	220 (68.3%)	91 (66.4%)	0.69
Body mass index, kg/m <sup>2</sup>	28 ± 6	28 ± 7	28 ± 6	0.82
<b>Cardiovascular risk factors, n (%)</b>				
Diabetes mellitus	131 (28.5%)	93 (28.9%)	38 (27.7%)	0.80
Hypertension	287 (62.5%)	197 (61.2%)	90 (65.7%)	0.36
Dyslipidemia	204 (44.4%)	144 (44.7%)	60 (43.8%)	0.85
Current or previous smoking	285 (63.1%)	196 (60.9%)	89 (65.0%)	0.62
COPD	38 (8.3%)	26 (8.1%)	12 (8.8%)	0.75
<b>History of CV disease, n (%)</b>				
Known CAD	107 (23.3%)	70 (21.7%)	37 (27.0%)	0.15
Renal failure	96 (20.9%)	63 (19.6%)	33 (24.1%)	0.27
History of HF hospitalization	65 (14.2%)	42 (13.0%)	23 (16.8%)	0.29

### Clinical parameters at admission

Mean arterial pressure, mmHg	98 ± 21	98 ± 21	98 ± 21	0.87
Heart rate, bpm	91 ± 27	92 ± 26	89 ± 30	0.19
Oxygen saturation, %	96.3 ± 3.6	96.2 ± 3.8	96.7 ± 3.2	0.21
Killip Class				
I	216 (47.1%)	155 (48.1%)	61 (44.5%)	0.47
II	169 (36.8%)	115 (35.7%)	54 (39.4%)	0.45
III	74 (16.1%)	52 (16.1%)	22 (16.1%)	0.98

### Biological parameters

Hemoglobin, g/dl	12.9 ± 2.2	12.9 ± 2.2	12.8 ± 2.1	0.60
Creatinemia, µmol/l	118 ± 91	115 ± 83	123 ± 107	0.44
High-sensitivity cardiac troponin peak, U/l	384 ± 1990	402 ± 2252	341 ± 1151	0.40
NTproBNP, pg/ml	13960 ± 23843	18549 ± 28083	12008 ± 21544	<b>0.034</b>

### Echocardiography data

LA dilatation ≥ 32 ml/m <sup>2</sup> , n (%)	144 (31.4%)	98 (30.4%)	46 (33.6%)	0.50
Mitral E/A ratio	1.38 ± 0.81	1.40 ± 0.82	1.32 ± 0.79	0.21
Mitral E/e' ratio	10.9 ± 4.6	11.0 ± 4.8	10.7 ± 4.2	0.82
LVEF, %	45 ± 16	46 ± 16	44 ± 16	0.31

LVEDV, ml/m <sup>2</sup>	127 ± 59	126 ± 58	131 ± 62	0.51
LVOT VTI, cm	18 ± 6	18 ± 6	18 ± 6	0.90
sPAP, mmHg	32 ± 16	32 ± 16	32 ± 16	0.70
RV dilatation, n (%)	63 (14.4%)	45 (14.5%)	18 (14.0%)	0.87
TAPSE, mm	19.1 ± 5.0	18.9 ± 5.2	19.7 ± 4.6	0.06
<b>Heart failure data</b>				
Worsening heart failure	95 (20.7%)	63 (19.6%)	32 (23.4%)	0.35
Ischemic etiology of heart failure	212 (46.2%)	141 (43.8%)	71 (51.8%)	0.114
Type of heart failure				<b>0.025</b>
HFrEF	155 (33.8%)	99 (30.7%)	56 (40.9%)	
HFmEF	87 (19.0%)	70 (21.7%)	17 (12.4%)	
HFpEF	217 (47.3%)	153 (47.5%)	64 (46.7%)	
<b>Previous treatment</b>				
Aldosterone antagonists	44 (9.6%)	30 (9.3%)	14 (10.2%)	0.76
ACEI/ARB, or Entresto	219 (47.7%)	152 (47.2%)	67 (48.9%)	0.73
Diuretics	136 (29.6%)	93 (28.9%)	43 (31.4%)	0.59
<b>Environmental factors</b>				
CO, ppm	4.7 ± 5.3	4.6 ± 5.1	4.9 ± 5.8	0.39

Alcohol consumption	223 (50.0%)	158 (50.3%)	65 (49.2%)	0.84
Recreational drug use, n (%)	42 (9.2%)	36 (11.2%)	6 (4.4%)	<b>0.021</b>
<b>Outcome, n (%)</b>				
In-hospital MAE	47 (10.2%)	33 (10.2%)	14 (10.2%)	0.99

---

*Abbreviations as in Table 1.*

**eSupplement 5. Evaluation of the Random Forest performances according to the gender.**

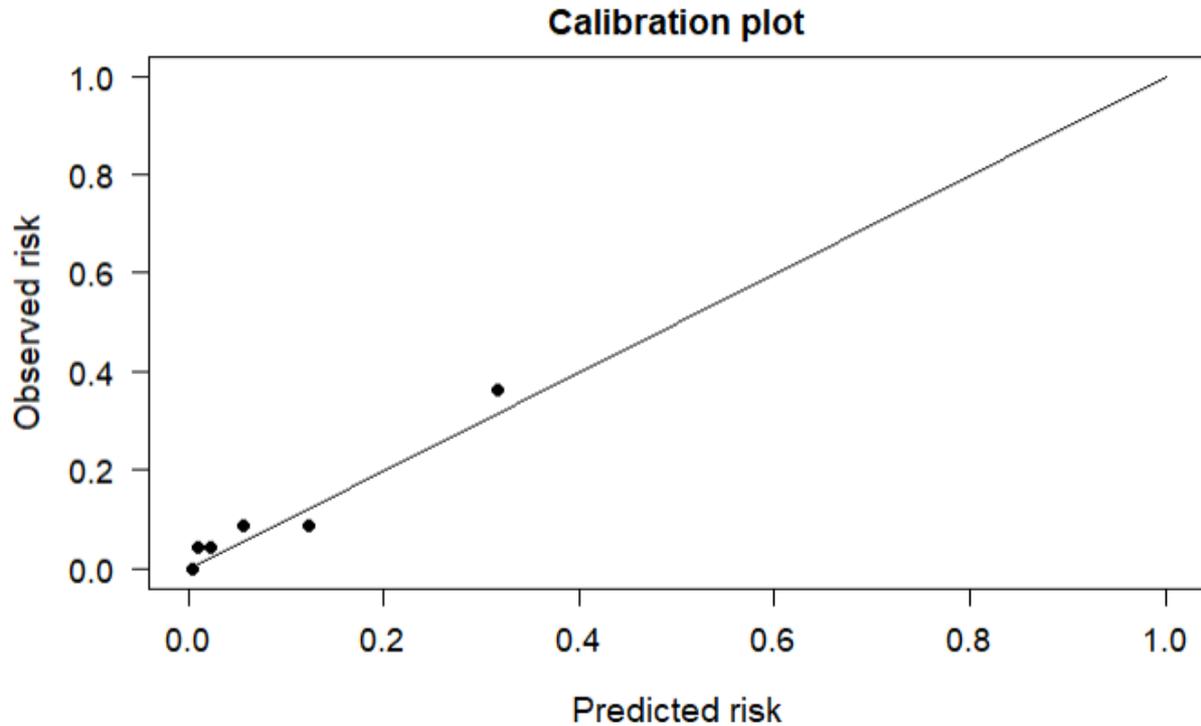
---

<b>Evaluation metrics</b>	<b>Males</b>	<b>Females</b>
ROC AUC	0.82	0.83
Precision-recall AUC	0.35	0.22
Accuracy	0.93	0.85
Cohen's Kappa	0.63	0.38
Sensitivity	0.67	0.60
Specificity	0.96	0.88
Precision	0.67	0.38
F1 score	0.67	0.46
Brier score	0.07	0.07

---

*Abbreviations: AUC: area under curve; ROC: receiver operating characteristic; LASSO: least absolute shrinkage and selection operator; XGBoost: Extreme gradient boosting*

**eSupplement 6.**



The calibration plot shows the relationship between the observed and predicted proportion of events, grouped by level of risk. The Random Forest model showed good calibration with the observed in-hospital MAEs. Of note, the Hosmer and Lemeshow goodness-of-fit test showed no significant difference between the proportions of observed and expected predicted in-hospital MAEs (p-value=0.91).

*MAE: major adverse events*

## eReferences

1. Collet JP, Thiele H, Barbato E, et al. 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. *Eur Heart J*. 2021;42(14):1289-1367. doi:10.1093/eurheartj/ehaa575
2. Ibanez B, James S, Agewall S, et al. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: The Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC). *Eur Heart J*. 2018;39(2):119-177. doi:10.1093/eurheartj/ehx393
3. McDonagh TA, Metra M, Adamo M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J*. 2021:ehab368. doi:10.1093/eurheartj/ehab368
4. Friedrich MG, Sechtem U, Schulz-Menger J, et al. Cardiovascular magnetic resonance in myocarditis: A JACC White Paper. *J Am Coll Cardiol*. 2009;53(17):1475-1487. doi:10.1016/j.jacc.2009.02.007
5. Ghadri JR, Wittstein IS, Prasad A, et al. International Expert Consensus Document on Takotsubo Syndrome (Part II): Diagnostic Workup, Outcome, and Management. *Eur Heart J*. 2018;39(22):2047-2062. doi:10.1093/eurheartj/ehy077
6. Konstantinides SV, Meyer G, Becattini C, et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). *Eur Heart J*. 2020;41(4):543-603. doi:10.1093/eurheartj/ehz405

## Partie 3 : DISCUSSION

---

### 1. Variables environnementales

Parmi les variables de notre modèle pronostique, **deux facteurs environnementaux** ont été sélectionnés par le ML. Il est important de noter que le niveau de CO exhalé était la **3<sup>ème</sup> variable d'importance** dans le modèle. Bien que ce résultat soit **inédit**, plusieurs études, de manière indirecte, retrouvent des résultats insistant sur l'**association entre survenue d'évènements cardiovasculaires et pollution atmosphérique**. Notamment une étude portant sur des patients de la cohorte de Framingham a montré une association entre le taux de CO exhalé et la survenue d'évènements cardiovasculaires, **indépendamment du statut tabagique du patient**.(20) Ce résultat suggère que le taux de CO exhalé pourrait être un marqueur intéressant pour évaluer l'impact cardiovasculaire de la pollution de l'air.(50) De surcroît, une importante méta-analyse a montré une forte association entre les augmentations quotidiennes des **polluants atmosphériques gazeux** (particules, monoxyde de carbone, dioxyde de soufre, dioxyde d'azote, ozone) et la survenue d'**affections cardio-vasculaires sévères**.(28) Plus récemment, une étude a montré qu'un **niveau élevé de CO expiré** est indépendamment associé à une multiplication par 10 des **évènements cardiovasculaires graves intra-hospitaliers** et à une **multiplication par 6 de la mortalité** à un an chez les patients hospitalisés pour des évènements cardiaques aigus.(51)

Le deuxième facteur environnemental sélectionné par notre modèle était la consommation de **drogues récréatives**. Ce résultat peut s'expliquer chez les patients en insuffisance cardiaque aigüe notamment par l'hypertonie adrénergique qui, comme décrit dans l'introduction, augmente la fréquence cardiaque, la pression artérielle, la température et par conséquent la demande en oxygène du myocarde.(52) Ce résultat est conforme à une analyse antérieure de l'étude ADDICTO-USIC montrant que la consommation de drogues récréatives était fortement associée à la survenue d'évènements cardiovasculaires graves intra-hospitaliers. (53)

## 2. Variables cliniques

Notre modèle a sélectionné aussi **deux variables cliniques** : la **pression artérielle moyenne** et **l'origine ischémique** de l'insuffisance cardiaque. Un point important à souligner est la **facilité d'accès** à ces données pour le clinicien qui évalue le patient aux soins intensifs. L'étiologie ischémique comme facteur prédictif de notre modèle a déjà été mis en évidence dans un score publié par une équipe espagnole (score MEESSEI-HF)(43,41). L'hypothèse physiopathologique sous-jacente est que ces patients sont plus sujets aux troubles du rythme ventriculaires graves, notamment en raison de la **fibrose cicatricielle pro-arythmogène**.(54) Quant à la pression artérielle moyenne, c'est une donnée qui fait **partie intégrante de tout examen clinique** en médecine, tant sa valeur est riche d'information pour le clinicien. Pression artérielle moyenne et systolique sont reliées et cette dernière est présente dans la plupart des scores prédictif de sévérité chez le patient insuffisant cardiaque.(38,40) Il est bien précisé dans notre travail, que les patients **admis pour un état de choc cardiogénique pré-hospitalier** étaient exclus de l'analyse.

## 3. Variables échographiques

Enfin, notre modèle a sélectionné **trois variables échocardiographiques** : l'ITV sous aortique, le rapport E/e' et le TAPSE. Ces trois variables sont particulièrement d'intérêt pour les raisons suivantes :

-Avant toute chose, ce sont des **paramètres de routine, facilement accessibles** pour le cardiologue et qui sont **inscrits dans les recommandations européennes**. Ce travail souligne l'intérêt de l'échographie cardiaque dans l'évaluation du patient en insuffisance cardiaque aigüe.

-De plus, ces trois paramètres offrent ensemble un **aperçu global de la fonction cardiaque** avec une approche du débit sous aortique, l'estimation des pressions de remplissage et la fonction longitudinale du ventricule droit. **Ces résultats sont soutenus par la littérature**. En effet, plusieurs études ont montré qu'une baisse du débit cardiaque chez l'insuffisant cardiaque était associée à un moins bon pronostic. (6,55) Concernant le TAPSE, plusieurs études ont montré qu'une baisse de la fonction ventriculaire droite utilisant ce paramètre étaient associés aussi à un moins bon pronostic en présence et en l'absence de syndrome coronarien aigu.(56,57) Des résultats similaires ont été publiés sur le rapport E/e' dans l'insuffisance cardiaque aigüe. (6,58,59)

#### 4. Utilisation du machine learning

Notre modèle construit à l'aide d'algorithmes de ML a montré de **meilleures performances** que le modèle de régression logistique dans la prédiction d'événements cardiovasculaires graves intra-hospitaliers. Bien que les résultats en termes de ROC-AUC étaient similaires entre le random forest, le LASSO et la régression logistique, **la différence était plus importante pour les PR-AUC**. En effet, nos modèles ont dû être développés sur des **données déséquilibrées** ce qui peut prendre à défaut certains paramètres d'évaluation de performance comme la ROC-AUC ; sa valeur est influencée par le taux de vrais négatifs qui est ici la classe majoritaire. La PR-AUC quant à elle se concentre majoritairement sur les cas positifs ce qui est plus adaptée dans le contexte de notre travail. Enfin, notre modèle a montré de **meilleures performances comparées** à celle d'un score validé déjà publié dans la littérature.

#### 5. Limites

Notre travail présente cependant certaines limites.

Premièrement, sur le plan statistique, la **taille de notre échantillon** a dû être réduite pour l'élaboration de la population d'entraînement et de validation ce qui a pu affecter les performances d'algorithmes de machine learning. Cette condition est néanmoins primordiale dans l'élaboration d'un modèle prédictif afin de limiter le phénomène de **surapprentissage** ou *overfitting* qui risque de faire apprendre à notre modèle non plus l'importance des variables en elles-mêmes mais le « bruit » (ou structure de la population) **limitant ainsi l'application à une autre population** et donc la **généralisabilité** des résultats.

Deuxièmement, bien que les données aient été récoltées de manière prospective, nous avons été confrontés à la **gestion des données manquantes**. Ici la méthode des plus proches voisins a été utilisée après avoir exclu les variables pour lesquelles le taux de données manquante était supérieur à 40%. (56,59,60) Le rapport E/e' avait 25% de données manquantes et a été retenu dans notre modèle. Aucun MAE n'a cependant été imputé.

Troisièmement, le seul score validé dans l'insuffisance cardiaque aiguë avec lequel notre modèle a pu être comparé est l'ACUTE-HF score qui, comme nous l'avons vu dans l'introduction, est

un score qui mesure la mortalité à 30 jours et non pas la mortalité intra-hospitalière.(61) Ces résultats doivent donc être pris avec précautions.

Enfin, notre modèle intègre des données environnementales comme le CO expiré et un dépistage urinaire des drogues récréatives qui **ne sont pas inscrites à ce jour dans la pratique médicale courante** ce qui limite l'applicabilité de notre modèle.

## Conclusion

---

L'IC aigüe est un des **principaux motifs d'admission** en France et représente un réel **enjeu de santé publique**. L'importance du taux d'évènements intra-hospitaliers soulignent la nécessité pour les cliniciens de disposer **d'outils performants** pour cibler les patients à haut risque de **complications**.

Portée par les progrès technologiques, **l'intelligence artificielle** ouvre de nouvelles perspectives dans le domaine de la santé. Cet outil, en dépassant la rigidité intrinsèque des statistiques traditionnelles, permet d'intégrer de nouvelles données à l'image des **facteurs environnementaux**, reconnus pour avoir un impact cardiovasculaire et **pourtant peu étudiés**.

Notre modèle prédictif de survenue d'évènements cardiovasculaires graves intra-hospitaliers développé à partir de méthodes de ML a permis de sélectionner sept variables d'intérêt : **l'ITV sous aortique**, le **rapport E/e'**, le **TAPSE**, la **pression artérielle moyenne**, **l'origine ischémique de l'insuffisance cardiaque** et des données environnementales (le **CO expiré** et les **drogues récréatives**).

En plus de surpasser les autres modèles de ML, le Random Forest a montré de **meilleures performances** que la régression logistique. Notre modèle n'a cependant pas pu être étudié dans une population de validation externe en **l'absence de données environnementales disponibles** au sein d'autres cohortes d'insuffisants cardiaques.

D'autres études sont alors **nécessaires** pour explorer les associations décrites dans ce travail avant d'intégrer ces données dans la pratique courante et ainsi **mieux sélectionner** les patients les plus à risque dans l'espoir **d'améliorer leur pronostic**.

## Bibliographie

---

1. Principales causes de décès et de morbidite.pdf [Internet]. [cité 31 août 2022]. Disponible sur: <https://drees.solidarites-sante.gouv.fr/sites/default/files/2021-01/Principales%20causes%20de%20d%C3%A9c%C3%A8s%20et%20de%20morbidite.pdf>
2. Géodes - Santé publique France - Indicateurs : cartes, données et graphiques [Internet]. 2022 [cité 31 août 2022]. Disponible sur: [https://geodes.santepubliquefrance.fr/#c=indicator&f=0&i=cardio\\_morta\\_ic0&s=2016&t=a01&view=map2](https://geodes.santepubliquefrance.fr/#c=indicator&f=0&i=cardio_morta_ic0&s=2016&t=a01&view=map2)
3. Cardiovascular diseases (CVDs) [Internet]. 2022 [cité 31 août 2022]. Disponible sur: [https://www.who.int/fr/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/fr/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
4. Insuffisance cardiaque : définition et causes [Internet]. 2022 [cité 31 août 2022]. Disponible sur: <https://www.ameli.fr/paris/assure/sante/themes/insuffisance-cardiaque/definition-causes>
5. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J.* 14 juill 2016;37(27):2129-200.
6. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J.* 21 sept 2021;42(36):3599-726.
7. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2023 Focused Update of the 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the task force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J.* 1 oct 2023;44(37):3627-39.
8. L'insuffisance cardiaque [Internet]. FFC. 2022 [cité 31 août 2022]. Disponible sur: <https://fedecardio.org/je-m-informe/l-insuffisance-cardiaque/>
9. Mentz RJ, O'Connor CM. Pathophysiology and clinical evaluation of acute heart failure. *Nat Rev Cardiol.* janv 2016;13(1):28-35.
10. Cours [Internet]. 2022 [cité 31 août 2022]. Disponible sur: [http://campus.cerimes.fr/cardiologie-et-maladies-vasculaires/enseignement/cardio\\_250/site/html/4.html](http://campus.cerimes.fr/cardiologie-et-maladies-vasculaires/enseignement/cardio_250/site/html/4.html)
11. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in-Congestive Heart Failure (MERIT-HF). *The Lancet.* 12 juin 1999;353(9169):2001-7.

12. Jourdain P, Lefèvre G, Oddoze C, Sapin V, Dievart F, Jondeau G, et al. NT-proBNP en pratique « De la biologie à la clinique ». *Ann Biol Clin (Paris)*. 1 mai 2009;67(3):255-71.
13. Kotecha D, Manzano L, Krum H, Rosano G, Holmes J, Altman DG, et al. Effect of age and sex on efficacy and tolerability of  $\beta$  blockers in patients with heart failure with reduced ejection fraction: individual patient data meta-analysis. *BMJ*. 20 avr 2016;353:i1855.
14. Anker SD, Butler J, Packer M. Empagliflozin in Heart Failure with a Preserved Ejection Fraction. Reply. *N Engl J Med*. 26 mai 2022;386(21):e57.
15. Solomon SD, de Boer RA, DeMets D, Hernandez AF, Inzucchi SE, Kosiborod MN, et al. Dapagliflozin in heart failure with preserved and mildly reduced ejection fraction: rationale and design of the DELIVER trial. *Eur J Heart Fail*. 2021;23(7):1217-25.
16. Stehlik J, Estep JD, Selzman CH, Rogers JG, Spertus JA, Shah KB, et al. Patient-Reported Health-Related Quality of Life Is a Predictor of Outcomes in Ambulatory Heart Failure Patients Treated With Left Ventricular Assist Device Compared With Medical Management. *Circ Heart Fail*. juin 2017;10(6):e003910.
17. Assistance Publique - Hôpitaux de Paris. Assessment of ADDICTION to Psychoactive Drug Use in Intensive Cardiac Care Units - The ADDICT-CCU Study [Internet]. [clinicaltrials.gov](https://clinicaltrials.gov); 2022 mars [cité 29 août 2022]. Report No.: NCT05063097. Disponible sur: <https://clinicaltrials.gov/ct2/show/NCT05063097>
18. intracto. Centre Antipoisons Belge. 2022 [cité 31 août 2022]. Quelles sont les concentrations toxiques de CO? Disponible sur: <https://www.centreantipoisons.be/monoxyde-de-carbone/le-monoxyde-de-carbone-co-en-d-tail/quelles-sont-les-concentrations-toxiques-de>
19. Curran L, Nah G, Marcus GM, Tseng Z, Crawford MH, Parikh NI. Clinical Correlates and Outcomes of Methamphetamine-Associated Cardiovascular Diseases in Hospitalized Patients in California. *J Am Heart Assoc*. 2022;11(16):e023663.
20. Nishimura M, Bhatia H, Ma J, Dickson SD, Alshawabkeh L, Adler E, et al. The Impact of Substance Abuse on Heart Failure Hospitalizations. *Am J Med*. 2020;133(2):207-213.e1.
21. Varga ZV, Ferdinandy P, Liaudet L, Pacher P. Drug-induced mitochondrial dysfunction and cardiotoxicity. *Am J Physiol Heart Circ Physiol*. 2015;309(9):H1453-1467.
22. Thyagaturu HS, Bolton AR, Li S, Gonuguntla K, Kumar A, Bianco C, et al. Effect of Cocaine, Amphetamine, and Cannabis Use Disorders on 30-day Readmissions of Patients with Heart Failure. *Curr Probl Cardiol*. 2022;101189.
23. Schwartz BG, Rezkalla S, Kloner RA. Cardiovascular effects of cocaine. *Circulation*. 2010;122(24):2558-69.
24. Page RL, Allen LA, Kloner RA, Carriker CR, Martel C, Morris AA, et al. Medical Marijuana, Recreational Cannabis, and Cardiovascular Health: A Scientific Statement From the American Heart Association. *Circulation*. 2020;142(10):e131-52.

25. Auer R, Sidney S, Goff D, Vittinghoff E, Pletcher MJ, Allen NB, et al. Lifetime marijuana use and subclinical atherosclerosis: the Coronary Artery Risk Development in Young Adults (CARDIA) study. *Addiction*. 2018;113(5):845-56.
26. Chami T, Kim CH. Cannabis Abuse and Elevated Risk of Myocardial Infarction in the Young: A Population-Based Study. *Mayo Clin Proc*. août 2019;94(8):1647-9.
27. Deborde T, Brisacier AC. Usage de cannabis et complications cardio-vasculaires et neuro-vasculaires. *Psychotropes*. 2016;22(3-4):157-83.
28. Shah AS, Langrish JP, Nair H, McAllister DA, Hunter AL, Donaldson K, et al. Global association of air pollution and heart failure: a systematic review and meta-analysis. *The Lancet*. 2013;382(9897):1039-48.
29. Tun B, Ehrbar R, Short M, Cheng S, Vasani RS, Xanthakis V. Association of Exhaled Carbon Monoxide With Ideal Cardiovascular Health, Circulating Biomarkers, and Incidence of Heart Failure in the Framingham Offspring Study. *J Am Heart Assoc*. 2020;9(21):e016762.
30. Reboul C, Boissière J, André L, Meyer G, Bideaux P, Fouret G, et al. Carbon monoxide pollution aggravates ischemic heart failure through oxidative stress pathway. *Sci Rep*. 2017;7(1):39715.
31. Lampropoulos CE, Sklavou E, Anastogiannis C, Papanikolaou V, Tsilivarakis D, Vasilioi V, et al. Acute Heart Failure in a Healthy Young Patient after Severe Carbon Monoxide Poisoning. *Eur J Case Rep Intern Med*. 13 déc 2019;6(12):001340.
32. JVC J. Data Transition Numérique. 2022 [cité 31 août 2022]. Introduction au Machine Learning avec Python. Disponible sur: <https://www.data-transitionnumerique.com/machine-learning-python/>
33. Moons KGM, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart*. 2012;98(9):683-90.
34. Wussler D, Michou E, Belkin M, Kozhuharov N, Diebold M, Gualandro DM, et al. Mortality prediction in acute heart failure: scores or biomarkers? *Swiss Med Wkly* [Internet]. 13 août 2020 [cité 29 août 2022];(33). Disponible sur: <https://smw.ch/article/doi/smw.2020.20320>
35. Horne BD, Roberts CA, Rasmusson KD, Buckway J, Alharethi R, Cruz J, et al. Risk score-guided multidisciplinary team-based Care for Heart Failure Inpatients is associated with lower 30-day readmission and lower 30-day mortality. *Am Heart J*. 1 janv 2020;219:78-88.
36. Engelsingjerd EK, Benziger CP, Horne BD. Validation of the Intermountain Risk Score and Get with the Guidelines—Heart Failure Score in predicting mortality. *Open Heart*. 1 août 2021;8(2):e001722.
37. Evans RS, Benezillo J, Horne BD, Lloyd JF, Bradshaw A, Budge D, et al. Automated identification and predictive tools to help identify high-risk heart failure patients: pilot

evaluation. *J Am Med Inform Assoc.* 2016;23(5):872-8.

38. Lee DS, Stitt A, Austin PC, Stukel TA, Schull MJ, Chong A, et al. Prediction of Heart Failure Mortality in Emergent Care: A Cohort Study. *Ann Intern Med.* 2012;156(11):767.

39. Cameli M, Pastore MC, Carli GD, Henein MY, Mandoli GE, Lisi E, et al. ACUTE HF score, a multiparametric prognostic tool for acute heart failure: A real-life study. *Int J Cardiol.* 2019;296:103-8.

40. Peterson PN, Rumsfeld JS, Liang L, Albert NM, Hernandez AF, Peterson ED, et al. A Validated Risk Score for In-Hospital Mortality in Patients With Heart Failure From the American Heart Association Get With the Guidelines Program. *Circ Cardiovasc Qual Outcomes.* 2010;3(1):25-32.

41. Stiell IG, Clement CM, Brison RJ, Rowe BH, Borgundvaag B, Aaron SD, et al. A risk scoring system to identify emergency department patients with heart failure at high risk for serious adverse events. *Acad Emerg Med Off J Soc Acad Emerg Med.* 2013;20(1):17-26.

42. Collins SP, Jenkins CA, Harrell FE, Liu D, Miller KF, Lindsell CJ, et al. Identification of Emergency Department Patients With Acute Heart Failure at Low Risk for 30-Day Adverse Events. *JACC Heart Fail.* 2015;3(10):737-47.

43. Miró Ò, Rossello X, Gil V, Martín-Sánchez FJ, Llorens P, Herrero-Puente P, et al. Predicting 30-Day Mortality for Patients With Acute Heart Failure in the Emergency Department: A Cohort Study. *Ann Intern Med.* 2017;167(10):698-705.

44. Gök G, Karadağ M, Sinan ÜY, Zoghi M. A New Risk Score to Predict In-Hospital Mortality in Elderly Patients With Acute Heart Failure: On Behalf of the Journey HF-TR Study Investigators. *Angiology.* 2020;71(10):948-54.

45. Yap J, Lim FY, Chia SY, Allen JC, Jaufeerally FR, Macdonald MR, et al. Prediction of Survival in Asian Patients Hospitalized With Heart Failure: Validation of the OPTIMIZE-HF Risk Score. *J Card Fail.* 2019;25(7):571-5.

46. Austin DE, Lee DS, Wang CX, Ma S, Wang X, Porter J, et al. Comparison of machine learning and the regression-based EHMRG model for predicting early mortality in acute heart failure. *Int J Cardiol.* 2022;365:78-84.

47. Li F, Xin H, Zhang J, Fu M, Zhou J, Lian Z. Prediction model of in-hospital mortality in intensive care unit patients with heart failure: machine learning-based, retrospective analysis of the MIMIC-III database. *BMJ Open.* 2021;11(7):e044779.

48. Luo C, Zhu Y, Zhu Z, Li R, Chen G, Wang Z. A machine learning-based risk stratification tool for in-hospital mortality of intensive care unit patients with heart failure. *J Transl Med.* 2022;20(1):136.

49. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci Data.* 2018;5:180178.

50. Cheng S, Enserro D, Xanthakis V, Sullivan LM, Murabito JM, Benjamin EJ, et al. Association of exhaled carbon monoxide with subclinical cardiovascular disease and their conjoint impact on the incidence of cardiovascular outcomes. *Eur Heart J*. 2014;35(42):2980-7.
51. Dillinger JG, Pezel T, Delmas C, Schurtz G, Trimaille A, Piliero N, et al. Carbon monoxide and prognosis in smokers hospitalised with acute cardiac events: a multicentre, prospective cohort study. *eClinicalMedicine*. 3 janv 2024;67:102401.
52. Ghuran A, van der Wieken LR, Nolan J. Cardiovascular complications of recreational drugs. *BMJ*. 1 sept 2001;323(7311):464-6.
53. 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*. 1 nov 2023;109(21):1608-16.
54. Albert NM. Ventricular Dysrhythmias in Heart Failure. *J Cardiovasc Nurs*. déc 2004;19(6):S11.
55. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2023 Focused Update of the 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J*. 2023;44(37):3627-39.
56. Zehender Manfred, Kasper Wolfgang, Kauder Elisabeth, Schonhaler Martin, Geibel Annette, Olschewski Manfred, et al. Right Ventricular Infarction as an Independent Predictor of Prognosis after Acute Inferior Myocardial Infarction. *N Engl J Med*. 1993;328(14):981-8.
57. Berrill M, Ashcroft E, Fluck D, John I, Beeton I, Sharma P, et al. Right Ventricular Dysfunction Predicts Outcome in Acute Heart Failure. *Front Cardiovasc Med*. 2022;9:911053.
58. Lombardi C, Peveri G, Cani D, Latta F, Bonelli A, Tomasoni D, et al. In-hospital and long-term mortality for acute heart failure: analysis at the time of admission to the emergency department. *ESC Heart Fail*. 2020;7(5):2650-61.
59. 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. *J Am Heart Assoc*. 2019;8(17):e013675.
60. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):162.

## Table des figures

---

<b>Figure 1</b> : Représentation de l'appareil cardiovasculaire. D'après <i>Chung,MK., and Rich, M.W. Introduction au système cardiovasculaire.</i> .....	15
<b>Figure 2</b> : Arbre diagnostique de l'IC chronique (traduit des recommandations ESC 2021). 20	
<b>Figure 3</b> : Prise en charge des patients en insuffisance cardiaque aigue (traduit des recommandations ESC 2021).....	22
<b>Figure 4</b> : Prise en charge de l'insuffisance cardiaque chronique à FEVG réduite (traduit des recommandations ESC 2021).....	23
<b>Figure 5</b> : Kit de dosage urinaire utilise pour l'étude ADDICTO-USIC (NarcoCheck®, Kappa City Biotech SAS, Montluçon, France) .....	26
<b>Figure 6</b> : courbe de dissociation de l'hémoglobine en fonction de son affinité. ....	27
<b>Figure 7</b> : Figure adaptée d'une figure tirée de l'étude : effets aigus de la cocaïne.....	29
<b>Figure 8</b> : Schéma synthétiques des trois principales formes de ML.....	32
<b>Figure 9</b> : Fonctionnement et étapes du ML supervisé(32).....	33
<b>Figure 10</b> : méthode des k plus proches voisins.....	35

## Table des tableaux

---

**Tableau 1:** Définition de l'insuffisance cardiaque à FEVG préservée, à FEVG modérément réduite et à FEVG réduite adapté selon les recommandations ESC 2021(10)..... 18

**Tableau 2 :** récapitulatif des variables des scores. .... 41

## Annexes

---

**Soumission de l'article à l'*European Journal of Heart Failure* le 05/05/2024**

## **Modèle pronostique de machine learning supervisé, incluant des facteurs environnementaux, pour prédire la sévérité intra-hospitalière chez les patients admis pour insuffisance cardiaque aiguë.**

### **Résumé :**

**Objectif:** Étudier la faisabilité et la robustesse d'un modèle de machine learning supervisé incluant des facteurs environnementaux pour prédire les événements indésirables majeurs intra-hospitaliers chez les patients hospitalisés pour insuffisance cardiaque aiguë aux soins intensifs.

**Méthode:** Nous avons sélectionné les patients admis pour insuffisance cardiaque aiguë issus du registre ADDICTO-USIC, une étude multicentrique prospective nationale française qui a inclus tous les patients consécutifs hospitalisés aux soins intensifs de cardiologie. Le critère de jugement principal était un critère composite de survenu d'événements majeurs intra-hospitaliers définis comme le décès, l'arrêt cardiaque récupéré ou le choc cardiogénique nécessitant une assistance médicamenteuse ou circulatoire. Plusieurs algorithmes de ML ont été entraînés puis comparés à un outil de statistique traditionnelle pour sélectionner les variables prédictives et construire notre modèle dans un échantillon d'entraînement. Les performances ont ensuite été validées dans un échantillon de validation et comparés à l'ACUTE-HF score.

**Résultats:** Parmi les 459 patients inclus (âge  $68 \pm 14$  ans, 68 % d'hommes), 47 soit 10,2% de la cohorte ont présenté un événement majeur intra-hospitalier. Sept variables ont été sélectionnées par la régression de LASSO pour prédire l'événement dans le *train-set* (N=322) : la pression artérielle moyenne, l'étiologie ischémique, l'ITV sous-aortique, le rapport E/e', le TAPSE, l'utilisation de drogues récréatives et le taux de CO expiré. Le Random forest a présenté les meilleures performances (AUROC=0,82, PR-AUC=0,48, F1-score=0,56). Enfin, notre modèle a été comparé à l'ACUTE-HF score et a montré une meilleure ROC-AUC (random forest : 0,82 vs Acute HF-score : 0,57 ;  $p < 0,001$ ).

**Conclusion:** Notre modèle de ML, incluant sept variables dont deux variables environnementales, a montré les meilleures performances pour prédire les événements majeurs intra-hospitaliers chez les patients admis pour insuffisance cardiaque aiguë.

### **Spécialité :**

[Médecine cardiovasculaire](#)

### **Mots clés français :**

Insuffisance cardiaque aiguë ; facteurs environnementaux ; monoxyde de carbone; drogues récréatives; machine learning; soins intensifs de cardiologie ; échocardiographie.

### **Forme ou Genre :**

[fMeSH : Dissertation universitaire](#)

[Rameau : Thèses et écrits académiques](#)

**Université Paris Cité  
UFR de médecine  
15 Rue de l'École de Médecine  
75006 Paris**