

Intelligence artificielle et ECG : revue critique

*Jean-Marie Grégoire , Cédric Gilon , Hugues Bersini ,
Stéphane Carlier*

Les modèles d'intelligence artificielle peuvent maintenant être utilisés non seulement pour l'interprétation des ECG, mais également pour le dépistage et la prévision de diverses pathologies cardiaques grâce à leur capacité d'analyse en profondeur des données.

Introduction

L'intelligence artificielle (IA) a remis l'ECG à l'ordre du jour. Actuellement de nombreuses applications d'IA se développent et certains modèles montrent des résultats qui égalent l'analyse faite par les cardiologues pour les interprétations de base. Plus largement, l'IA est devenue la méthode de pointe à appliquer à toutes les tâches : traitement des images obtenues par RX, scanners et IRM, mais aussi pour l'analyse des dossiers médicaux et la connexion entre les bases de données médicales, des organismes de soins de santé et des assurances. Cet article se limite à l'apport des techniques d'IA à l'ECG, qui reste l'outil cardiologique de base.

Les possibilités d'analyse sont très éloignées de ce qui pouvait être obtenu à partir des appareils analogiques des années 1960. Il faut rappeler qu'on n'a pas attendu l'IA pour réaliser des protocoles automatisés d'ECGs : ce processus a commencé dès la transformation des appareils ECG analogiques en numériques dans les années 1950. Une excellente revue des débuts des ECG numérisés a été réalisée par Mac Farlane¹. La performance des premiers programmes dans le diagnostic de l'infarctus aigu du myocarde se situait entre 0 et 42 % de faux positifs et 22 et 30 % de faux négatifs². Plus récemment, un réseau de neurones à 34 couches créé pour diagnostiquer une occlusion coronaire aiguë en salle de coronarographie ne présente pas de bien meilleur résultat³. L'oeil de l'expert

reste donc indispensable. Il faut toutefois reconnaître que l'IA devient extrêmement performante dans le diagnostic des troubles du rythme, des infarctus STEMI et de l'insuffisance cardiaque.

L'apprentissage automatique, *machine learning* (ML), est un sous-domaine de l'IA dans lequel le modèle n'est pas explicitement programmé pour suivre un ensemble d'instructions afin de résoudre une tâche donnée. En effet, le modèle va apprendre comment agir en acquérant son propre ensemble de règles à partir de ses expériences (figure 1). Le terme 'apprentissage' est toutefois quelque peu trompeur, car les machines n'apprennent pas comme les humains : l'apprentissage automatique est dans de nombreux cas une application des statistiques. Il existe de nombreuses relations entre les algorithmes statistiques et de ML, tels que l'analyse discriminante, la régression logistique et d'autres modèles linéaires. La modélisation statistique suppose l'existence d'un modèle probabiliste générant les données (la culture de la modélisation des données); l'IA se concentre sur la mise en correspondance des entrées et des sorties à travers un modèle (la culture de la modélisation algorithmique⁴). La première approche commence par une question et/ou des données, tandis que la seconde correspond à une approche centrée sur les tâches (tableau 1). Le ML propose un ensemble d'algorithmes qui peuvent prendre un ensemble de données en entrée, *input*, et produire un résultat, *output*.

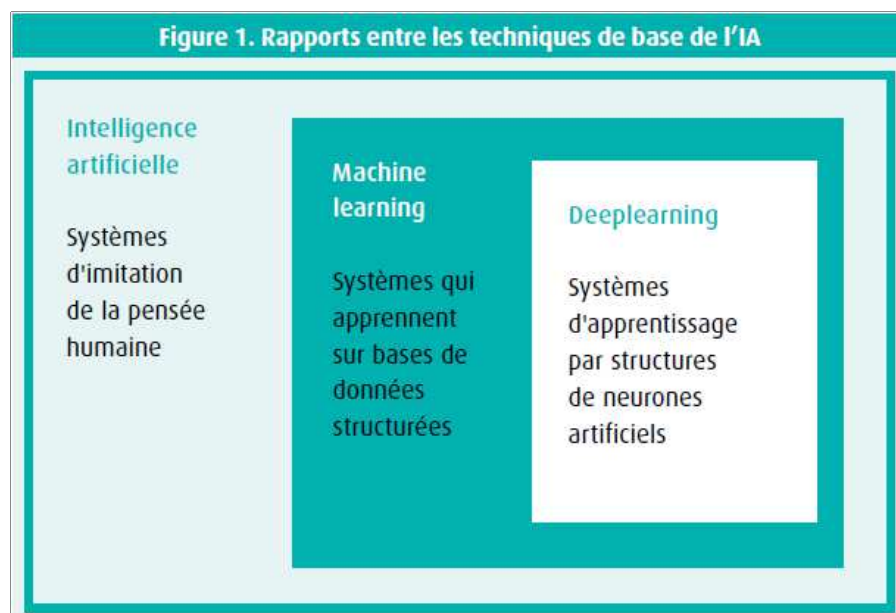


Tableau 1. Statistiques : inférences sur la population à partir d'un échantillon - machine learning : modèles prédictifs généralisables	
Machine learning	Statistiques
Prédictions	Inférences
Erreurs d'apprentissage possibles pas de problème avec la complexité	Préférence pour les modèles simples même si les modèles complexes performent mieux
Emphase sur la performance	Emphase sur l'interprétabilité
Généralisation sur des nouvelles données	Inférences liées à la population d'intérêt
Le souci : performance et robustesse	Le souci : inférences et robustesse
Résultats via la performance	Hypothèses posées a priori

La diversité des techniques d'enregistrement ECG est devenue pléthorique et il est parfois difficile de s'y retrouver. Il est possible d'utiliser des techniques de ML à partir des enregistrements ECG 12 dérivations de 10 secondes utilisés en routine, jusqu'à des ECG à une seule dérivation pour des durées prolongées, via les enregistreurs en boucle, *implantable loop recorders* (ILR). Les données placées en entrée des modèles peuvent être très variées depuis le signal ECG entier échantillonné à 1000 Hz jusqu'aux intervalles RR échantillonnés à 200 Hz, voire moins. On peut essayer d'obtenir comme output une interprétation, une prévision, le diagnostic d'une ischémie, mais aussi d'un trouble du rythme. On peut enregistrer l'ECG de manière répétitive (par ex. durant 30 secondes 2x/j) ou à la demande (durant de longues périodes). D'autre part des enregistrements de phénomènes cardiaques autres que l'ECG peuvent également être utilisés : photopléthysmographie (PPG) ou ballistographie, retrouvant les techniques d'antan telles les mécanocardiogrammes. Il importe donc de bien situer le cadre et la méthodologie des résultats des études rapportées dans la littérature, afin de ne pas attribuer à l'IA des qualités qu'elle ne possède pas.

Le terme 'prédiction' a un sens différent selon son domaine d'utilisation et les confusions sont multiples. En ML, ce terme est utilisé pour rapporter les données sortantes du modèle, mais attention : prédire n'est pas prévoir. Selon le Robert une prédiction est une chose qui a une faible probabilité de se produire et du temps des Romains, une prédiction relevait de la divination. De plus, il faut distinguer la prévision de la prévention qui consiste à

déterminer et corriger les facteurs de risque.

Techniques d'IA

Machine learning et deep learning

Le *deep learning* (DL) n'a pas besoin de données structurées. Cette approche est particulièrement adaptée pour les tâches complexes, lorsque tous les aspects des objets à traiter ne peuvent pas être classifiés ou catégorisés en amont. Le réseau identifie lui-même les caractéristiques discriminantes.

Tandis que le ML fonctionne à partir d'une base d'une quantité de données contrôlable, le DL a besoin d'un volume de données bien plus considérable (en général plusieurs centaines de milliers d'éléments) pour optimiser ses paramètres.

Modèles de deep learning

Réseaux de neurones récurrents

Les réseaux de neurones récurrents, *recurrent neural network* (RNN), ont été introduits pour que le modèle puisse traiter plus efficacement les séquences de données où l'ordre des entrées a de l'importance. Les RNN traitent la séquence d'entrée un élément à la fois et utilisent le retour d'information des éléments précédents dans le calcul des éléments suivants de la séquence. Ils conviennent donc parfaitement au traitement des données de nature séquentielle, comme l'ECG. Le modèle de mémoire à long terme (LSTM) est une version améliorée du RNN conventionnel.

Réseaux de neurones convolutionnels

Les réseaux de neurones convolutionnels, *convolutional neural network* (CNN), sont des réseaux capables d'identifier des modèles simples dans les données. Plus le nombre de couches est élevé, plus les modèles peuvent être complexes. Les CNN peuvent être appliqués à des séquences de données unidimensionnelles (par ex. ECG), bidimensionnelles (par ex. images) ou même tridimensionnelles (par ex. vidéo).

Applications de l'IA pour les ECG

ECG 12 dériviatiions

Chang et al. ont développé un modèle LSTM capable à la fois d'identifier les infarctus STEMI et 12 rythmes cardiaques à partir de 60 537 ECG 12 dériviatiions enregistrés chez 35 981 patients. L'AUC de leur modèle est de 0,987, ce qui est supérieur à celle des cardiologues (0,898), des urgentistes (0,820), des internistes (0,765) et d'un algorithme commercial (0,845) Ils suggèrent que leur algorithme puisse être utilisé pour optimiser le processus de triage des patients présentant des douleurs thoraciques aiguës et accélérer le traitement de reperfusion des patients STEMI.⁵

Réaliser des interprétations simples un peu mieux que les cardiologues présente certainement un intérêt, mais l'utilisation de l'IA en tant qu'outil pouvant permettre au médecin de voir ce que lui-même est incapable de voir semble plus intéressant. Dans ce cadre, l'ECG peut s'utiliser comme capteur biométrique semblable à un fond d'oeil ou à une empreinte digitale, car il est spécifique à chaque sujet⁶. Pour cette raison, il n'est pas interdit de croire que l'ECG pourrait être utilisé comme technique pour prévoir la susceptibilité à certaines maladies cardiovasculaires, au même titre que la génomique, dans le cadre de la médecine de précision.

Attia et al. ont développé un DNN capable de donner le sexe et l'âge à 7 ans près⁷. Un patient dont l'ECG montre une différence de plus de 7 ans entre l'âge estimé par le DNN et l'âge réel présenterait un risque de surmortalité qui pourrait être évitée par une correction de facteurs de risque ou un traitement préventif. La même équipe de la Mayo Clinic a ensuite utilisé un DNN pour analyser des ECG 12 dériviatiions standard dans une cohorte de 25 144 sujets de plus de 30 ans. Ils ont montré que le DNN pouvait prédire l'âge des patients et que la différence entre l'ECG et l'âge chronologique était un prédicteur indépendant de la mortalité cardiovasculaire et toutes causes confondues⁸. Cette idée n'est d'ailleurs pas nouvelle : les effets de l'âge, du sexe et de l'ethnie sur l'ECG ont été documentés, en premier lieu par Simonson à partir de

mesures manuelles sur les ECG⁹. En 1994, Mac Farlane a confirmé ces résultats sur des ECG numérisés¹⁰. Encore récemment, Hnatkova a montré l'effet du sexe et de l'ethnie sur la durée du QRS¹¹.

L'ECG peut même être utilisé pour dépister l'insuffisance cardiaque. Attia et al. de la Mayo Clinic ont utilisé des données appariées d'ECG 12 dérivations et d'échocardiogramme, provenant de 44 959 patients. Ils ont entraîné leur CNN en utilisant uniquement l'ECG pour identifier les patients présentant un dysfonctionnement ventriculaire, défini comme une fraction d'éjection inférieure à 35 %. Testé sur un ensemble indépendant de 52 870 patients, le modèle a donné des valeurs pour l'AUC, la sensibilité, la spécificité et la précision de 93 %, 86,3 %, 85,7 % et 85,7 %, respectivement. Chez les patients ne présentant pas de dysfonction ventriculaire, ceux dont le dépistage par le CNN était positif avaient un risque 4 fois plus élevé de développer une future dysfonction ventriculaire par rapport à ceux dont le dépistage était négatif². Joon-myung Kwon et al. ont développé un modèle de DL qui est capable de détecter l'insuffisance cardiaque à fonction préservée en utilisant un ECG à 12 dérivations¹³. Ces résultats suggèrent que l'insuffisance cardiaque à fonction préservée pourrait aussi être dépistée grâce aux ECG conventionnels.

C'est presque devenu une habitude pour les éditeurs et les reviewers d'accepter, apparemment sans broncher, le fait que les algorithmes surpassent les cardiologues et de publier les résultats rapportés sans en tester les modèles et sans même voir le code utilisé. Il faut certainement relativiser cet enthousiasme : les algorithmes sont plus performants que les cardiologues certifiés (pas les experts) et pour des interprétations d'ECG de routine (non complexes). Un des intérêts des techniques diagnostiques d'IA réside en réalité dans le fait de décharger les cardiologues des tâches considérées comme secondaires, telles l'interprétation d'un ECG et de leur permettre de se consacrer à des tâches beaucoup plus complexes et gratifiantes dans les salles de cathétérisme et d'électrophysiologie. Notons d'emblée qu'en ce qui concerne l'interprétation de l'ECG, l'IA, même si elle est

déjà très performante pour les diagnostics de base, reste cependant encore éloignée d'une performance optimale pour les troubles du rythme plus compliqués. Des progrès substantiels vont probablement s'accomplir dans les années à venir.

EKG à une ou deux dérivations

Acharya et al.¹⁴ ont conçu un CNN pour classer un EKG brut en quatre classes différentes : rythme sinusal normal, fibrillation auriculaire, flutter auriculaire et fibrillation ventriculaire. La sensibilité était de 99,13 et la spécificité de 81,44 sur des segments d'EKG de cinq secondes.

Le fait que l'IA puisse être utilisée dans le dépistage de la FA a donné lieu au développement de son utilisation dans les wearables. Lors de l'étude Apple Watch réalisée avec la montre Apple 3, 419 297 participants ont été recrutés. Si un algorithme de détection de pouls irrégulier basé sur la smartwatch identifiait une possible FA, une visite de télé-médecine était lancée et un patch EKG était envoyé par courrier au participant, qui devait le porter pendant 7 jours maximum. Les principaux objectifs étaient d'estimer la proportion de participants notifiés présentant une FA indiquée sur un patch EKG et la valeur prédictive positive des intervalles de pouls irréguliers avec une largeur d'intervalle de confiance ciblée de 10 %. Parmi les participants qui ont reçu une notification de pouls irrégulier, seulement 20,8 % ont renvoyé les patchs EKG pour analyse, dont 34 % ont présenté une FA lors des lectures ultérieures et 84 % des notifications étaient concordantes avec une FA ; la valeur prédictive positive de l'algorithme de détection était de 71 %. Cette étude a montré la faisabilité d'un dépistage à grande échelle grâce à une montre connectée, mais elle a surtout mis en évidence la faible rentabilité diagnostique de ce dépistage réalisé dans une population à faible prévalence de la maladie (0,5 % des participants ont reçu des notifications de pouls irrégulier) et la faible participation des sujets recrutés jusqu'à la fin de l'étude¹⁵. Actuellement, l'Apple Watch 4 permet de réaliser directement un enregistrement EKG d'une dérivation et qui est ensuite envoyé sur le smartphone pour interprétation par un algorithme de ML, puis éventuellement par un médecin

pour validation. Bien sûr, pratiquement toutes les firmes telles que Samsung, Huawei, Garmin et Polar développent leurs propres algorithmes de ML. Évidemment tous ces algorithmes sont brevetés et généralement maintenus secrets, ce qui peut nuire à la transparence des données et des résultats.

Giudicessi et al. ont très récemment démontré la capacité d'un DNN à reconnaître un allongement cliniquement significatif de l'intervalle QT corrigé ($QTc > 500$ ms) sur des tracés d'appareils ECG mobiles, qui était similaire aux mesures du QTc basées sur l'ECG à 12 dérivations, déterminées à la fois par un cardiologue expert en QT et un laboratoire central commercial¹⁶. Ils concluent qu'à partir d'un ECG deux dérivations obtenues par un smartphone, leur DNN peut prédire le QTc aussi bien qu'un ECG standard à 12 dérivations. On aurait évidemment souhaité qu'ils utilisent leur DNN comme prédicteur des patients qui vont présenter des torsades de pointe, mais cela n'est déjà pas si mal.

Enregistrements de longue durée par patch ECG

Hannun et al.¹⁷ ont développé un DNN pour identifier 12 troubles du rythme différents à partir de 91 232 ECG à une seule dérivation tirés du Zio Patch monitor et provenant de 53 549 patients. Ils ont utilisé un CNN à 34 couches de neurones en prenant comme entrée un signal ECG brut de 30 secondes échantillonné à 200 Hz. Le réseau produit une séquence de prédictions diagnostiques toutes les secondes. Ils ont validé leur modèle sur 328 ECG annotés par des cardiologues experts, mais ont prudemment comparé ses performances aux interprétations de cardiologues 'certifiés'. Leur DNN a atteint une AUC de 0,97. Avec une spécificité fixée à la spécificité moyenne obtenue par les cardiologues, la sensibilité du DNN a dépassé la sensibilité moyenne des cardiologues pour les 12 classes de rythme.

Enregistrements d'ECG périodiques, en boucle par ILR

Une arythmie a d'autant plus de chance d'être détectée que la durée d'observation est longue. Ces appareils, au format d'une clé USB et implantables en sous-cutané,

permettent l'enregistrement en boucle de l'ECC avec mise en mémoire des épisodes d'arythmies détectés. Ils sont programmables de sorte que leurs seuils de détection et la durée d'enregistrement puissent être adaptés à chaque patient. La durée de surveillance est de plusieurs années. Initialement développés pour permettre l'enregistrement des troubles de la conduction à l'origine d'épisodes syncopaux, les ILR ont vu leurs capacités diagnostiques augmentées grâce à des algorithmes de détection de FA (tableau 2). Les algorithmes utilisés sont relativement simples étant donné que l'implémentation doit se faire en tenant compte d'une capacité mémorielle réduite.

Tableau 2. Performances des ILR par rapport aux Holter (tous, sauf le biomonitor 2 (6'), considèrent une FA à partir de 2 minutes d'enregistrement)							
Etude	ILR	sens (%)	spéc(%)	rec	an	ép	VPP
XPECT ¹⁸	Reveal-XT	89	91,3	247	206	76	79,3
AF detect study ¹⁹	Biomonitor 1	95,4	67	63	39	146	76,3
DETECT-AF ²⁰	Confirm-DM2102	100	85,7	90	79	16	64
Linq usability study ²¹	Reveal-LINQ	93,7	99,6	151	138	365	90,6
Piorkowski ²²	Biomonitor 2	100 %	88,1	23	15	23	72,5
Bisignani ²³	Biomonitor III m	96,9	96	19	19	234	82

sens : sensibilité ; spéc : spécificité ; rec : nombre d'enregistrements ; an : nombre de patients analysables ; ép : nombre d'épisodes ; VPP : valeur prédictive positive.

Pour un dispositif conçu pour détecter et surveiller la FA, la sensibilité doit être considérée comme le paramètre de performance le plus pertinent, car une sous-détection peut entraîner un sous-diagnostic, exposant ainsi le patient à un risque d'événements cliniques liés à la FA, notamment une syncope ou un accident vasculaire cérébral. Les paramètres les plus considérés dans toutes ces études sont donc la sensibilité et la valeur prédictive positive (VPP).

Enregistrements non basés directement sur l'ECC

Les technologies utilisant la photopléthysmographie (PPG) permettent de détecter les arythmies à l'aide d'un matériel déjà présent sur la plupart des appareils grand public (smartwatches et autres smartphones) par le biais d'une appli téléchargeable. Les capteurs de la plupart de ces wearables reprennent la technologie utilisée par les oxymètres pour mesurer le rythme cardiaque via le monitoring des variations du spectre lumineux réfléchi par la peau, consécutives au pouls capillaire. Ce signal pulsatile est interprété comme équivalent à l'onde R de

l'ECG. Un algorithme, basé principalement sur l'irrégularité du signal pose le diagnostic de FA. La performance des algorithmes interprétant ces signaux PPG s'est avérée être concordante avec les ondes R de l'ECG^{24,25}. Il est également possible d'utiliser le smartphone pour mesurer l'énergie cinétique transmise par le choc de pointe du ventricule G à la cage thoracique et d'en tirer des renseignements concernant la fonction myocardique, reproduisant en mode digital les bons vieux apexogrammes.²⁶

Le big data de la médecine connectée

Le DL n'est vraiment optimal que grâce à deux facteurs principaux : l'augmentation de la puissance de calcul et la quantité de données disponible. C'est dans ce cadre qu'il faut comprendre l'intérêt des appareils connectés en cardiologie. Le nombre d'utilisateurs de smartphones ne cesse d'augmenter. En 2021, ils étaient 6,4 milliards dans le monde, soit une augmentation annuelle de 5,3 %²⁷. Les smartphones peuvent soit collecter des données directement via les variations de couleur de la peau détectées par la caméra, soit recueillir des informations envoyées par les wearables. Le nombre de wearables connectés dans le monde a plus que doublé en l'espace de trois ans, passant de 325 millions en 2016 à 722 millions en 2019. Selon les prévisions, le nombre de ces appareils devrait atteindre plus d'un milliard d'ici à 2022.²⁸

La santé mobile - ou '*mHealth*' - est une composante de la santé numérique, définie par l'OMS comme une pratique médicale et de santé publique soutenue par des dispositifs mobiles, tels que les téléphones mobiles, les dispositifs de suivi des patients, les assistants numériques personnels et autres dispositifs sans fil.

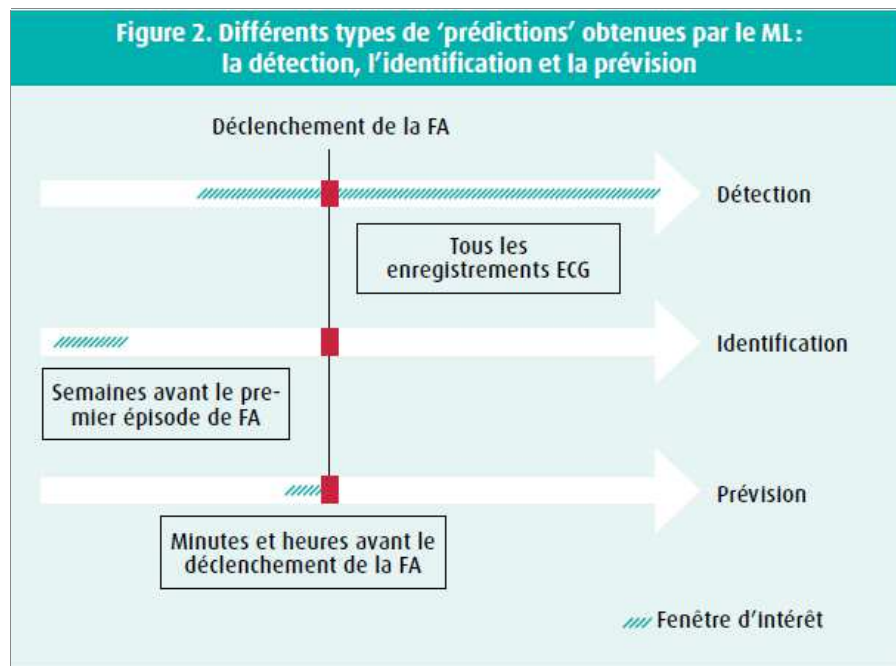
Ses différents aspects en cardiologie ont récemment fait l'objet d'un consensus²⁹. Un groupe de travail de l'ESC vient également de faire le point sur l'e-cardiologie.³⁰

La plupart des wearables peuvent maintenant enregistrer des ECG limités à une seule dérivation. Des études de validation sont réalisées en comparant les performances d'ECG à une ou deux dérivations ECG à celles d'ECG 12

dérivations. Le *Physio Net Challenge 2021*³¹ a d'ailleurs pour objectif d'identifier des diagnostics cliniques à partir d'enregistrements ECG à douze dérivations, six dérivations périphériques, trois dérivations (I, II et V2) et deux dérivations (II et V5). Tout récemment, l'appareil Kardia Mobile 6L permet un enregistrement de 30 secondes des 6 dérivations périphériques, ce qui a comme avantage de permettre la mesure du QT en plus de la détection de la FA. L'impact de ces technologies sur la prise de décision clinique vient d'être évalué afin de définir l'intégration des données issues de ces dispositifs.³²

Études prévisionnelles

Plusieurs études ont étudié le signal ECG en rythme sinusal afin de déterminer si un réseau pouvait distinguer les patients qui allaient présenter ultérieurement un trouble du rythme. Dans ce cas, l'utilisation de l'IA permettrait d'adopter une médecine préventive personnalisée, ce qui est un pas plus loin qu'une médecine préventive basée sur les facteurs de risque. Il reste encore à déterminer à quel moment précis ce trouble du rythme va se manifester (figure 2).



Galloway et al. ont montré qu'un réseau neuronal profond était capable de trouver une signature de FA à partir d'ECG une dérivation de 20 secondes avec un rythme sinusal

normal en utilisant un appareil compatible avec les smartphones. Ils ont divisé les patients souffrant de FA en deux groupes, selon qu'ils avaient plus ou moins 30 % de FA. Leur modèle de DL, exécuté sur des ECG en rythme sinusal normal enregistrés avant la FA, a pu prédire si l'ECG provenait d'un patient sans FA ou avec 30 % ou plus de FA. En utilisant un point de référence avec une sensibilité et une spécificité égales, la sensibilité et la spécificité du modèle étaient de 73,1 %.³³

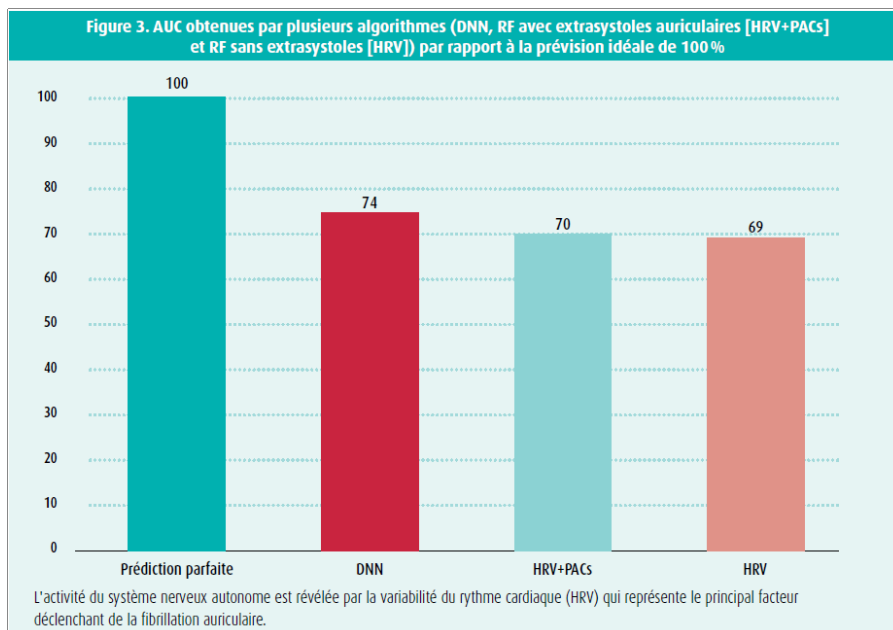
Attia et al. ont utilisé un CNN pour identifier une signature de FA chez les patients atteints de FA paroxystique en rythme sinusal normal en utilisant des ECG standard de 10 secondes et 12 dérivations. La fenêtre d'intérêt pour les patients atteints de FA paroxystique a commencé 31 jours avant le premier ECG enregistré de FA. Leur modèle identifie la FA avec une AUC de 0,87, une sensibilité de 79,0 %, une spécificité de 79,5 %, un score F1 de 39,2 % et une précision globale de 79,4 %.³⁴

Dans une étude rétrospective multicentrique effectuée sur 12 955 patients, Yong-Yeo Jo et al. ont développé un modèle de DL capable d'identifier les patients qui vont présenter un épisode de tachycardie supraventriculaire alors qu'ils sont encore en rythme sinusal³⁵. Leur modèle a été développé à partir de 31 147 ECGs et a présenté une AUC de 0,966. La précision, la sensibilité, la spécificité, la valeur prédictive positive et la valeur prédictive négative étaient respectivement de 0,970, 0,868, 0,972, 0,255 et 0,998. Les éléments importants servant à la prédiction comprenaient l'intervalle QT et l'existence ou non d'une onde delta.

Les problèmes de l'interprétabilité et de la reproductibilité
Si le clinicien veut considérer l'IA comme une technique qui apporte des certitudes et non comme une technique de magie divinatoire, il est impératif d'en souligner les limites et d'en préciser les méthodes

L'explicabilité des modèles de DL reste un sujet brûlant et l'extraction des caractéristiques utilisées par ces modèles fait l'objet de travaux intenses. Par ailleurs, créer des modèles qui soient interprétables en premier lieu, plutôt

que d'essayer de les expliquer a posteriori semble une autre alternative peut-être plus valable³⁶. Interprétabilité et explicabilité sont deux concepts complémentaires, mais différents. Par exemple, nous avons utilisé 2 algorithmes de ML différents pour essayer de déterminer à partir de quelle durée de temps le déclenchement d'une crise de FA était prévisible. Nous avons programmé sur les mêmes données ECG en entrée: d'une part un DNN prenant uniquement la succession de tous les intervalles RR et d'autre part un algorithme de random forest (RF) prenant 9 paramètres de variabilité du rythme cardiaque. Il n'y avait pas de différence de performance entre les 2 algorithmes, suggérant que la 'boîte noire' du DNN utilise les paramètres de variabilité pour réaliser sa prédiction³⁷ (figure 3). L'intérêt de connaître les données en entrée nous a permis d'établir des corrélations avec la pathophysiologie, rendant en cela l'algorithme de DL interprétable. Nous avons ainsi pu confirmer le rôle déterminant du système nerveux autonome dans les minutes qui précèdent les épisodes de FA.



Il ne faut cependant pas rejeter l'utilisation de la boîte noire de l'IA sous prétexte qu'on ne dispose pas d'explications suffisamment détaillées sur son fonctionnement. Connaître le rôle physiologique de tous les paramètres n'est pas indispensable pour développer une application clinique individuelle qui fonctionne : ignorer le détail du fonctionnement des pièces du moteur

d'une voiture n'empêche pas de savoir conduire correctement.

Par contre, une description détaillée de la méthodologie est cruciale, d'autant qu'en ce domaine, les méthodes utilisées peuvent se révéler extrêmement complexes. Olorisade et al. nous révèlent qu'il leur a été difficile, voire impossible, de reproduire de manière indépendante les résultats publiés dans 33 études de ML qu'ils ont examinées en détail. Plus grave encore, le manque d'informations sur les bases de données utilisées en limitait la reproductibilité d'environ 80 % et les informations sur les algorithmes d'apprentissage automatique étaient jugées insuffisantes dans 27 % des cas³⁸. Le problème est tel que certains évoquent une véritable 'crise de reproductibilité'³⁹.

Seule une minorité de publications a étudié la reproductibilité et la généralisation des méthodes de ML en cardiologie et il existe très peu d'essais cliniques prospectifs. Il devient urgent d'améliorer la communication de l'évaluation et la transparence des données et des méthodes utilisées⁴⁰.

Conclusions

L'IA doit encore trouver sa place exacte en cardiologie et en particulier dans son utilisation à partir de l'ECG. Elle n'en est encore qu'à ses balbutiements aussi bien au niveau de l'interprétation qu'en ce qui concerne son utilisation pour la prévention cardiovasculaire et la prévision des arythmies.

Les possibilités accrues de l'e-santé permettent d'obtenir des quantités de données très importantes qui peuvent être optimalement exploitées par les réseaux de DL. Elle est déjà utilisée pour le diagnostic automatique de la FA dans le cadre de la médecine connectée en utilisant des ECG à une ou deux dérivations.

L'ECG 12 dérivations, d'une durée de 10 secondes, pourrait dans un futur proche être utilisé comme une signature (au même titre qu'une empreinte digitale) permettant d'identifier un patient et de prévoir son risque de développer de la FA, de la TSV, une insuffisance cardiaque

à fonction préservée ou non.

Toutes ces stratégies doivent clairement encore trouver leur place et on devrait être capable de valider indépendamment chaque algorithme et d'en prouver l'utilité clinique avant d'approuver son utilisation en routine.

Le DL nécessite l'utilisation de bases de données très importantes pour dépasser la performance d'algorithmes de ML plus simples. Ceux-ci ont l'avantage d'être interprétables, dans la mesure où on contrôle les données entrées. L'explicabilité des boîtes noires du DL donne lieu à des travaux importants. Si l'on veut que l'IA soit acceptée par tous, il est crucial que la méthodologie des différentes études soit reproductible et que les bases de données utilisées soient accessibles.

La mise en concurrence de l'homme et la machine semble nécessaire pour la validation des résultats, mais il nous semble que l'intérêt principal de l'IA réside plutôt dans son utilisation comme un outil collaboratif, permettant d'obtenir des informations impossibles à obtenir par une autre approche, informations invisibles à l'oeil du cardiologue expert, comme le microscope a permis et permet encore des avancées importantes.

Références

1. Macallan, P. A brief history of computer-assisted electrocardiography. *Methods Inf Med*, 1990, 29 (04), 272-281.
2. Smulyan, H. The computerized ecg: friend and foe. *Am J Med*, 2019, 132 (2), 153-160.
3. Brisk, R., Bond, R., Finlay, D., McLaughlin, J., Piodlo, A., Leslie, S.J. et al. The effect of confounding data features on a deep learning algorithm to predict complete coronary occlusion in a retrospective observational setting. *Eur Heart J - Digital Health*, 2021, 2 (1), 127-134.
4. Breiman, L. Statistical modeling: the two cultures (with comments and a rejoinder by the author). *Statist Sci*, 2001, 16 (3), 199-231.
5. Chang, K-C., Hsieh, P-H., Wu, M-Y., Wang, Y-C., Wei, J-T., Shih, E.S. et al. Usefulness of multi-labelling artificial intelligence in detecting rhythm disorders and acute st-elevation myocardial infarction on 12-lead electrocardiogram. *Eur*

- Heart J - Digital Health*, 2021, 2 (2), 299-310.
6. Lourenço, A., Silva, H., Fred, A. Unveiling the biometric potential of finger-based ecg signals. *Comput Intell Neurosci*, 2011, (720971), 1-8.
 7. Attia, Z.I., Friedman, P.A., Noseworthy, P.A., Lopez-Jimenez, F., Ladewig, D.J., Satam, G. et al. Age and sex estimation using artificial intelligence from standard 12-lead ecgs. *Circ Arrhythm Electrophysiol*, 2019, 12 (9), e007284.
 8. Ladejobi, A.O., Medina-Inojosa, J.R., Shelly Cohen, M., Attia, Z.I., Scott, C.G., LeBrasseur, N.K. et al. The 12-lead electrocardiogram as a biomarker of biological age. *Eur Heart J - Digital Health*, 2021, doi:10.1093/ehjdh/ztabo43.
 9. Simonson, E. Differentiation between normal and abnormal in electrocardiography. *Acad Med*, 1962, 37 (2), 161.
 10. Macfarlane, P., McLaughlin, S., Devine, B., Yang T. Effects of age, sex, and race on ecg interval measurements. *J Electrocardiol*, 1994, 27, 14-19.
 11. Hnatkova, K., Smetana, P., Toman, O., Schmidt, G., Malik, M. Sex and race differences in QRS duration. *Ep Europace*, 2016, 18 (12), 1842-1849.
 12. Attia, Z.I., Kapa, S., Lopez-Jimenez, F., McKie, P.M., Ladewig, D.J., Satam, G. et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med*, 2019, 25 (1), 70-74.
 13. Kwon, J-M., Kim, K-H., Eisen, H.J., Cho, Y., Jeon, K-H., Lee, S.Y. et al. Oh, Artificial intelligence assessment for early detection of heart failure with preserved ejection fraction based on electrocardiographic features. *Eur Heart J - Digital Health*, 2021, 2 (1), 106-116.
 14. Acharya, U.R., Fujita, H., Lih, O.S., Hagiwara, Y., Tan, J.H., Adam, M. Automated detection of arrhythmias using different intervals of tachycardia ecg segments with convolutional neural network. *Information Sciences*, 2017, 405, 81-90.
 15. Perez, M.V., Mahaffey, K.W., Hedlin, H., Rumsfeld, J.S., Garcia, A., Ferris, T. et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med*, 2019, 381 (20), 1909-1917.
 16. Giudicessi, J.R., Schram, M., Bos, J.M., Galloway, C.D., Shreibati, J.B., Johnson, P.W. et al. Artificial intelligence-enabled assessment of the heart rate corrected qt interval using a mobile electrocardiogram device. *Circulation*, 2021, 143 (13), 1274-1286.
 17. Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., Tison, G.H., Bourn, C., Turakhia, M.P. et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med*, 2019, 25 (1), 65-69.
 18. Hindricks, G., Pokushalov,

- E., Urban, L., Taborsky, M., Kuck, K-H., Lebedev, D. et al. Performance of a new leadless implantable cardiac monitor in detecting and quantifying atrial fibrillation results of the xpect trial. *Circ Arrhythm Electrophysiol*, 2010, 3 (2), 141-147.
19. Ciconte, G., Saviano, M., Giannelli, L., Calovic, Z., Baldi, M., Ciaccio, C. et al. Atrial fibrillation detection using a novel threevector cardiac implantable monitor: the atrial fibrillation detect study. *Ep Europace*, 2017, 19 (7), 1101-1108.
 20. Noelker, G., Mayer, J., BOLDT, L-H., Seidl, K., Van Driel, V., Massa, T. et al. Performance of an implantable cardiac monitor to detect atrial fibrillation: results of the detect af study. *J Cardiovasc Electrophysiol*, 2016, 27 (12), 1403-1410.
 21. Sanders, P., Pürerfellner, H., Pokushalov, E., Sarkar, S., Di Bacco, M., Maus, B. et al. Performance of a new atrial fibrillation detection algorithm in a miniaturized insertable cardiac monitor: results from the reveal linq usability study. *Heart Rhythm*, 2016, 13 (7), 1425-1430.
 22. Piorkowski, C., Busch, M., Nölker, G., Schmitt, J., Roithinger, F.X., Young, G. et al. Clinical evaluation of a small implantable cardiac monitor with a long sensing vector. *Pacing Clin Electrophysiol*, 2019, 42 (7), 1038-1046.
 23. Bisignani, A., De Bonis, S., Mancuso, L., Ceravolo, G., Giacomelli, D., Pelargonio, G. et al. Are implantable cardiac monitors reliable tools for cardiac arrhythmias detection? an intra-patient comparison with permanent pacemakers. *J Electrocardiol*, 2020, 59, 147-150.
 24. McManus, D.D., Chong, J.W., Soni, A., Saczynski, J.S., Esa, N., Napolitano, C. et al. Pulse-smart: pulse-based arrhythmia discrimination using a novel smartphone application. *J Cardiovasc Electrophysiol*, 2016, 27 (1), 51-57.
 25. Proesmans, T., Mortelmans, C., Van Haelst, R., Verbrugge, F., Vandervoort, P., Vaes, B. Mobile phone-based use of the photoplethysmography technique to detect atrial fibrillation in primary care: diagnostic accuracy study of the fibrichk app. *JMIR mHealth and uHealth*, 2019, 7 (3), e12284.
 26. Oberlo. How many people have smartphones in 2021? URL: <https://www.oberlo.com/statistics>. (accessed: 22/07/2021).
 27. Statista. Number of connected wearable devices worldwide from 2016 to 2022. URL: <https://www.statista.com/statistics/487291/global-connectedwearable-devices/>. (accessed: 22/07/2021).
 28. Varma, N., Cygankiewicz, I., Turakhia, M.P., Heidbuchel, H., Hu, Y-F., Chen, L.Y. et al.

- 2021 ishne/hrs/ehra/aphrs expert collaborative statement on mhealth in arrhythmia management: digital medical tools for heart rhythm professionals: from the international society for holter and noninvasive electrocardiology/heart rhythm society/ european heart rhythm association/asiapacific heart rhythm society. *Circ Arrhythm Electrophysiol*, 2021, 14 (2), e009204.
29. Jensen, M.T., Treskes, R.W., Caiani, E.G., Casado-Arroyo, R., Cowie, M.R., Dilaveris, P. et al. Esc working group on e-cardiology position paper: use of commercially available wearable technology for heart rate and activity tracking in primary and secondary cardiovascular prevention—in collaboration with the european heart rhythm association, european association of preventive cardiology, association of cardiovascular nursing and allied professionals, patient forum, and the digital health committee. *Eur Heart J - Digital Health*, 2021, 2 (1), 49-59.
30. Physionet. Will two do? varying dimensions in electrocardiography: the physionet/ computing in cardiology challenge 2021 URL: <https://physionetchallenges.org/2021/>. (accessed: 22/07/2021).
31. Manninger, M., Zweiker, D., Svennberg, E., Chatzikiyriakou, S., Pavlovic, N., Zaman, J.A. et al. Current perspectives on wearable rhythm recordings for clinical decisionmaking: the wehrables 2 survey. *EP Europace*, 2021, 23 (7), 1106-1113.
32. Galloway, C., Treiman, D., Schreibati, J., Schram, M., Karbaschi, Z., Valys, A. et al. 5105 a deep neural network predicts atrial fibrillation from normal ecgs recorded on a smartphone-enabled device. *Eur Heart J*, 2019, 40 (Suppl 1), ehz746-0041.
33. Attia, Z.I., Noseworthy, P.A., Lopez-Jimenez, F., Asirvatham, S.J., Deshmukh, A.J., Gersh, B.J. et al. An artificial intelligence-enabled ecg algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *The Lancet*, 2019, 394 (10201), 861-867.
34. Jo, Y-Y., Kwon, J-M., Jeon, K-H., Cho, Y-H., Shin, J-H., Lee, Y-J. et al. Artificial intelligence to diagnose paroxysmal supraventricular tachycardia using electrocardiography during normal sinus rhythm. *Eur Heart J - Digital Health*, 2021, 2 (2), 290-298.
35. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 2019, 1 (5), 206-215.

36. Gregoire, J-M., Gilon, C., Carlier, S., Bersini, H. Unravelling the black box of machine learning for atrial fibrillation forecast: role of heart rate variability and of premature beats. *Eur Heart J*, 2020, 41 (Suppl 2), ehaa946-0671.
37.
38. Resnik, D.B., Shammoo, A.E. Reproducibility and research integrity. *Account Res*, 2017, 24 (2), 116-123.
39. Friedrich, S., Groß, S., König, I.R., Engelhardt, S., Bahls, M., Heinz, J. et al. Applications of ai/ml