



UNIVERSITÉ DE LILLE

FACULTÉ DE MÉDECINE HENRI WAREMBOURG

Année : 2024

THÈSE POUR LE DIPLÔME D'ÉTAT
DE DOCTEUR EN MÉDECINE

**Apport de l'Intelligence Artificielle dans la prise en charge des
lésions ostéoarticulaires du squelette appendiculaire aux urgences.**

Présentée et soutenue publiquement le 4 octobre 2024 à 16 heures
au Pôle Recherche

par Agathe VINOT

JURY

Président :

Monsieur le Professeur Éric WIEL

Assesseurs :

Monsieur le Docteur Jean-Marie RENARD

Monsieur le Docteur Jérôme MIZON

Directeur de thèse :

Monsieur le Docteur Erwan DUPIRE

AVERTISSEMENT

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

TABLE DES MATIERES

ABREVIATIONS	1
RESUME	2
INTRODUCTION	3
I. LE CENTRE HOSPITALIER DE VALENCIENNES	3
II. L'INTELLIGENCE ARTIFICIELLE	5
MATERIEL ET METHODE	16
I. TYPE D'ETUDE	16
II. CRITERES D'INCLUSION ET D'EXCLUSION	16
III. POPULATION ETUDIEE	17
IV. OBJECTIFS DE L'ETUDE	18
V. CRITERE DE JUGEMENT	18
VI. MODE DE RECUEIL DES DONNEES	20
VII. AUTRES DONNEES RECUEILLIES	20
VIII. MODELE STATISTIQUE	21
IX. SUR LE PLAN ETHIQUE	21
RESULTATS	23
I. ÉPIDEMIOLOGIE DESCRIPTIVE	23
II. CRITERE DE JUGEMENT PRINCIPAL	29
III. CRITERES DE JUGEMENT SECONDAIRES	30

DISCUSSION	41
I. CRITERE DE JUGEMENT PRINCIPAL	43
II. CRITERES DE JUGEMENT SECONDAIRES	45
III. FORCES ET FAIBLESSES	58
CONCLUSION	60
LISTE DES FIGURES ET DES TABLEAUX	61
BIBLIOGRAPHIE	62
ANNEXES	65

ABREVIATIONS

CERCL : Comité Éthique de la Recherche Clinique

CH : Centre Hospitalier

CHV : Centre Hospitalier de Valenciennes

GHT : Groupement Hospitalier du Territoire de santé

IA : Intelligence artificielle

IRM : Imagerie par résonance magnétique

MAO : Médecin d'accueil et d'orientation

ORL : Oto-Rhino-Laryngologie

PEC : Prise en charge

RP : Radiographies Pulmonaires

Se : Sensibilité

Sp : Spécificité

URC : Unité de Recherche Clinique

VPP : Valeur prédictive positive

VPN : Valeur prédictive négative

RESUME

Titre : Apport de l'Intelligence Artificielle dans la prise en charge des lésions ostéoarticulaires du squelette appendiculaire aux urgences.

Contexte : L'Intelligence Artificielle prend une place de plus en plus importante dans le domaine médical et spécialement dans l'aide diagnostique. Les traumatismes de membre sont un motif courant de consultation aux urgences et de nombreuses radiographies sont réalisées. Depuis 2022, les urgences adultes du Centre Hospitalier de Valenciennes sont dotées d'un logiciel d'intelligence artificielle (logiciel SmartUrgences® par la société MILVUE®) pour l'aide au diagnostic des lésions pulmonaires et ostéoarticulaires. Ce logiciel a déjà fait ses preuves et ses tests de performance ont été analysés. L'impact de ce logiciel en pratique clinique n'a pas été évalué à ce jour. L'intérêt de cette étude est d'évaluer ce logiciel comme aide diagnostique pour les urgentistes en traumatologie au moment de la prise en charge des patients et d'étudier les variables qui exercent une influence sur cette aide diagnostique.

Méthode : Il s'agit d'une étude monocentrique, observationnelle et descriptive, réalisée aux urgences adultes du Centre Hospitalier de Valenciennes. Elle a été réalisée par recueil de données rétrospectif, du 1er décembre 2023 au 1er mai 2024. Le *Gold Standard* est le compte rendu du radiologue édité à postériori et récupéré à distance de la prise en charge.

Résultats : 8,5% des diagnostics posés par l'urgentiste étaient erronés alors que l'IA était en accord avec le radiologue. L'IA est suivie lorsqu'elle a raison dans 96,1% des cas. Les facteurs associés au suivi de l'IA sont : l'absence de diagnostic de lésion par l'urgentiste, tout type de lésion et lorsque l'IA pose le bon diagnostic. Le taux de changement d'avis de l'urgentiste quand l'IA a raison est de 8,8%. La prise en charge a été modifiée vers celle adéquate dans 6,2%. Les facteurs associés à l'erreur de l'urgentiste sont : les lésions au niveau du pied, de la main, du genou et de la cheville, le doute initial, la demande d'avis complémentaire, et le fait d'être un urgentiste thésé. Le ressenti moyen des médecins vis-à-vis de l'IA est de 5,08 (+/- 2,98), pour des notes allant de 0 à 10.

Conclusion : L'IA constitue une aide diagnostique non négligeable pour le médecin urgentiste, et permet des modifications de prise en charge ainsi que des rectifications de diagnostic. Cependant, l'IA ne reste pas la seule décisionnaire face au diagnostic final. Elle a un impact dans la prise en charge des patients, en apportant au médecin une assistance supplémentaire. L'alliance IA et urgentiste reste la meilleure association pour diminuer les erreurs de diagnostic.

Mots clefs : intelligence artificielle, radiologie, urgence, lésions ostéoarticulaires

INTRODUCTION

I. LE CENTRE HOSPITALIER DE VALENCIENNES

A. Description

Situé au sein du Groupement Hospitalier du Territoire de santé (GHT) du Hainaut-Cambrésis dont il est l'hôpital de recours et de référence, le Centre Hospitalier de Valenciennes (CHV) est la 3ème structure hospitalière des Hauts de France et est régulièrement cité comme l'un des trois plus grands hôpitaux généraux de France. Il couvre un important bassin de population, soit 20% de la population totale du Nord-Pas-de-Calais (1).

Depuis plusieurs années, le service des urgences du CHV enregistre une hausse constante de sa fréquentation. En effet, entre 2021 et 2022, l'activité a augmenté de 6,6%, et entre 2022 et 2023 de 2,6 %. En 2023, 114 010 patients ont consulté aux urgences du CH de Valenciennes, (urgences adultes, urgences pédiatriques, et urgences gynécologiques). (2) Parmi ceux-ci, 75 520 patients ont consulté aux urgences adultes tout secteur confondu, soit une moyenne de 206 patients par jour. Face à cette augmentation de consultations l'Intelligence Artificielle et la lecture automatique des radiographies sont l'une des solutions pour optimiser la prise en charge des patients.

B. Organisation des urgences

En France, environ 40% des patients consultant aux urgences réalisent une imagerie médicale (4). Les radiographies sont les examens complémentaires de première ligne les plus fréquemment prescrits aux urgences (71% des imageries prescrites) du fait de leur accessibilité, de leur faible taux d'irradiation, de leur faible coût et de leur rapidité de réalisation (4). Réalisés dans un contexte d'urgence, ces clichés nécessitent aussi une interprétation radiologique 24h/24, 7j/7 (5). Or, dans de nombreux centres hospitaliers français, la couverture radiologique pour des examens aussi basiques n'existe pas ou très peu. Dans beaucoup de CH, dont le CHV, les images sont ainsi interprétées par un radiologue en différé.

La prise en charge d'un patient consultant aux urgences pour un traumatisme de membre est réalisée dans le secteur dit « ambulatoire » de 8h30 à minuit, tous les jours de la semaine. De minuit au lendemain matin, ces patients sont pris en charge du côté « médico-chirurgical ». La prise en charge est faite de façon conjointe avec les spécialistes. Un examen clinique initial est réalisé par un médecin urgentiste, qui prescrit ou non une radiographie. Pour accélérer la prise en charge sans attendre l'interprétation du cliché par l'expert, l'interprétation de l'image est réalisée par le médecin urgentiste. Cependant, ces derniers ne sont pas des experts dans ce domaine. Par conséquent, les erreurs de diagnostic sur l'analyse des images peuvent être non négligeables et préjudiciables pour la santé des patients. (6)

Les interprétations finales des clichés radiographiques sont réalisées par le radiologue à distance, tandis que la primo interprétation et la prise en charge initiale sont effectuées par le médecin urgentiste.

Lors de la relecture des images par le radiologue, si une discordance entre la conclusion de l'urgentiste et celle du radiologue est observée, une alerte est donnée au médecin d'accueil et d'orientation (MAO). Plusieurs situations sont envisageables : le patient peut être reconvoqué pour une adaptation thérapeutique aux urgences, rappelé pour une modification en externe, ou la discordance peut être classée sans suite.

Par conséquent, s'il y a eu une erreur diagnostique, celle-ci est rattrapée par la suite.

II. L'INTELLIGENCE ARTIFICIELLE

A. Qu'est-ce que l'intelligence artificielle ?

L'Intelligence Artificielle est un domaine de l'informatique, définie par le Parlement européen (7) (8), comme la possibilité pour une machine de reproduire des comportements humains, comme le raisonnement, la planification ou la créativité. (9) (10)

Ses origines remontent aux années 50, grâce à Alan Turing, qui a soulevé pour la première fois la question phare « une machine peut-elle penser ? ». L'idée *princeps* était de créer des machines pouvant simuler l'intelligence biologique. Le terme «intelligence artificielle» est apparu en 1956, lors d'une conférence à l'université de Dartmouth, dans le New Hampshire (USA) au cours de laquelle elle a été établie en tant que discipline scientifique (11). Les décennies qui ont suivi ont connu des avancées significatives par à-coups suivant le développement informatique. Depuis quinze ans, l'intelligence artificielle connaît un véritable bond. En effet, lors des prémices de son développement, elle était très limitée et ne pouvait réaliser qu'un seul

programme : celui pour lequel l'Humain l'avait paramétrée. Cependant, le traitement de données massives a permis un renouveau spectaculaire et des avancées majeures. Elle est donc devenue rapidement très performante entraînant une évolution des objectifs des domaines de recherche.

Dans un premier temps, les chercheurs ont voulu retrouver le rêve d'Alan Turing : concevoir une machine capable de raisonner comme l'être humain. Dans un second temps, ils ont voulu développer des machines permettant d'aider l'être humain dans ses tâches.

Il existe plusieurs branches et subdivisions de l'IA bien établies, liées à des domaines de recherche et des calculs informatiques différents. Nous allons nous focaliser sur une catégorie et une sous-catégorie de celles-ci.

- Le « *Machine Learning* », soit un apprentissage automatique. C'est une branche de l'intelligence artificielle qui se focalise sur le développement informatique afin d'apprendre aux machines, à partir de données et d'expériences, à tirer des conclusions sans l'avoir explicitement programmée pour cette tâche au préalable. Au plus il y aura de données, au plus les machines seront performantes.
- Le « *Deep Learning* » ou apprentissage profond, est quant à lui, une branche de l'apprentissage automatique et du *Machine Learning* en général. Il va se baser sur les réseaux de neurones déjà existants dans le *Machine Learning* et les améliorer. Ces réseaux de neurones artificiels profonds sont faits pour apprendre à partir de données. Le nom « réseaux de neurones » provient de son organisation similaire

au cerveau humain, en couches et sous couches. De cette façon, la machine peut apprendre des modèles complexes de données en utilisant plusieurs niveaux de représentation. *Le Deep Learning* est particulièrement utilisé pour les tâches nécessitant d'extraire des données brutes pour arriver à des résultats de modèles complexes tels que la reconnaissance d'image, la reconnaissance vocale ou la traduction automatique. Il nécessite cependant une quantité importante de données.

B. Dans le domaine médical

L'IA et le *Big Data*, défini comme une quantité de données que le cerveau humain ne peut pas traiter, s'immiscent dans tous les domaines car leur puissance de calcul et d'apprentissage nous surpasse. Dans le domaine de la santé où la connaissance est décisive, ses champs d'application sont nombreux. Elle peut s'ancrer dans une démarche de médecine prédictive, d'aide à la décision, de prévention, de chirurgie assistée avec les robots dits « compagnons », et est utilisée dans de nombreuses spécialités telles que l'oncologie, les spécialités chirurgicales, la radiologie etc. (12). L'enjeu de l'IA dans la recherche médicale est d'améliorer les performances des systèmes déjà en place ou en cours de développement, et de devenir les alliés des médecins pour l'aide à la décision, pour la stratégie diagnostique ou la réalisation d'actes (13).

En radiologie, l'augmentation exponentielle des articles scientifiques s'intéressant à l'IA met en évidence le potentiel et les perspectives très prometteuses

de cette dernière (14)(15)(16)(17)(18). Aux urgences, l'intelligence artificielle a aussi de nombreux rôles à jouer (19)(20).

Par exemple, elle peut apporter une aide au diagnostic dans la détection des lésions telles que les fractures en imagerie musculosquelettique (21). Cela est un enjeu majeur puisque les fractures non diagnostiquées représentent plus de 40% des erreurs de diagnostic aux urgences. Ces erreurs sont plus fréquentes la nuit et démontrent l'importance du facteur humain dans la prise en charge (22)(23). Ces erreurs n'entraînent pas de conséquences forcément létales mais elles peuvent entraîner des complications à moyen ou long terme et portent atteinte à la crédibilité du système de prise en charge. (24). En revanche, la « machine » n'est pas influencée par les conditions extérieures lors de son analyse.

Actuellement, l'enjeu de l'IA consiste à établir la meilleure alliance possible entre l'humain et la machine pour majorer le potentiel du médecin grâce au bon usage de la technologie (12).

C. L'Intelligence artificielle au sein du CHV

1. Domaines de compétences

Le CHV a mis en place plusieurs logiciels d'IA permettant l'amélioration des pratiques dans de multiples domaines au sein de l'établissement. Cette mise en place s'est déroulée sur plusieurs années.

Tout d'abord, PharmIA® est une intelligence artificielle au service de la pharmacie hospitalière. Elle permet une aide à la décision pharmaceutique. Ce logiciel alerte les pharmaciens sur les problématiques liées aux médicaments, détectées à partir d'une

variété de paramètres contextuels (biologiques, pharmacologiques, physiologiques). Il permet également de prioriser l'analyse sur des profils de patients plus à risque.

De même, l'IA Prévia Médical © permet de prévoir et de prévenir les urgences vitales en milieu hospitalier. Cela est possible grâce à une analyse en temps réel des données tels que les paramètres vitaux, les résultats de laboratoire ou les antécédents du patient. Des modèles prédictifs détectent les anomalies signalant une urgence vitale imminente. Depuis 2021 Prévia© collabore avec l'équipe de maladies infectieuses et tropicales du CHV sur le SEPSI SCORE © dans la détection du sepsis (25).

Depuis quelques années, le projet du CHV et de ses urgences est de s'adapter à l'afflux toujours plus important de patients ainsi qu'aux sur-sollicitations. C'est pourquoi, en septembre 2019, le CHV s'est doté d'un logiciel d'IA, Saniia® , élaboré avec la société suisse Calyps® (26) qui permet de prédire le nombre de patients attendu chaque jour dans les locaux des urgences et de coordonner leur parcours hospitalier. Cela permet ainsi aux soignants d'adapter les besoins et les ressources en prévision du flux. La prédiction des flux proposée par CALYPS Saniia® permet une indication du nombre d'admissions quotidiennes et d'hospitalisations pour les 5 jours suivants et plus précisément sur les 24 heures en cours. Cette vision permet d'anticiper les éventuelles tensions à venir en adaptant les organisations tant au niveau du service des urgences qu'au niveau de l'établissement avec l'anticipation du capacitaire d'aval nécessaire. Les variables qui peuvent être à l'origine de flux plus ou

moins intenses sont multiples : de la météo aux événements sportifs, du trafic routier aux dates de calendrier.

La prochaine étape de CALYPS Saniia® aux urgences adultes du CHV est de proposer au patient et à son entourage d'accéder à une prédiction des différentes étapes horodatées de son parcours aux urgences, de son admission jusqu'à son orientation finale, tout en intégrant les examens cliniques et paracliniques. L'objectif est d'améliorer la communication du service des urgences aux patients et de réduire les interruptions de tâches pour les soignants.

Calyps Saniia®, a pu progresser et s'ouvrir à d'autres problématiques, comme les demandes de brancardage pour la mobilité des patients entre les différents services. La ligne directrice de ces avancées est de fluidifier les prises en charge et de permettre une organisation de service rigoureuse.

Une des caractéristiques de Calyps Saniia® est sa capacité à optimiser l'allocation des ressources. Depuis 2022, le CHV a mis en place une cellule de « *bed management* » afin de mieux répondre aux besoins des urgences.

Grâce à un algorithme, cette solution d'IA propose le lit le plus adapté ce qui limite les changements de chambre. Depuis la mise en place de ces nouveaux usages, la mise en lit peut être faite directement au chevet du patient, ce qui permet de faire gagner du temps aux professionnels de santé, et bien-sûr du confort au patient.

2. Instauration de l'IA sur la radiologie au CHV

Le 11 octobre 2021, l'intelligence artificielle pour la radiologie conventionnelle a été instaurée au CHV. La société Milvue® a mis en place une IA (SmartUrgences®)

venant se greffer au logiciel de radiologie déjà présent pour l'analyse des radiographies pulmonaires et ostéo articulaires à l'exception des radiographies du crâne et du rachis.

Il s'agissait d'une version 2, la première ayant été testée en 2020 mais restait très perfectible.

Ce logiciel est en auto-apprentissage. Plus il interprètera de radiographies, plus il sera performant. Cela peut se comparer à l'expérience humaine du médecin : plus il prendra en charge des patients, plus il sera performant dans son domaine.

Au contraire, il reste perfectible pour l'interprétation des RP. En effet, à l'heure actuelle seule la recherche d'un épanchement liquidien ou d'un pneumothorax ont des indicateurs de performance équivalents à ceux de l'IA en ostéoarticulaire.

En parallèle, les radiographies standards seront toujours interprétées par les radiologues et pourront toujours faire l'objet de discordances après comparaison des conclusions du radiologue et du clinicien. Les radiologues disposent également de l'aide de cet outil pour l'interprétation des radiographies, l'association IA + urgentiste étant la meilleure combinaison pour renforcer la qualité des interprétations à l'instant T.

Pour illustrer, un cliché supplémentaire apparaît lors de la lecture des radiographies et propose 3 types de réponse :

- Pour l'ostéoarticulaire : POSITIF/DOUTE/NÉGATIF (fracture, luxation, épanchement). L'IA précise ce qu'elle suspecte ou ce qu'elle élimine à chaque fois.

- Pour la radiographie pulmonaire : POSITIF/DOUTE/NÉGATIF (épanchement, opacité, pneumothorax). De même, l'IA précise ce qu'elle suspecte ou ce qu'elle élimine.

Elle précise aussi la localisation de la lésion : sur le cliché apparaît un rectangle encadrant la lésion suspectée : en trait plein si l'IA est sûre d'elle ou en pointillé si elle a un doute.

Il ne faut pas oublier que l'IA est un outil d'aide au diagnostic, et que la décision finale appartient toujours au médecin prenant en charge le patient.

L'interprétation des clichés par l'IA ne peut pas se substituer à celle du médecin.

Une étude a été menée en partenariat avec la société Milvue® et le CH de Valenciennes afin d'étudier l'efficacité du logiciel SmartUrgences® pour la détection des anomalies radiographiques osseuses et pulmonaires sur les clichés réalisés à l'hôpital.(4)

Cette étude a été réalisée entre juillet et octobre 2020, et a concerné 1772 patients admis aux urgences. SmartUrgences® a permis de détecter de nombreuses anomalies. Une comparaison des comptes-rendus des radiologues et des médecins urgentistes avec ceux de l'IA a été réalisée.

Les résultats ont révélé dans 9,7% des cas un désaccord de diagnostic entre le médecin urgentiste et le radiologue. Parmi ces désaccords, dans 90,1% des cas, le diagnostic de la solution Milvue® était en accord avec celui du radiologue. Cette étude montre que les performances de SmartUrgences® étaient fiables et stables avec une sensibilité de 92,8% et une spécificité de 85,6%, quel que soit l'âge du patient aux urgences. De plus, la VPP indique que dans 77,4% il existe réellement une fracture,

lorsque l'intelligence artificielle en montre une. La VPN quant à elle indique qu'il n'y a pas de lésions lorsque l'IA n'en détecte pas dans 95,7% des cas. Par ailleurs, l'IA a démontré une précision similaire dans la détection des anomalies, qu'elles se trouvent au niveau des os ou des poumons.

Cette étude a donc démontré le potentiel de Smarturgences® pour aider le médecin urgentiste dans le diagnostic, lorsqu'une radiographie pulmonaire ou osseuse est réalisée.

Au CHV, 9,5% des diagnostics posés sont des diagnostics de traumatologie.

Comme décrits précédemment, lors du parcours patient, si l'interprétation des clichés à distance par le radiologue expert ne corrobore pas avec celle du médecin urgentiste, une « discordance » est alors mise en évidence et le MAO peut gérer de plusieurs façons cette inadéquation.

Une comparaison des chiffres de taux de discordance avant et après IA a été réalisée.

L'étude avant IA a porté du 1er janvier 2021 au 31 octobre 2021.

- En 2021, il y a eu 68288 passages aux urgences tout secteur confondu. Nous avons eu 154 discordances radiologiques soit 0,225%, dont 136 portaient sur des patients ambulatoires (les 18 patients restants étant des patients hospitalisés).
- En 2022 il y a eu 72842 passages aux urgences (+6.6% d'activité). Sur la même période en 2022 de janvier à fin octobre, nous avons eu 116 discordances radiologiques, soit 0,159%. Nous avons donc eu 31% de réduction de discordances. Si on compare à activité équivalente à 2021, cela aurait donné une projection de 108 discordances aux urgences. Sur les 116 cas, 106 discordances concernaient des patients ambulatoires (10 patients hospitalisés).

La répartition des 136 discordances en 2021 et des 106 discordances en 2022 est décrite dans le tableau ci-dessous.

Nous considérons que l'adaptation thérapeutique en externe correspond à la modification du traitement antalgique et la remise d'une ordonnance d'attelle pour un retour à domicile dans les suites.

	Reconvocation aux urgences	Adaptation thérapeutique en externe	Convocation vers un spécialiste ou examen programmé	Classé sans suite	Total
2021	26 cas (19,1%)	27 cas (19,9%)	40 cas (29,4%)	43 cas (31,6%)	136
2022	10 cas (9,4%)	21 cas (19,9%)	19 cas (17,9%)	56 cas (52,8%)	106

Tableau 1 : Répartition de la prise en charge des discordances 2021/2022.

En 2023, nous avons comptabilisé 75520 passages, soit 2,6% de plus que 2022, et 162 discordances de janvier à fin octobre soit 0,214%.

Pour comparer à activité équivalente à 2021, projection de 146 discordances en 2023 versus 154 en 2021.

Nous constatons donc une diminution du nombre de discordances radiologiques depuis l'implantation du logiciel d'IA fin 2021.

De nombreux logiciels différents proposent cette aide diagnostique tels que BoneView™, Rayvolve ®, et ont été comparés dans une étude de 2023 (27). Les performances retenues sont à peu près équivalentes quel que soit le logiciel utilisé.

Au vu du flux continu de patients consultant aux urgences, de la non-disponibilité des experts 7j/7 et 24h/24h pour l'interprétation des radiographies, l'existence et la mise à disposition de logiciels performants d'IA tels que Smarturgences® a-t-elle une réelle influence sur la décision thérapeutique de l'urgentiste face à son patient à l'instant T ?

L'objectif principal de cette thèse est d'évaluer le taux de différence de diagnostic entre l'IA et l'urgentiste lorsque l'IA est en accord avec l'avis du radiologue, lors de diagnostic de lésions ostéoarticulaires chez les patients consultant aux urgences du CHV pour un traumatisme du squelette appendiculaire.

Le *Gold Standard* est le diagnostic du radiologue lors de l'analyse des radiographies à distance.

Le critère de jugement principal est le pourcentage des cas où l'IA et l'urgentiste sont en désaccord, parmi les cas où l'IA avait raison. Nous considérons que l'IA a raison si elle n'a rien détecté et que l'expert non plus, ou si elle a détecté une lésion avec doute ou certitude et que le radiologue a détecté lui aussi une lésion.

MATERIEL ET METHODE

I. TYPE D'ETUDE

Il s'agit d'une étude monocentrique, observationnelle et descriptive, réalisée aux urgences adultes du Centre Hospitalier de Valenciennes. Elle a été réalisée par recueil de données rétrospectif, du 1er décembre 2023 au 1er mai 2024.

II. CRITERES D'INCLUSION ET D'EXCLUSION

Nous avons inclus 480 patients, ayant bénéficié de radiographies d'une ou plusieurs parties du squelette appendiculaire, mettant en évidence ou non une lésion, lors de leur passage aux urgences.

Parmi ceux-ci, 126 patients ont été exclus de l'étude car nous n'avons pas pu récupérer les comptes rendus des experts : 119 dont l'identité du patient manquait, et 7 où les radiologues n'ont pas édité de compte rendu.

A. Critères d'inclusion

- Adultes consultant aux urgences adultes du CHV soit > 16 ans
- Suspicion de lésion du squelette appendiculaire
- Radiographie ostéoarticulaire réalisée aux urgences
- Questionnaire rempli par les médecins lors de la prise en charge du patient
- Analyse de la radiographie par un radiologue

B. Critères exclusion :

- Radiographie du rachis ou radiographie thoracique
- Défaut d'analyse des radiographies par un radiologue
- Défaut d'identité reportée sur le questionnaire
- Population pédiatrique soit patient < 16 ans

III. POPULATION ETUDIEE

Les patients dont les données ont été étudiées sont des patients consultant aux urgences adulte du CHV entre le 1er décembre et le 1er mai 2024, dont l'âge est supérieur à 16 ans, pour qui une radiographie du squelette appendiculaire a été réalisée et pour qui un questionnaire a été rempli par le professionnel médical prenant en charge le patient.

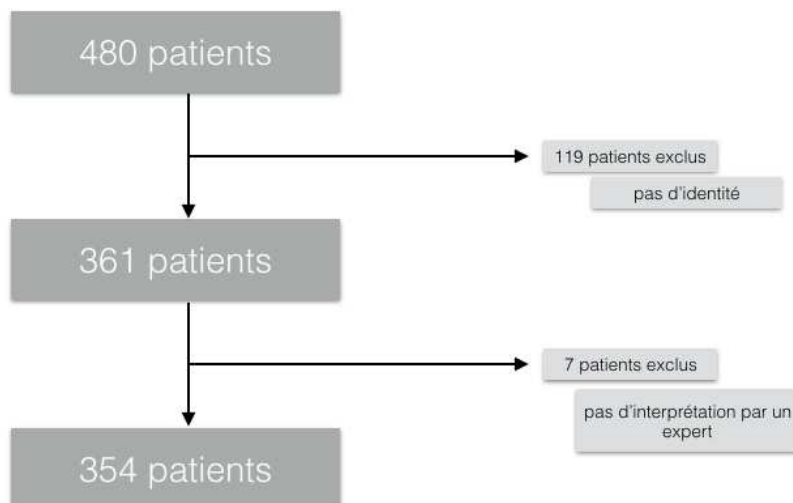


Figure 1 : Flow chart inclusion

IV. OBJECTIFS DE L'ETUDE

A. Objectif principal

L'objectif principal de cette étude est d'évaluer le taux de différence de diagnostic entre l'IA et l'urgentiste, lorsque l'IA a raison, c'est-à-dire lorsque l'IA est en accord avec l'avis du radiologue.

B. Objectifs secondaires

Les objectifs secondaires sont :

- De comparer le taux de suivi de l'IA en fonction de la véracité de son diagnostic
- De rechercher les facteurs associés au suivi de l'IA
- De comparer le taux de modification de prise en charge à la suite de l'avis de l'IA, entre les cas où l'IA avait raison et les cas où l'IA avait tort.
- De rechercher les facteurs associés à l'erreur de l'urgentiste alors que l'IA avait raison.
- D'analyser le ressenti des médecins en fonction de leur statut sur l'apport de l'IA dans ce cas précis.

V. CRITERE DE JUGEMENT

Le *gold standard* permettant de déterminer la présence ou non d'une lésion est le diagnostic posé par le radiologue à la suite de son interprétation de la radiographie, rendu à distance de l'admission aux urgences et donc de la prise en charge du patient.

Nous considérons que l'IA a détecté une lésion si elle l'a signalée avec un doute ou avec certitude.

A. Principal

Le critère de jugement principal est le pourcentage des cas où l'IA et l'urgentiste sont en désaccord, parmi les cas où l'IA avait raison.

On considère que l'IA a raison si elle n'a rien détecté et l'expert non plus, ou si elle a détecté une lésion avec doute ou certitude et que le radiologue a détecté lui aussi une lésion.

B. Secondaires

- Nous avons analysé le taux de suivi de l'IA en fonction de la véracité de son diagnostic.
- Nous avons réalisé une analyse multivariée afin de rechercher les facteurs associés au suivi de l'IA.
- Nous avons comparé le taux de modification de prise en charge à la suite de l'avis de l'IA, entre les cas où l'IA avait raison et les cas où l'IA avait tort.
- Nous avons réalisé une seconde analyse multivariée pour étudier les facteurs associés à l'erreur de l'urgentiste alors que l'IA avait le bon diagnostic.
- Nous avons analysé le ressenti des médecins en fonction de leur statut sur l'apport de l'IA dans ce cas précis par la question de satisfaction à la fin du questionnaire.

VI. MODE DE RECUEIL DES DONNEES

Nous avons récolté les données grâce à un questionnaire papier (annexe 1), destiné aux médecins thésés, internes, FFI et externes prenant en charge des patients consultant pour traumatisme des membres aux urgences adultes du CHV.

Ce questionnaire visait à reproduire la démarche diagnostique du médecin lors de la prise en charge de ces patients.

A distance, nous avons récupéré le compte rendu du radiologue et fait une corrélation entre le diagnostic posé par le médecin urgentiste avec son interprétation du cliché et le compte rendu de l'expert.

VII. AUTRES DONNEES RECUEILLIES

Les données recueillies peuvent être divisées en plusieurs parties.

Tout d'abord, une partie épidémiologie (identité du prescripteur et localisation de la fracture), puis la partie démarche diagnostique, et pour finir une question sur l'intérêt des médecins quant à l'utilité de l'IA dans ce cas précis.

A distance, nous avons récupéré les résultats de l'interprétation des clichés radiologiques par le radiologue, expert en la matière, grâce aux logiciels Millenium® et Firsnet®.

La totalité des données a ensuite été pseudonymisée puis anonymisée pour permettre son étude.

VIII. MODELE STATISTIQUE

Tout d'abord, nous avons réalisé les analyses statistiques avec le logiciel Python. Nous avons réalisé une série de statistiques descriptives concernant la base de données, puis des analyses multivariées.

Deux modèles sont couramment utilisés pour analyser les variables binaires : le modèle *logit* (régression logistique) ou son alternative le modèle *probit*. Nous avons choisi d'utiliser le modèle *logit*.

Dans le modèle *logit*, nous ne pouvons pas avoir des variables entièrement dépendantes. Nous considérons donc que les variables indépendantes doivent être parfaitement non corrélées entre elle. Nous avons donc étudié leur corrélation pour nous assurer que nous sommes bien en face de variables indépendantes.

Pour résumer, le modèle statistique traite 354 variables, qui finalement, pour notre étude est peu. Il s'attache à décrire l'effet de variables explicatives, binaires, sur la divergence entre le diagnostic posé par l'urgentiste et celui posé par l'IA qui est en accord avec celui de l'expert, et du suivi de l'avis de l'IA dans la prise en charge des traumatismes du squelette appendiculaire aux urgences de Valenciennes.

IX. SUR LE PLAN ETHIQUE

Les patients ont pu être informés de l'utilisation de leurs données médicales par différentes modalités : pancartes en salle d'attente et à l'accueil des urgences, livret

d'accueil remis aux patients lors de l'hospitalisation si tel a été le cas et sur le site internet du CHV. (Annexe 3).

Cette étude n'a pas impliqué la personne humaine. La méthodologie se base sur des données hors loi Jardé (MR004). Une déclaration auprès de la Commission Nationale de l'Informatique et des Libertés (CNIL) n'a pas été nécessaire.

L'étude a fait l'objet d'une déclaration auprès de l'URC du CHV, et le protocole de recherche a été validé. Elle a ensuite été présentée à la Commission Éthique de la Recherche Clinique (CERCL) de l'hôpital, rendant un avis favorable pour la réalisation de cette étude.

Le numéro d'enregistrement au registre interne des traitements des données de l'établissement du CHV est le CHV-2024-019. (Annexe 2)

RESULTATS

I. ÉPIDEMIOLOGIE DESCRIPTIVE

A. Catégorie médicale prenant en charge le patient

Le patient a été pris en charge dans 216 cas par un interne (61%), 94 fois par un urgentiste (26,6%), 23 fois par un FFI (6,5%), et 21 fois par un externe (5,9%).

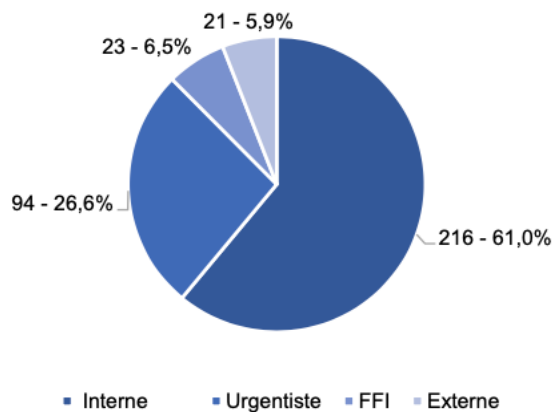


Figure 2 : Catégorie médicale prenant en charge le patient

B. Siège de la lésion

Sur les 354 radiologies réalisées, les localisations se répartissent comme suit :

- poignet : 55 cas soit 15,5%
- genou : 54 cas soit 15,3%
- fémur (incluant extrémité supérieure qui est le plus courant, extrémité inférieure et diaphyse) : 46 cas soit 13%
- cheville : 45 cas, soit 12,7%

- humérus : 38 cas soit 10,7%
- pied : 36 cas soit 10,2%
- main : 36 cas soit 10,2%
- coude : 16 cas soit 4,5%
- doigt : 13 cas soit 3,7%
- jambe (incluant les extrémités et la diaphyse) : 6 cas soit 1,7%
- avant-bras : 5 cas soit 1,4%
- orteils : 4 cas soit 1,1%

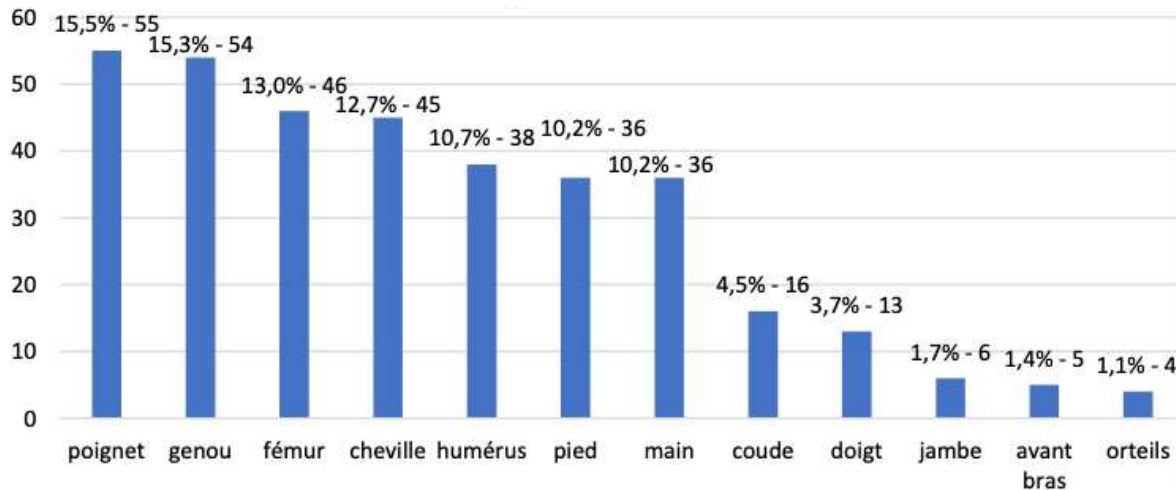


Figure 3 : Siège de la lésion

C. Matériel d'ostéosynthèse

Lors des études des radiographies, il a été demandé de notifier la présence de matériel d'ostéosynthèse. En effet, la présence de matériel peut modifier

l'interprétation de la radiographie par l'intelligence artificielle, mais aussi par les médecins.

Dans 332 cas, soit 93,8%, il n'y avait pas de matériel d'ostéosynthèse, alors que dans 22 cas, soit 6,2%, on notait la présence de matériel.

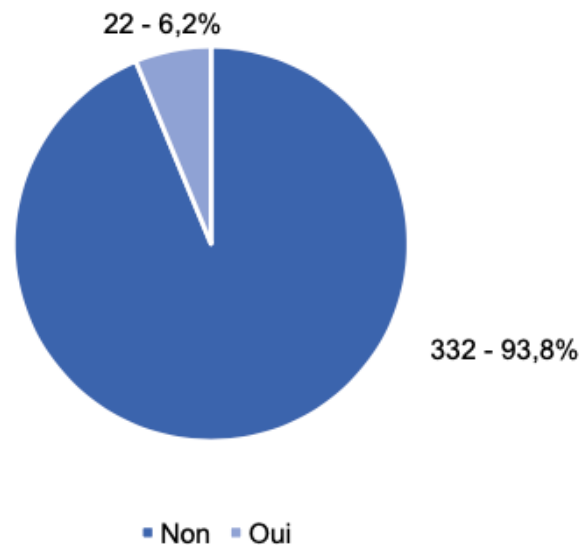


Figure 4 : Présence de matériel d'ostéosynthèse

D. Compte rendu de l'expert

Notre *gold standard*, donné par le compte rendu du radiologue à distance retrouve une lésion dans 135 cas, soit 38,1%. Dans 219 cas, il n'y a pas de lésion retrouvée par le radiologue (soit 61,8%).

Dans 118 cas, une fracture est diagnostiquée, 4 cas le diagnostic est une association fracture et luxation, 5 cas où c'est une luxation simple, 7 cas où on retrouve un

épanchement intra articulaire et 1 cas où il est retrouvé un épanchement intra articulaire avec une fracture associée.

E. Diagnostics de l'urgentiste et de l'IA

D'une part, nous avons constaté que le médecin a diagnostiqué une lésion dans 88 cas soit 24,9%, a eu un doute dans 46 cas soit 13% et n'a pas diagnostiqué de lésion dans 220 cas.

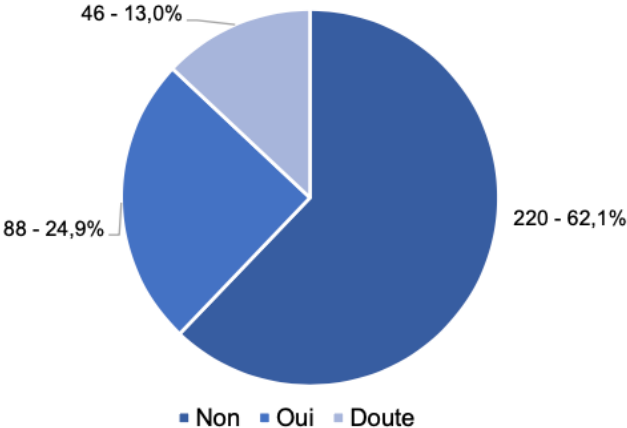


Figure 5 : Lésion selon l'urgentiste prenant en charge le patient.

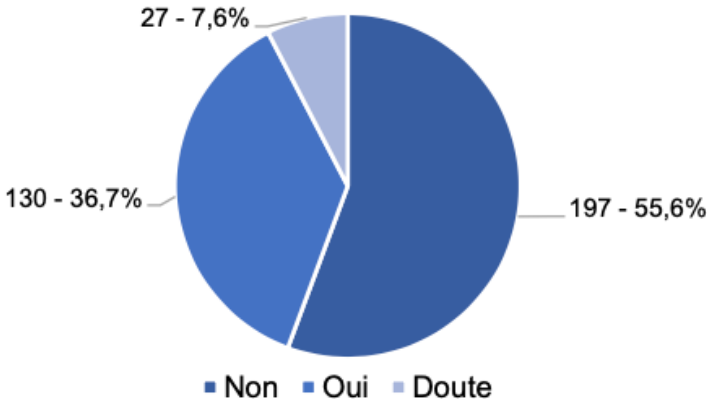


Figure 6 : Lésion selon l'IA

D'autre part, l'IA a diagnostiqué une lésion dans 130 cas (36,7%), a eu un doute dans 27 cas (7,6%), n'a pas diagnostiqué de lésion dans 197 cas (55,6%).

Nous voyons que dans 283 cas (79,9%) les diagnostics des médecins et de l'IA corroborent, mais que dans 71 cas (20,1%) ils sont en désaccord.

Pourtant, dans 316 cas (89,3%), l'avis de l'IA est suivi pour la suite de la prise en charge, mais n'est pas suivi dans 38 cas (10,7%).

F. Avis complémentaire

Lors de la prise en charge du patient, le médecin urgentiste a eu la possibilité de demander des avis complémentaires à plusieurs personnes différentes : à un chirurgien (69 cas soit 19,5%), à un radiologue (24 cas soit 6,8%), au sénior des urgences (28 cas soit 7,9%). Parfois plusieurs avis sont demandés en simultané : radiologue et chirurgien (11 cas soit 3,1%), sénior des urgences et radiologue (8 cas soit 2,3%), sénior des urgences et chirurgien (6 cas soit 1,7%), ou les 3 combinés (4 cas soit 1,1%).

Tous les avis demandés au sénior des urgences, ont été demandés par des étudiants (soit internes, soit FFI soit externes). Nous ne retrouvons pas de cas où un avis à un sénior a été demandé par un autre médecin urgentiste thésé considéré lui-même comme sénior.

Dans tous les autres cas, soit 204 cas (57,6%), personne n'a été sollicité pour donner d'avis complémentaire.

G. Examen complémentaire

Un examen complémentaire a pu aussi être demandé lorsque la radiographie seule n'était pas contributive. Il existe plusieurs cas de figure dans lesquels les examens sont demandés : lorsqu'il y a une discordance de l'humain et de l'IA, ou pour une aide à l'orientation thérapeutique (traitement orthopédique ou chirurgical, par exemple).

Nous avons recensé 34 demandes d'examen complémentaire (9,6%) : 29 scanners, 2 où un scanner et une IRM sont associés, une échographie instantanée, une échographie à distance, une IRM. Ces examens en sus permettent une prise en charge plus adaptée au patient.

Dans 320 cas (90,4%), il n'y a pas eu de réalisation d'examen complémentaire.

H. Modification des prises en charge

L'intervention de l'IA n'a pas modifié la prise en charge initiale de l'urgentiste dans 315 cas (89%).

Au contraire, 39 fois (11%), l'IA a permis une modification des prises en charge : 9 fois (2,6%) sans précision, 16 fois où elle a permis la mise en place d'une attelle (4,5%), 8 fois d'un plâtre (2,2%) et 6 fois (1,7%) où elle a permis une prise en charge chirurgicale.

II. CRITERE DE JUGEMENT PRINCIPAL

Le critère de jugement principal est le pourcentage des cas où l'IA et l'urgentiste sont en désaccord, parmi les cas où l'IA a raison.

Nous considérons que l'IA a raison si elle n'a rien détecté et que l'expert non plus, ou si elle a détecté une lésion avec doute ou certitude et que le radiologue a détecté lui aussi une lésion.

Il existe une discordance entre le diagnostic de l'urgentiste et celui de l'IA, avec l'expert en accord avec le diagnostic de l'IA lorsque nous avons les cas de figure suivants :

- Lésion non vue par l'urgentiste – Lésion vue par l'IA – Lésion vue par l'expert
- Lésion non vue par l'urgentiste – Lésion vue avec un doute par l'IA – Lésion vue par l'expert
- Lésion vue avec un doute par l'urgentiste – Lésion non vue par l'IA – Lésion non vue par l'expert
- Lésion vue par l'urgentiste – Lésion non vue par l'IA – Lésion non vue par l'expert

Au total : nous avons 30 cas de discordance entre l'urgentiste et l'IA, avec l'IA du même avis que l'expert, soit 8,5%.

De plus, dans 78,5% (278 cas), nous avons une concordance entre les diagnostics posés par le médecin en charge du patient, l'IA et le radiologue.

Dans 8,5% (30 cas), nous avons une discordance de l'avis des humains par rapport à celui de l'IA : c'est-à-dire que l'IA se trompe mais que l'urgentiste a le bon diagnostic car il est en adéquation avec celui du radiologue.

Dans 4,5% des cas, nous avons une discordance de l'urgentiste et de l'IA par rapport au radiologue.

III. CRITERES DE JUGEMENT SECONDAIRES

Nous avons regroupé des sous catégories entre elles pour limiter le nombre de données d'analyses :

- Concernant la localisation des lésions suspectées : les orteils, l'avant-bras, la jambe, les doigts et le coude sont considérés comme « Autre » car il n'y avait pas assez de variables pour réaliser des analyses représentatives.
- Concernant le type de lésions, hormis les fractures, toutes les autres lésions sont considérées comme « Autre », soit les luxations, les associations fracture et luxation, les épanchements intra articulaires, les associations fractures et épanchements.
- Concernant les avis complémentaires, ils ont tous été regroupés dans la même catégorie, qu'ils aient été donnés par un chirurgien, par un sénior des urgences, par un radiologue ou par plusieurs d'entre eux combiné.
- Les FFI sont considérés comme des internes.

A. Taux de suivi de l'IA

Nous avons analysé le taux de suivi de l'IA par le médecin prenant en charge le patient, en fonction des diagnostics posés par l'IA et par le radiologue.

Dans les résultats des analyses descriptives de nos questionnaires, nous voyons que dans 283 cas (79,9%), les diagnostics des médecins et de l'IA corroborent, mais que dans 71 cas (20,1%) ils sont en désaccord.

Pourtant, dans 316 cas (89,3%), l'avis de l'IA est suivi pour la suite de la prise en charge, mais n'est pas suivi dans 38 cas (10,7%).

Nous pouvons voir que l'IA a raison dans 308 cas (87%) et l'IA a tort dans 46 cas (13%). L'IA a raison si son diagnostic est identique à celui du radiologue et a tort si les deux diagnostics diffèrent.

Ainsi, après la réponse donnée par le compte rendu de l'expert, nous avons pu remarquer que l'IA est suivie 296 fois lorsqu'elle a raison (96,1%) et 20 fois lorsqu'elle a tort (43,5%).

L'IA, au contraire, n'est pas suivie alors qu'elle a raison, c'est à dire que l'IA est non suivie à tort, dans 12 cas (3,9%). Elle n'est pas suivie à raison 26 fois (56,5%) c'est-à-dire qu'elle n'est pas suivie et que son diagnostic est incorrect.

	IA raison (308)	IA a tort (46)
IA est suivie (316)	296 cas (96,1%)	20 cas (43,5%)
IA non suivie (38)	12 cas (3,9%)	26 cas (56,5%)

Tableau 2 : Suivi de l'IA

Pour éclaircir, nous considérons que le suivi de l'IA à tort par le médecin durant la prise en charge, ou le non suivi à raison de l'IA revient à dire que finalement l'IA et

l'expert ont un diagnostic différent (soit les cas de « discordance humain vs IA » et les cas de « discordance IA + urgentiste vs expert »).

Au contraire, le suivi de l'IA a raison ou le non suivi de l'IA à tort, revient à dire que l'IA et l'expert ont un avis similaire (soit les cas de « concordance » et les cas de « discordance urgentiste vs IA + expert »).

B. Rechercher les facteurs associés au suivi de l'IA

Nous cherchons à expliquer la variable « suivi de l'IA », *via* différentes variables explicatives.

Dans cette analyse multivariée, nous étudions les variables associées au suivi de l'IA. Peu importe le type de lésion diagnostiqué, ces variables sont associées au suivi de l'IA. En effet leurs p-value sont inférieures à 0,05.

De plus, si l'IA a raison, cela est significativement associé au suivi de l'IA par le médecin prenant en charge le patient.

Nous remarquons aussi que le fait que le médecin ait un doute n'est pas significativement associé au suivi de l'IA par ce dernier. En revanche, si le médecin ne voit pas de lésion, cela est associé au suivi de l'IA (p-value < 0,05).

	Coefficient	P value	[0,025]	[0,975]
Urgentiste	- 0.806	0.519	-3.253	1.641
Interne	- 0.228	0.849	-2.578	2.122
Doute sur une lésion selon le médecin	0.6079	0.545	-1.359	2.575
Pas de lésion selon le médecin	2.2512	0.032	0.188	4.315
Lésion autre que fracture selon radiologue	2.8032	0.023	0.383	5.223
Fracture selon radiologue	3.4716	0.001	1.384	5.559
Lésion au niveau de la cheville	-1.3863	0.185	-3.434	0.662
Lésion au niveau du fémur	-1.4193	0.192	-3.549	0.711
Lésion au niveau genou	0.4104	0.747	-2.087	2.907
Lésion au niveau humérus	-0.5627	0.618	-2.773	1.648
Lésion au niveau main	1.2407	0.388	-1.576	4.058
Lésion au niveau pied	-0.5790	0.626	-2.905	1.747
Lésion au niveau poignet	-0.7729	0.465	-2.846	1.300
IA a raison	4.1952	0.000	3.057	5.333

Tableau 3 : Variables associées au suivi de l'IA.

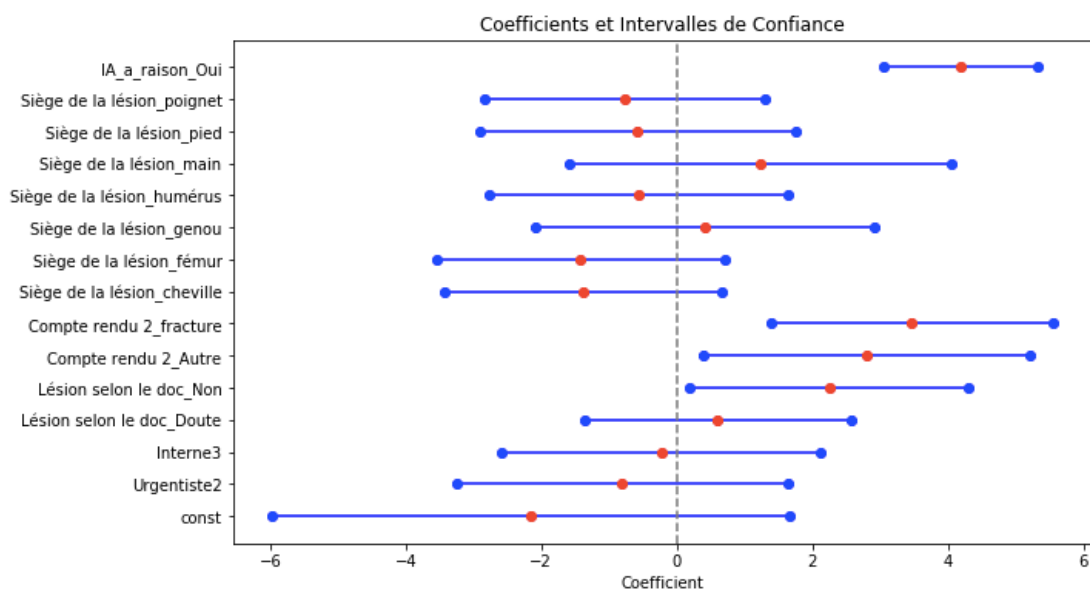


Figure 7 : Forest Plot des variables associées au suivi de l'IA

C. Comparaison du taux de changement d'avis de l'urgentiste

Nous avons analysé les taux de changement d'avis de l'urgentiste, puis les avons comparés entre le groupe de patient pour lequel l'IA avait raison et pour lequel l'IA avait tort.

Dans les données que nous avons analysées, nous avons souligné que l'urgentiste a changé 34 fois d'avis, et n'a pas changé d'avis dans 320 cas.

Nous définissons le changement d'avis de l'urgentiste comme le suivi de l'IA par ce dernier alors que leurs diagnostics initiaux étaient différents.

Le taux de changement d'avis de l'urgentiste grâce à l'IA, c'est-à-dire quand l'IA avait raison est de 8,8% (27 cas).

Le taux de changement d'avis de l'urgentiste alors que l'IA a tort est de 15,2%.

Le médecin prenant en charge le patient a ainsi changé d'avis 7 fois à cause de l'IA alors que le logiciel n'avait pas posé le bon diagnostic.

	IA a raison (308)	IA a tort (46)
Changement avis (34)	27 (8,8%)	7 (15,2%)
Pas de changement d'avis (320)	281 (91,2%)	39 (84,8%)

Tableau 4 : Changement d'avis de l'urgentiste

D. Comparaison du taux de modification de prise en charge

Nous avons analysé les taux de modification de prise en charge décidée par le médecin, entre les cas où l'IA avait raison et les cas où l'IA avait tort.

Il y a eu 39 modifications de prise en charge sur 354 dossiers analysés (11%).

Dans 2 cas (4,3%), nous observons une modification de la prise en charge en sachant que l'IA a tort. Au contraire, dans 37 cas (12%), nous observons une modification de prise en charge, en sachant que l'IA a raison.

	IA a raison (308)	IA a tort (46)
Modification de PEC (39)	37 (12%)	2 (4,3%)
Pas de modification de PEC (315)	271 (88%)	44 (95,7%)

Tableau 5 : Modification de la prise en charge en fonction du diagnostic de l'IA

Nous avons réalisé des analyses complémentaires, en prenant en compte la concordance ou la discordance du diagnostic posé par l'urgentiste et celui posé par l'IA, pour analyser plus finement la modification de la prise en charge.

	Diagnostic initial différent		Diagnostic initial identique	
	IA tort	IA raison	IA tort	IA raison
Modification de PEC	1	22	1	15
Pas de modification de PEC	32	16	12	235

Tableau 6 : Modification de prise en charge en fonction du résultat initial de l'urgentiste et de l'IA

Nous avons étudié les cas où les diagnostics initiaux n'étaient pas les mêmes entre celui de l'urgentiste et celui du logiciel d'IA.

- Dans un cas isolé, l'urgentiste a modifié sa prise en charge, alors que nous avons su à posteriori que l'IA avait tort dans son diagnostic.
- Dans 22 cas (6,2%), nous avons relevé des modifications de prise en charge vers une thérapeutique adéquate.
- Au contraire, lors de 16 prises en charge, il n'y a pas eu de modification de prise en charge, alors que l'IA avait le bon diagnostic.
- Dans 32 cas, il n'y a pas eu de modification de prise en charge en sachant que l'urgentiste avait le diagnostic correct.

Ensuite, nous avons étudié les cas où les diagnostics initiaux étaient les mêmes :

- Dans 12 cas, en apprenant par la suite que les deux diagnostics ne correspondaient pas à celui de l'expert, il n'y a pas eu de modification de prise en charge.

- Dans 1 cas, étant passé à côté du bon diagnostic, la prise en charge a tout de même été modifiée, et probablement de part les avis complémentaires demandés.
- De plus, alors que le médecin et l'IA avaient tous les deux le bon diagnostic, il y a eu 15 modifications de la prise en charge.

E. Recherche des facteurs associés à l'erreur de l'urgentiste

Nous avons réalisé une analyse statistique multivariée, sur un modèle de régression logistique pour rechercher les facteurs associés à l'erreur de l'urgentiste alors que l'IA avait le bon diagnostic.

Plusieurs variables exercent une influence sur la différence de diagnostic entre l'urgentiste et celui de l'IA.

Tout d'abord la nature des lésions exerce une influence positive. En effet, si la lésion suspectée est une fracture, elle a un effet positif sur les chances d'observer une discordance de diagnostic (p value = 0,000).

De plus, certaines localisations exercent une influence. En particulier, si la lésion se situe au niveau de la cheville, celle-ci exerce une influence positive sur le risque d'erreur de l'urgentiste (p value = 0,002). De même pour une lésion au niveau du fémur (p value = 0,003), du genou (p-value = 0,028) de la main (p value = 0,02) et du pied (p value = 0,006).

Le fait de demander un avis complémentaire exerce un effet positif. Cela veut dire que le médecin n'était pas sûr de lui au départ, n'avait statistiquement pas la bonne

réponse, et a dû demander un avis complémentaire. Quand on demande un avis complémentaire, on a plus de risque de s'être trompé initialement que si on n'en demande pas.

Le fait d'avoir un doute initialement, exerce un effet positif sur la discordance de diagnostic. En effet, quand on n'est pas sûr de soi, cela montre qu'on a statistiquement plus de risque de se tromper et cela a un effet sur la différence de diagnostic.

De plus, le fait d'être un urgentiste, est significativement associé au risque d'erreur.

Ces variables ont un effet sur la discordance et les coefficients associés à ces dernières nous indiquent qu'elles exercent une influence positive

	Coefficient	P-value	[0.025	0.975]
Avis complémentaire demandé	1.8346	0.000	1.144	2.525
Urgentiste	1.8412	0.018	0.321	3,361
Interne	0.9305	0.205	-0.509	2.370
Doute selon le médecin	2.2595	0.000	1.593	2.926
Fracture selon le radiologue	0.7342	0.019	0.120	1.349
Lésion au niveau de la cheville	2.7494	0.000	1.441	4.058
Lésion au niveau fémur	2.1110	0.001	0.825	3.397
Lésion au niveau du genou	1.6004	0.028	0.172	3.029
Lésion au niveau humérus	1.1593	0.108	-0.253	2,572
Lésion au niveau main	1.7903	0.011	0.414	3.166
Lésion au niveau pied	2.2573	0.001	0.892	3.622
Lésion au niveau poignet	1.0776	0.115	-0.262	2,417

Tableau 7 : Variables associées à l'erreur des urgentistes

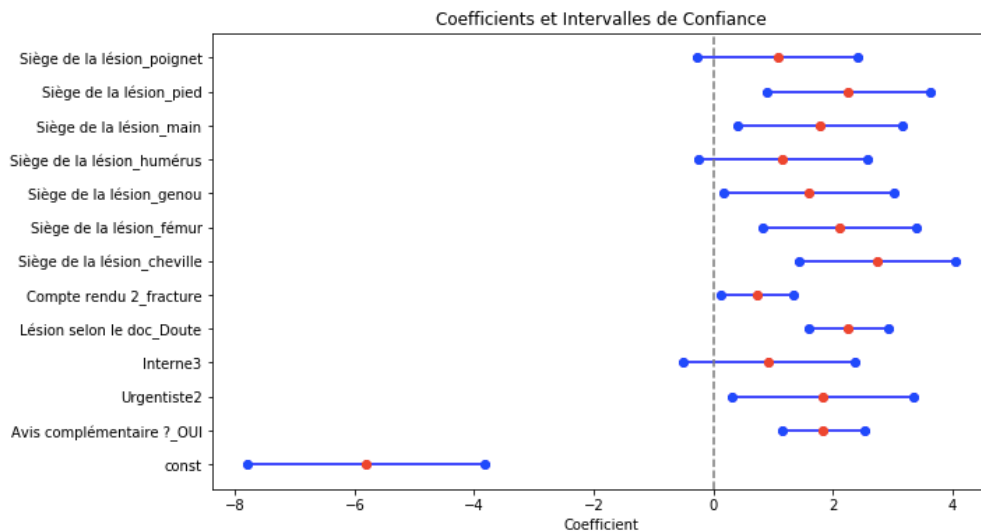


Figure 8 : Forest plot des variables associées à l'erreur de l'urgentiste

F. Analyse du ressenti des médecins

Nous avons pu analyser le ressenti de chaque médecin grâce à une question quant à l'intérêt de l'intelligence artificielle dans le diagnostic de la lésion ostéoarticulaire dans ce cas précis, allant de 0 à 10. La moyenne du ressenti des médecins est de 5,08 (+/- 2,98). La médiane est 5 (2 ; 8).

Nous avons calculé les moyennes par type d'expérience médicales, ainsi que les écarts type correspondants.

- La moyenne des ressentis exprimés par les externes est 6,52 +/- 2,73.
- Les FFI ont un ressenti moyen à 4 +/- 2,61.
- Les internes, tout statut confondu, ont un ressenti en moyenne à 5,9 +/- 2,59. Les internes en stage, en phase socle ont une moyenne à 5,97 +/- 2,54.

- Les urgentistes ont une moyenne à 3,16 +/- 2,98.
- Les docteurs juniors (DJ) évaluent l'IA en moyenne à 1,57 +/- 3,23.
- Les médecins de moins de 5 ans d'expérience ont un ressenti estimé en moyenne 3,44 +/- 3,29.
- Les médecins entre 5 et 10 ans d'expérience ont quant à eux un ressenti à 4,25 +/- 2,63 en moyenne
- Les médecins entre 10 et 20 ans d'expérience ont un ressenti à 2,5 +/- 1,73.
- Les médecins de plus de 20 ans d'expérience ont un ressenti à 2,17 +/- 2,06.

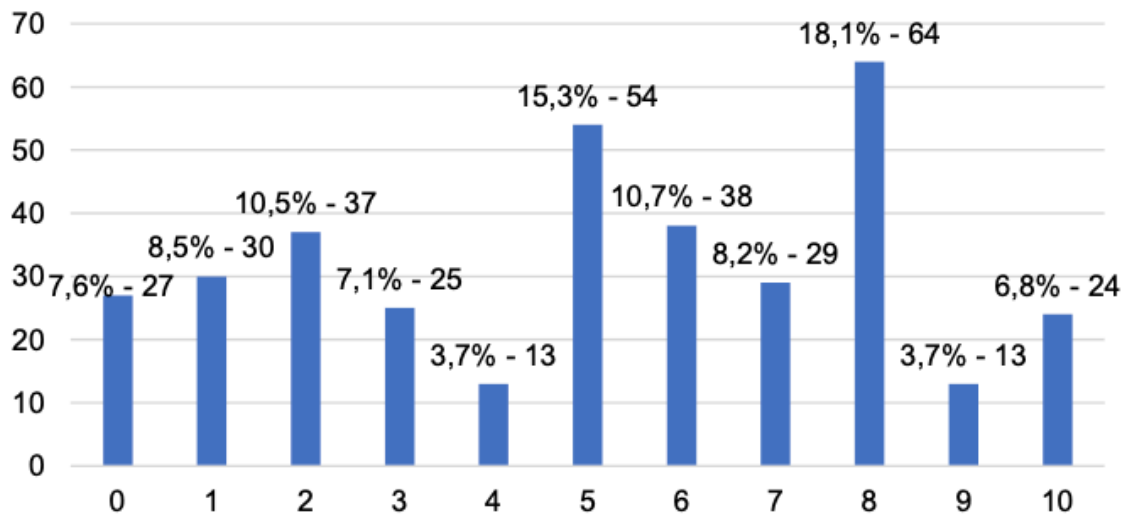


Figure 9 : Échelle de ressenti concernant l'IA

DISCUSSION

Pour commencer, l'étude de 2020 de Milvue® (4), avait mis en évidence le potentiel de SmartUrgences® pour aider les médecins urgentistes dans le diagnostic des pathologies traumatologiques des patients consultant aux urgences. Elle avait comparé les diagnostics posés respectivement par les urgentistes, l'intelligence artificielle et les radiologues sur un même cliché. A partir des réponses obtenues, ils avaient pu démontrer les performances diagnostiques de l'intelligence artificielle et prouver que ce logiciel d'IA établissait un diagnostic de très bonne qualité. La sensibilité, la spécificité, la valeur prédictive positive et la valeur prédictive négative du modèle atteignaient 92,8%, 85,6%, 77,4%, et 95,7%. De plus, lorsque les urgentistes et le logiciel d'IA présentaient des avis contradictoires, celui de l'IA était confirmé par le radiologue dans 90,1% des cas. La finalité de ce logiciel est donc, en collaboration étroite avec l'urgentiste, de garantir un diagnostic le plus juste possible afin de ne pas nuire à la prise en charge de la santé du patient.

Notre étude consiste à évaluer l'influence d'un logiciel d'intelligence artificielle dans le processus décisionnel des médecins urgentistes lors de la prise en charge de patients aux urgences. Nous avons ainsi cherché à démontrer la valeur ajoutée et l'efficacité de l'IA dans le diagnostic et l'interprétation des radiographies pour les patients se présentant aux urgences à la suite d'un traumatisme de membre. Les médecins urgentistes recherchent avant tout de la fiabilité lors de l'utilisation d'un logiciel de détection clinique et notre étude est, à notre connaissance, la première à

se pencher sur un contexte clinique réel. A noter que la référence absolue de notre étude, est le diagnostic final posé par le radiologue.

Au sein des 354 données que nous avons récoltées, nous avons analysé les comptes rendus de l'expert afin d'obtenir notre *gold standard* que nous avons ensuite comparé aux diagnostics des médecins urgentistes et de l'intelligence artificielle. D'abord, le radiologue a diagnostiqué une lésion dans 135 cas (38,1%), tandis que dans 219 cas (61,8%) aucune lésion n'a été repérée.

Ensuite, les médecins urgentistes ont diagnostiqué de leur côté une lésion dans 88 cas (24,9%) et l'absence de lésion dans 220 cas (62,1%). Pour les 46 cas restants (13%), les médecins ont émis un doute sur l'existence ou non d'une lésion. Enfin, le logiciel d'intelligence artificielle a diagnostiqué une lésion dans 130 cas (36,7%), n'a pas diagnostiqué de lésion dans 197 cas (55,6%) ou a eu un doute dans 27 cas (7,6%). A partir de cette première comparaison nous pouvons d'ores et déjà constater que l'intelligence artificielle a diagnostiqué plus de lésions que les urgentistes avec une plus grande certitude. A ce stade, deux hypothèses sont les plus probables : un surdiagnostic de lésion par l'IA, ou alors un défaut de diagnostic par les médecins urgentistes.

Pour discuter de cela, il faudrait analyser des performances de l'urgentiste seul pour la lecture des radiographies, ce qui n'était pas l'objectif de cette étude.

I. CRITERE DE JUGEMENT PRINCIPAL

Dans notre étude, nous avons relevé 30 cas de divergence entre le diagnostic de l'urgentiste et celui l'IA, dans le cas où le celui de de l'IA concordait avec celui du radiologue. Autrement dit, dans 8,5% des cas l'urgentiste a commis une erreur que l'IA aurait évité. Ces erreurs humaines peuvent être d'origine multiple, et dans la plupart des cas, c'est une lésion qui passe inaperçue. L'objectif serait que l'IA puisse rectifier cette erreur pour permettre la décision thérapeutique adéquate pour le patient lors de sa prise en charge. Pour que l'urgentiste puisse faire confiance à l'IA, qu'il puisse se fier à son diagnostic initial et qu'il puisse changer d'avis et ainsi modifier la prise en charge, il faut que le logiciel soit performant.

De plus, nous avons constaté 16 cas de discordance entre l'avis de l'urgentiste et de l'IA (qui sont similaires) et celui du radiologue (4,5%).

A noter que ces discordances ne veulent pas dire que la prise en charge finale n'a pas été la bonne. L'étude porte en effet sur le début de l'algorithme décisionnel. Au cours de la prise en charge, des avis complémentaires ont pu être demandés et ont pu permettre de rectifier et de proposer la thérapeutique adéquate au patient. Ces deux situations-là ont également pu déclencher des alertes, à distance, après la prise en charge du patient et son retour à domicile entraînant une nouvelle prise en charge par le MAO. En effet, si les diagnostics sont différents entre l'urgentiste et le radiologue interprétant la radiologie à postériori, la discordance est notifiée aux secrétaires du service puis rectifiée par le MAO.

Enfin, dans 30 cas (8,5%), nous avons observé une discordance de l'avis des humains et par rapport à celui de l'IA. Autrement dit, l'IA s'est trompée mais l'urgentiste a posé le même diagnostic que celui de l'expert. Ce type de situation survient principalement lorsque l'IA détecte des images construites, des séquelles de fractures, des variations anatomiques, ou que la lésion détectée ne correspond pas à la situation clinique. En effet, l'urgentiste bénéficiant de la connaissance de la clinique du patient, peut analyser l'image de manière plus précise en tenant compte de la localisation de la douleur, du mécanisme du traumatisme, des antécédents personnels du patient. L'IA peut avoir du mal à capter les nuances que seul l'œil humain peut percevoir. Il est donc essentiel que le diagnostic soit basé sur une analyse conjointe de l'image radiographique et des données cliniques. C'est la principale limite de ce logiciel d'IA. Les raisonnements diagnostics de l'IA et de l'humain ont été expliqués dans une étude de 2023, nous permettant de mieux comprendre leurs similitudes et leurs différences. (28)

En somme, nous mettons en évidence que l'IA et l'urgentiste se trompent dans des proportions similaires dans notre étude. Un plus grand nombre de données nous permettrait d'analyser plus en profondeur les taux d'erreurs de l'urgentiste et celui de l'IA.

Enfin, dans 78,5% des cas, soit 278 dossiers, nous avons une concordance entre les diagnostics du médecin, de l'IA et du radiologue, indiquant dans ces cas, que le diagnostic est unanime dès le départ et confirmé par l'expert, à distance. Ce taux de concordance est élevé mais reste perfectible.

Nous nous sommes interrogés sur la capacité de l'IA à rester performante en dehors des conditions contrôlées des études cliniques ainsi que sur son adoption par les médecins urgentistes dans la prise de décision.

II. CRITERES DE JUGEMENT SECONDAIRES

A. Taux de suivi de l'IA

D'après l'analyse de notre jeu de données, nous avons observé que l'avis de l'IA est suivi dans 316 cas (83,9%), et n'est pas suivi dans 38 cas (16,1%).

De plus, elle est suivie à raison dans 296 cas (96,1%) mais « à tort » dans 20 cas (43,5%).

Le taux de suivi de l'IA à raison, c'est-à-dire pour rappel, que l'urgentiste a suivi l'avis de l'IA et que cette dernière avait un diagnostic correct, peut provenir de deux situations distinctes.

La première, qui est la plus fréquente dans notre étude, survient lorsque l'urgentiste et l'IA posent un diagnostic identique. L'urgentiste se sentant conforté par l'avis de l'IA, décide de suivre cette orientation. Il s'avère ensuite que les deux diagnostics étaient justes. Même si cela n'a pas été évalué objectivement, on peut envisager à ce stade la plus-value de l'IA pour l'urgentiste car ce logiciel jugé comme compétent a posé le même diagnostic que lui.

Le deuxième cas de figure que nous rencontrons, qui est le plus intéressant, survient lorsque l'IA a le bon diagnostic, que l'urgentiste se trompe, mais qu'après l'avis de l'IA il fait évoluer son avis pour finalement suivre celui du logiciel. L'urgentiste a pu se

décider à suivre l'IA pour plusieurs raisons : il n'avait initialement pas vu la lésion, ou alors il pensait qu'il y en avait une et l'IA a dit le contraire.

Dans ces deux cas, l'IA est une plus-value et un outil d'aide à la décision majeure pour le corps médical.

Dans les paramètres du logiciel,(4) nous retrouvons une VPP (soit la probabilité qu'il y ait réellement une lésion lorsque le résultat affiché est positif) de 77,4%, ce qui signifie qu'il y a une grande proportion de « faux positifs ». Cela pourrait nous amener à diminuer la confiance dans l'algorithme pour détecter les lésions du squelette ostéoarticulaire. A contrario, la VPN (soit la probabilité qu'il n'y ait réellement pas de lésion lorsque le résultat affiché est négatif) est de 95,7%. Cela signifie qu'il y a une forte probabilité que les patients sans lésion détectée n'aient effectivement pas de lésion. Ces indicateurs permettent de nous conforter sur la performance du test à confirmer l'absence de lésion. En résumé, grâce à cette VPN élevée, nous pouvons savoir dans quelle mesure nous fier à un résultat négatif décrit par l'IA pour exclure une lésion.

B. Facteurs associés au suivi de l'IA

Afin d'analyser les facteurs explicatifs qui influencent le suivi de l'IA par le médecin, nous avons réalisé une analyse multivariée à partir de plusieurs variables. Quel que soit le type de lésion diagnostiqué, ils sont tous significativement associés au suivi de l'IA. Il n'y a donc pas de différence que ce soit une fracture, un épanchement articulaire, une luxation ou deux d'entre eux associés.

Si l'IA a raison, cela est significativement associé au suivi de l'IA par le médecin prenant en charge le patient. Pour rappel, nous savons si l'IA a raison à posteriori. Cela dit, grâce à cette donnée, cela conforte le médecin à pouvoir suivre l'IA avant d'avoir la confirmation du diagnostic par l'expert.

Tout ceci peut être mis en corrélation avec les résultats des performances du logiciel : l'IA détecte une lésion à tort dans 23% des cas puisque la VPP est de 77% pour ce logiciel. Il ne faut donc pas suivre l'IA aveuglement. Le taux de réussite de l'IA n'est pas de 100%, et la décision finale appartiendra toujours au médecin.

De manière inattendue, nous découvrons que le fait que le médecin ait un doute n'est pas significativement associé au suivi de l'IA. Pourtant, nous aurions pu penser que lorsque le médecin doute il serait plus enclin à suivre l'analyse de l'IA. C'est ce à quoi l'IA pourrait servir à ce moment-là de la prise en charge. Peut-être que dans ces cas-là, le médecin va plutôt demander un avis complémentaire plutôt que se fier à l'IA. Nous attendons de l'IA qu'elle soit une aide au diagnostic et c'est au moment où le médecin doute qu'elle serait intéressante à consulter pour orienter la prise en charge. En effet, si cela peut confirmer un doute dans un sens ou dans un autre, cela aurait un impact positif non négligeable sur les prises en charge, la vitesse de diagnostic et probablement la diminution des demandes d'avis complémentaires.

Il serait intéressant d'étudier de nouveau les variables associées au suivi de l'IA avec des populations diverses telle que la population pédiatrique, et d'inclure les facteurs environnementaux tels que le nombre de passage aux urgences, l'heure de la prise en charge, la période de l'année, le nombre de patient consultant au même

moment, le nombre d'internes ou d'étudiants à encadrer. Ces facteurs pourraient être à même d'influencer le suivi de l'IA par le médecin, afin de réduire le risque d'erreur lié à la perte d'esprit critique due aux contraintes du métier.

De même, il serait intéressant, d'analyser la tendance à suivre l'IA en fonction de son diagnostic et pouvoir la mettre en corrélation avec les tests de performances déjà réalisés.

C. Changement d'avis de l'urgentiste

Les outils d'aide à la décision permettent de détecter avec précision les probables anomalies sur des clichés radiologiques, mais aussi d'attirer l'attention des médecins sur ces anomalies. En offrant un appui à la pose de leur diagnostic, les médecins peuvent ainsi se focaliser sur l'endroit suspect et ainsi prendre la décision en fonction de données telles que l'examen clinique, les antécédents du patient et leurs connaissances radiologiques.

Dans certaines situations, en se focalisant particulièrement sur une zone spécifique du cliché, grâce à l'expertise de l'IA et après analyse approfondie, le médecin urgentiste peut donc suivre l'avis de l'IA et reconsidérer son diagnostic. C'est ce que nous avons étudié.

Pour notre part, nous avons défini le changement d'avis de l'urgentiste, comme une modification du diagnostic du médecin pour rejoindre l'avis de l'IA après que l'IA et l'urgentiste ont proposé des avis différents. Il conviendrait de préciser, que tout au long

de la prise en charge, le médecin ignore la vérité, c'est-à-dire le *gold standard* et si l'avis de l'IA est juste ou erroné.

De plus, dans tous les cas, le médecin reste, dans tous les cas, le décisionnaire final de la prise en charge, et c'est sa responsabilité qui est engagée.

Le taux de changement d'avis de l'urgentiste parmi les cas où l'IA pose le bon diagnostic est de 8,8%. C'est-à-dire que le médecin change d'avis dans 8,8% des cas lorsque le diagnostic posé par l'IA est correct.

De plus, le taux de changement d'avis quand l'IA a raison parmi tous les changements d'avis est de 79,4% (27/34). Parmi toutes les fois où l'urgentiste change d'avis, il change d'avis dans 79,4% lorsque l'IA s'avère avoir raison : les changements d'avis permettent donc de rectifier le diagnostic posé initialement. Elle apporte une plus-value, et permet à l'urgentiste de corriger son diagnostic.

D. Taux de modification de prise en charge

Tout d'abord, les modifications de prises en charge sont considérées après l'avis de l'IA et les avis complémentaires demandés. De ce fait, nous ne pouvons pas conclure, si l'IA a fait changer une prise en charge, ou si celle-ci a été modifiée par les réponses des avis complémentaires.

Nous avons réalisé des analyses complémentaires, en prenant en compte la concordance ou la discordance du diagnostic posé par l'urgentiste et celui posé par le logiciel, pour analyser plus finement la modification de la prise en charge.

Au sein de nos 354 données, on observe 39 modifications de prise en charge. Pour 9 cas, la modification de la prise en charge n'a pas été précisée, pour 16 patients elle a entraîné la pose d'une attelle, 8 fois celle d'un plâtre et 1 fois une prise en charge chirurgicale.

La principale limite de ce critère secondaire, provient de l'absence de données sur l'intention de prise en charge du médecin.

Nous avons analysé les données en fonction des correspondances de diagnostic des urgentistes et de l'IA, puis la correspondance des diagnostics de l'IA et de l'expert.

Certaines données ont présenté des caractéristiques inattendues. Premièrement, nous retrouvons un unique cas où le diagnostic posé est différent entre l'IA et l'urgentiste, l'IA ayant tort, avec une modification de prise en charge : l'urgentiste aurait suivi l'IA a tort et aurait changé sa prise en charge vers une prise en charge erronée. L'IA a probablement induit en erreur l'urgentiste, qui lui a fait confiance.

Deuxièmement, pour 15 dossiers, alors que le médecin et l'IA avaient tous les deux le bon diagnostic, il y a eu une modification de la prise en charge. Nous pouvons l'expliquer de deux manières différentes : 1) le médecin avait un doute, l'IA a permis de le confirmer, et la prise en charge a évolué pour s'avérer différente entre le premier doute et le moment où l'IA l'a confirmé ; 2) les avis complémentaires demandés ont modifié la prise en charge. Par exemple, une prise en charge orthopédique par immobilisation plâtrée a été proposée par le spécialiste alors que l'urgentiste ne

pensait mettre qu'une attelle. Ou alors une prise en opération chirurgicale a été réalisée alors que l'urgentiste pensait faire une immobilisation orthopédique seule.

Ensuite, pour un des patients, ni l'IA ni l'urgentiste n'avaient le bon diagnostic, mais la prise en charge a quand même été modifiée, et ce probablement grâce aux avis complémentaires demandés.

En revanche, l'IA a joué un rôle important en ayant potentiellement été bénéfique lors de certaines prises en charge. Si nous ne pouvons pas affirmer avec certitude que l'IA a le rôle déterminant dans ce changement de prise en charge, elle a joué un rôle dans la modification de prise en charge l'orientant vers un traitement plus adéquat dans 6,2% des données (22/354). En effet, le diagnostic initial était différent entre l'urgentiste et l'IA, et le médecin se trompait à l'instar du logiciel. Dans ces 22 cas, l'IA a répondu aux attentes que nous avons d'un logiciel d'aide à la décision.

Nous devons cependant reconnaître que l'IA a pu être de mauvais conseil et a pu faire douter le médecin dans 32 cas. Alors que le médecin avait réalisé le bon diagnostic l'avis proposé par le logiciel était différent. Dans ce cas, l'urgentiste est resté sûr de lui et n'a pas modifié sa prise en charge, à raison.

En outre, dans 12 autres situations l'urgentiste et l'IA ont posé le même diagnostic, qui n'était cependant pas concordant avec celui de l'expert. L'IA a pu conforter à tort l'urgentiste malgré son erreur. Il n'y a donc pas eu de modification de prise en charge. L'avis complémentaire n'a probablement pas été demandé à ce moment-là, car les réponses entre l'urgentiste et l'IA étaient concordantes. Ces

patients-là ont logiquement dû être recontactés par la suite, lorsque la discordance a été mise en évidence, pour que le MAO a puisse rectifier la prise en charge finale. L'IA n'a donc pas apporté de bénéfice dans ces cas-là.

Au contraire, les médecins n'ont pas écouté l'IA dans 16 cas, ils sont restés sûrs de leur hypothèse initiale, alors qu'il aurait fallu suivre les propositions de l'IA.

Malheureusement ce ne sont que des suppositions, nous n'étions pas présents lors des réflexions de prise en charge, et les données que nous avons ne nous permettent pas de l'affirmer avec certitude.

Pour rappel, le compte rendu du radiologue n'est édité qu'à distance de la prise en charge, et de ce fait les médecins à l'instant T ne savent pas si l'IA dit vrai.

E. Taux d'erreur de l'urgentiste et variables associées

Lors de la prise en charge des patients, l'erreur humaine existe. Tout d'abord, l'urgentiste peut ne pas détecter des lésions ou parfois en détecter à tort. La radiographie réduit un modèle en trois dimensions en une image en deux dimensions, ce qui modifie la perception de la réalité des lésions. Par exemple la visualisation d'images construites. De surcroit, l'urgentiste n'a pas la formation d'un radiologue, et n'a donc pas non plus son expertise. D'une certaine façon, l'urgentiste et le logiciel de *Machine Learning*, se ressemblent, car les deux apprennent par expérience.

Les variables de notre étude qui sont associées de manière significative au taux d'erreur de l'urgentiste retrouvées sont multiples.

La fracture est la lésion la plus diagnostiquée mais derrière ce terme se cache une multitude de cas et de complexité. Certaines fractures peuvent donc passer inaperçues pour un œil moins expérimenté.

Il existe, dans la littérature, de nombreuses études, concernant les erreurs de diagnostic de l'urgentiste et les facteurs associées.

Nous retrouvons dans notre étude certains résultats similaires à ceux étudiés précédemment. En effet, nous retrouvons un risque d'erreur statistiquement lié au diagnostic de fracture, et à une localisation de la lésion au niveau la cheville. Dans cette étude de *Moonen et al*, (29), ils mettent en évidence la hausse des erreurs de diagnostic de l'urgentiste lorsque les lésions sont situées au niveau de la cheville et du pied, entre autres. En pratique, l'articulation de la cheville est composée de deux os de la jambe qui viennent pincer plusieurs os du pied. Indéniablement, devant la complexité de cette articulation, la mise en 2D des images par le cliché radiologique, peut limiter l'interprétation et faire apparaître des images construites ou masquer certaines lésions en présentant des incidences particulières. De même, des lésions situées au niveau du pied ou au niveau de la main sont aussi associées à risque d'erreur plus important. Du fait de leur localisation aux extrémités, les analyses doivent être plus fines. Au niveau du pied, le tarse est composé d'une multitude d'os différents qui s'intriquent et se superposent, plus favorable aux erreurs de diagnostic.

Contrairement à cette étude, nous n'avons pas retrouvé de résultat significatif quant aux lésions au niveau du poignet.

En outre, une lésion au niveau du fémur (nous considérons ici l'extrémité supérieure du fémur qui est la plus couramment retrouvée), est statistiquement liée au risque d'erreur de l'urgentiste. A savoir, c'est à cet endroit qu'on retrouve le plus de matériel d'ostéosynthèse : les prothèses de hanche peuvent donc empêcher l'analyse fine et correcte des radiographies. De plus, les remaniements ostéoporotiques, sont très présents au niveau de l'articulation coxo-fémorale et peuvent être pris à tort pour des lésions ou au contraire, empêcher la visualisation de lésions en pensant que ce sont des ostéophytes, par exemple.

De même, dans notre étude, comme nous avons pu retrouver dans plusieurs études antérieures, nous avons montré une différence en fonction du statut du médecin (30) (31).

En effet, ces études montraient qu'il y avait un risque accru d'erreur lorsqu'un jeune médecin prenait en charge le patient. Cependant, nos résultats indiquent que le statut d'interne n'est pas significativement associé à une augmentation des erreurs contrairement à celui de médecin urgentiste sénior.

Ces conclusions nécessitent des études complémentaires pour appuyer ces résultats.

Plusieurs hypothèses peuvent être envisagées : d'une part, les questionnaires ont été complétés par les médecins eux-mêmes ce qui pourrait entraîner des biais dans les réponses. De plus, d'une manière plus générale, les internes ayant moins d'expérience sont peut-être plus attentifs au cliché, prenant davantage de temps et ainsi réduisant le risque d'erreur, par rapport aux médecins plus expérimentés. Enfin,

les médecins séniors sont plus sujets aux interruptions de tâche, ce qui entraîne des distractions et donc des erreurs.

Cette inversion de situation pourrait suggérer que les internes en médecine sont mieux encadrés, que l'enseignement reçu par les médecins plus expérimentés et le temps qu'ils accordent à l'analyse permettent de limiter les erreurs, ou alors que leurs connaissances en radiologie sont plus solides qu'avant.

Cependant, d'autres études plus récentes ne mettaient pas en évidence de différence entre les statuts des médecins. (32)(33)

F. Les ressentis des médecins.

De plus en plus, l'IA fait partie intégrante du quotidien médical. Elle s'immisce dans de nombreux domaines, et elle est utilisée pour réaliser ce que les médecins savent faire, mais aussi pour ce qu'ils ne savent pas faire (comme la médecine prédictive).

Pour pouvoir intégrer la technologie dans le flux de travail aux urgences, il faut que les principaux acteurs l'adoptent. C'est un défi de faire accepter ces outils aux professionnels de santé. Existe-il une réticence due à la peur de l'erreur ou à la méfiance envers la technologie ?

Une enquête a été réalisée par Pulse Life (34) sur l'impact de l'IA dans la pratique médicale auprès des professionnels de santé et notamment sur la confiance qu'ils accordent à l'IA. Dans le cadre de l'aide au diagnostic, la confiance en l'IA reste mesurée, avec 63,5% des médecins déclarant avoir confiance. Les principales

préoccupations quant à l'émergence de l'IA dans l'aide au diagnostic sont les biais algorithmiques, la transparence des sources et la dégradation de la relation soignant-patient. Une autre étude est sortie en février 2024 concernant le regard des médecins français sur le recours à l'IA dans le diagnostic des maladies rares (35). 81% des médecins considèrent que l'IA pourrait les aider à améliorer le diagnostic des maladies rares. En effet, le diagnostic doit être posé rapidement et malheureusement il existe une errance médicale importante.

Nous avons demandé aux médecins qui ont rempli les questionnaires de partager leur ressenti sur la valeur ajoutée de l'IA dans chaque prise en charge.

Les jeunes médecins, fraîchement diplômés ou en cours d'apprentissage, sont-ils plus enclins à se laisser convaincre par le diagnostic de l'IA, ou bien, forts de leurs connaissances récentes, auront-ils davantage confiance en leur propre jugement ? De leur côté, les médecins plus expérimentés, s'appuieront-ils plutôt sur leurs expertises et leurs connaissances accumulées au fil des années et sont-ils plus réticents à faire confiance à l'IA, sachant que son taux de réussite n'atteint pas 100% ? En fin de compte, le ressenti moyen des médecins est de 5,08 sur 10, avec une médiane à 5.

La moyenne de ressenti la plus haute est celle des externes à 6,52. Les externes sont des étudiants en médecine, qui sont en formation et qui n'ont pas l'habitude d'analyser des radiologies. Ils sont donc plus à même de faire confiance au diagnostic posé par l'IA, et ainsi l'IA pourrait rattraper des erreurs plus régulièrement. Mais, les externes ne sont jamais seuls, ils sont toujours supervisés, ils ne peuvent pas rendre

un diagnostic en autonomie : leurs prises en charge ne sont pas définitives tant qu'un sénior ne les a pas validées. Le garde-fou reste le sénior.

Les internes ont une moyenne de ressenti de 5,9. Nous aurions pu nous attendre à une moyenne plus élevée, sachant que la plupart des internes ne sont pas habitués à ce type d'exercice.

De plus, nous nous rendons compte que plus le médecin a de l'expérience, plus le ressenti vis-à-vis de l'IA est faible.

Dans la prise en charge des traumatismes de membre, la démarche diagnostique est très protocolisée. L'IA pourrait être acteur d'homogénéisation des prises en charge avec son algorithme de diagnostic. Nous aurions pu attendre une moyenne de ressenti plus élevée de la part des médecins.

Pour aider au mieux les urgentistes et améliorer leur ressenti vis-à-vis de l'expertise de l'IA, il serait intéressant de créer un pourcentage de confiance de l'IA dans son diagnostic quand elle évoque un doute. Ainsi, cela permettrait d'accélérer le processus décisionnel et de diagnostic, ce qui est crucial dans un service d'urgences, particulièrement au CHV où il y a beaucoup de passage. La finalité serait de diminuer la variabilité inter examinateur, permettre de standardiser les diagnostics et d'améliorer la qualité des soins en diminuant le nombre de discordances.

III. FORCES ET FAIBLESSES

A. Les forces

L'analyse quasi systématique à distance des clichés par les radiologues, a permis un *gold standard* fiable.

Dans notre étude, les médecins ont intégré l'IA dans leur flux de travail aux urgences. De nombreux défis sont liés à l'adoption de l'IA par les professionnels de santé pour permettre une étroite collaboration homme/machine. L'IA est utilisée comme un outil d'aide à la décision plutôt que comme le remplacement de l'expertise humaine. Par exemple, dans la gestion des cas atypiques, les capacités de l'IA peuvent être remises en question et le radiologue sera encore présent et disponible pour donner son expertise.

B. Les faiblesses

La principale faiblesse de notre étude est le nombre de patients inclus. En effet, le recueil de données dépendait du nombre de questionnaires remplis par les médecins urgentistes au moment de la prise en charge des patients. Des dizaines de radiologies sont réalisées chaque jour, mais peu de questionnaires ont été remplis.

Nous avons de plus, un biais de sélection, car les professionnels de santé qui ont le plus répondu sont des internes. Nous n'avons pas eu d'échantillon significatif de chaque catégorie de médecin.

La troisième limite porte sur le caractère monocentrique de notre étude, limitant alors l'extrapolation des données. De nombreux CH ont accès à ces logiciels d'intelligence artificielle dans l'aide au diagnostic.

CONCLUSION

L'IA ne remplace pas le médecin et l'urgentiste demeure le principal acteur de la prise en charge du patient. Dans les structures d'urgence, l'utilisation de l'IA nécessite une garantie humaine et se doit d'être strictement encadrée. Les enjeux éthiques et légaux face à l'intégration de l'IA en médecine sont nombreux et doivent impérativement être considérés dans le développement des logiciels d'IA (36)(37).

Néanmoins, l'IA peut véritablement enrichir le processus de diagnostic, notamment dans l'interprétation des radiographies aux urgences. Devant ses performances et ses capacités diagnostiques, son indépendance vis-à-vis des facteurs extérieurs, l'IA permet aux urgentistes de corriger et d'ajuster leurs diagnostics ainsi que les décisions de prise en charge.

Nos analyses indiquent que l'IA constitue davantage une assistance supplémentaire pour les médecins qu'une réponse définitive. En fournissant une seconde opinion, l'IA peut renforcer la confiance du médecin en son diagnostic ou lui permettre de se concentrer sur des zones spécifiques d'une radiographie, influençant alors sa décision thérapeutique. A ce jour, l'alliance IA et urgentiste, reste la meilleure association pour diminuer les erreurs de diagnostic.

Cependant, des études supplémentaires impliquant un plus grand nombre de patients sont nécessaires pour valider ces premières observations.

LISTE DES FIGURES ET DES TABLEAUX

FIGURE 1 : FLOW CHART INCLUSION	17
FIGURE 2 : CATEGORIE MEDICALE PRENANT EN CHARGE LE PATIENT.....	23
FIGURE 3 : SIEGE DE LA LESION.....	24
FIGURE 4 : PRESENCE DE MATERIEL D'OSTEOSYNTHESE.....	25
FIGURE 5 : LESION SELON L'URGENTISTE PRENANT EN CHARGE LE PATIENT.	26
FIGURE 6 : LESION SELON L'IA	26
FIGURE 7 : FOREST PLOT DES VARIABLES ASSOCIEES AU SUIVI DE L'IA.....	33
FIGURE 8 : FOREST PLOT DES VARIABLES ASSOCIEES A L'ERREUR DE L'URGENTISTE	39
FIGURE 9 : ÉCHELLE DE RESENTI CONCERNANT L'IA	40
TABLEAU 1 : REPARTITION DE LA PRISE EN CHARGE DES DISCORDANCES 2021/2022.	14
TABLEAU 2 : SUIVI DE L'IA.....	31
TABLEAU 3 : VARIABLES ASSOCIEES AU SUIVI DE L'IA.....	33
TABLEAU 4 : CHANGEMENT D'AVIS DE L'URGENTISTE	34
TABLEAU 5 : MODIFICATION DE LA PRISE EN CHARGE EN FONCTION DU DIAGNOSTIC DE L'IA	35
TABLEAU 6 : MODIFICATION DE PRISE EN CHARGE EN FONCTION DU RESULTAT INITIAL DE L'URGENTISTE ET DE L'IA.....	36
TABLEAU 7 : VARIABLES ASSOCIEES A L'ERREUR DES URGENTISTES	38

BIBLIOGRAPHIE

1. Projet établissement CHV [Internet]. Disponible sur: https://www.ch-valenciennes.fr/app/uploads/2024/05/Interactions_Projet_etablissement_2024_2029.pdf
2. Centre Hospitalier de Valenciennes - CHV [Internet]. 2021 [cité 3 août 2024]. Disponible sur: <https://www.ch-valenciennes.fr/>
4. Parpaleix A, Parsy C, Cordari M, Mejdoubi M. Assessment of a combined musculoskeletal and chest deep learning-based detection solution in an emergency setting. *Eur J Radiol Open*. 2023;10:100482.
5. FLYER_ETUDE_European Journal of Radiology Open_Deep Learning et radiographie des Urgences _ Mars 2023.pdf.
6. Hussain F, Cooper A, Carson-Stevens A, Donaldson L, Hibbert P, Hughes T, et al. Diagnostic error in the emergency department: learning from national patient safety incident report analysis. *BMC Emerg Med*. 4 déc 2019;19:77.
7. European Parliament. Directorate General for Parliamentary Research Services. Artificial intelligence: how does it work, why does it matter, and what we can do about it? [Internet]. LU: Publications Office; 2020 [cité 25 janv 2024]. Disponible sur: <https://data.europa.eu/doi/10.2861/44572>
8. Intelligence artificielle [Internet]. [cité 29 juill 2024]. Disponible sur: <https://www.cnil.fr/fr/definition/intelligence-artificielle>
9. Larousse É. intelligence artificielle - LAROUSSE [Internet]. [cité 25 janv 2024]. Disponible sur: https://www.larousse.fr/encyclopedie/divers/intelligence_artificielle/187257
10. Inserm [Internet]. [cité 29 janv 2024]. Intelligence artificielle et santé · Inserm, La science pour la santé. Disponible sur: <https://www.inserm.fr/dossier/intelligence-artificielle-et-sante/>
11. Intelligence artificielle : définitions et défis.
12. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. janv 2019;25(1):44-56.
13. Gaube S, Suresh H, Raue M, Merritt A, Berkowitz SJ, Lerner E, et al. Do as AI say: susceptibility in deployment of clinical decision-aids. *NPJ Digit Med*. 19 févr 2021;4(1):31.
14. Liew C. The future of radiology augmented with Artificial Intelligence: A strategy for success. *Eur J Radiol*. mai 2018;102:152-6.
15. Oakden-Rayner L, Gale W, Bonham TA, Lungren MP, Carneiro G, Bradley AP, et al. Validation and algorithmic audit of a deep learning system for the detection of proximal femoral fractures in patients in the emergency department: a diagnostic accuracy study. *Lancet Digit Health*. mai 2022;4(5):e351-8.
16. Kuo RYL, Harrison C, Curran TA, Jones B, Freethy A, Cussons D, et al. Artificial Intelligence in Fracture Detection: A Systematic Review and Meta-Analysis. *Radiology*. juill 2022;304(1):50-62.
17. Kalmet PHS, Sanduleanu S, Primakov S, Wu G, Jochems A, Refaee T, et al. Deep

- learning in fracture detection: a narrative review. *Acta Orthop.* avr 2020;91(2):215-20.
18. Guermazi A, Tannoury C, Kompel AJ, Murakami AM, Ducarouge A, Gillibert A, et al. Improving Radiographic Fracture Recognition Performance and Efficiency Using Artificial Intelligence. *Radiology.* mars 2022;302(3):627-36.
 19. Boonstra A, Laven M. Influence of artificial intelligence on the work design of emergency department clinicians a systematic literature review. *BMC Health Serv Res.* déc 2022;22(1):669.
 20. Shaw G. AI in the ED? Maybe. *Emerg Med News.* août 2023;45(8):12-3.
 21. Jones RM, Sharma A, Hotchkiss R, Sperling JW, Hamburger J, Ledig C, et al. Assessment of a deep-learning system for fracture detection in musculoskeletal radiographs. *NPJ Digit Med.* 2020;3:144.
 22. Hallas P, Ellingsen T. Errors in fracture diagnoses in the emergency department--characteristics of patients and diurnal variation. *BMC Emerg Med.* 16 févr 2006;6:4.
 23. Arnaud E, Petitprez E, Ammirati C, Nemitz B, Dequen G, Gignon M, et al. L'intelligence artificielle dans les structures d'urgences : place de la formation et de la garantie humaine. *Ann Fr Médecine D'urgence.* mai 2023;13(3):169-78.
 24. Pesapane F, Tantrige P, Patella F, Biondetti P, Nicosia L, Ianniello A, et al. Myths and facts about artificial intelligence: why machine- and deep-learning will not replace interventional radiologists. *Med Oncol Northwood Lond Engl.* 3 avr 2020;37(5):40.
 25. previa. Previa Medical. [cité 3 août 2024]. Previa Medical. Disponible sur: <https://www.previa-medical.com>
 26. Solution CALYPS Saniia • CALYPS [Internet]. CALYPS. [cité 3 août 2024]. Disponible sur: <https://www.calyps.ch/en/medical-solutions/calyps-saniia-solution/>
 27. Bousson V, Attané G, Benoist N, Perronne L, Diallo A, Hadid-Beurrier L, et al. Artificial Intelligence for Detecting Acute Fractures in Patients Admitted to an Emergency Department: Real-Life Performance of Three Commercial Algorithms. *Acad Radiol.* 1 oct 2023;30(10):2118-39.
 28. Pelaccia T, Forestier G, Wemmert C. Une intelligence artificielle raisonne-t-elle de la même façon que les cliniciens pour poser des diagnostics ? *Rev Médecine Interne.* mars 2020;41(3):192-5.
 29. Moonen PJ, Mercelina L, Boer W, Fret T. Diagnostic error in the Emergency Department: follow up of patients with minor trauma in the outpatient clinic. *Scand J Trauma Resusc Emerg Med.* déc 2017;25(1):13.
 30. Guly HR. Diagnostic errors in an accident and emergency department. *Emerg Med J.* 1 juill 2001;18(4):263-9.
 31. André S, Lim MS, Iraqui I, Gestin S, Francart A, Takun K, et al. Facteurs favorisant les erreurs d'interprétation radiologique aux urgences. *J Eur Urgences.* mai 2007;20(1):15.
 32. Bouillon-Minois JB, Lambert C, Dutheil F, Raconnat J, Benamor M, Dalle B, et al. Assessment of discordance between radiologists and emergency physicians of RADIOgraphs among discharged patients in an emergency department: the RADIO-ED study. *Emerg Radiol.* avr 2024;31(2):125-31.

33. McLauchlan CA, Jones K, Guly HR. Interpretation of trauma radiographs by junior doctors in accident and emergency departments: a cause for concern? *J Accid Emerg Med.* sept 1997;14(5):295-8.
34. PulseLife [Internet]. [cité 9 août 2024]. Baromètre IA en santé : alliée ou menace ? Disponible sur: <https://pulselife.com/fr-fr/blog/post/barometre-ia-en-sante-alliee-ou-menace>
35. Rapport-Ifop_sanofi_malad.rares_2024.02.21.pdf [Internet]. [cité 9 août 2024]. Disponible sur: https://www.ifop.com/wp-content/uploads/2024/02/Rapport-Ifop_sanofi_malad.rares_2024.02.21.pdf?fbclid=IwY2xjawEi8vtleHRuA2FlbQIxMAABHUISGJlk5hNtDe5aYX8Jq4QYyfP9YqLVgixX4OR--BBE7ErajI4xEdkM1g_aem_iSnT3mVjXoJzPdeenyEshg
36. Manaouil C, Chamot S, Petit P. Le médecin confronté à l'IA (Intelligence artificielle) : Éthique et responsabilité. *Médecine Droit.* juin 2024;2024(186):50-66.
37. Langlois L, Dilhac MA, Dratwa J, Ménissier T, Ganascia JG, Weinstock D, et al. L'éthique au cœur de l'IA [Internet]. Observatoire international sur les impacts sociétaux de l'intelligence artificielle et du numérique; 2023 oct [cité 16 août 2024]. Disponible sur: <https://www.obvia.ca/ressources/lethique-au-coeur-de-lia>

ANNEXES

Annexe 1 : Questionnaire

ETIQUETTE PATIENT

QUESTIONNAIRE DE THESE SUR L'APPORT DE L'INTELLIGENCE ARTIFICIELLE *A DESTINATION DES MEDECINS AUX URGENCES ADULTES DU CH DE VALENCIENNES*

Ce questionnaire vise à évaluer l'efficacité du logiciel d'intelligence artificielle des urgences adultes du CH de Valenciennes pour le diagnostic des lésions ostéoarticulaires du squelette appendiculaire.

1) VOUS ETES :

- **Urgentiste** DJ < 5ans 5-10 ans 10-20 ans > 20ans
- **Interne** en stage aux urgences de garde
 phase socle phase d'approfondissement phase de consolidation/DJ
- **FFI** < 1an > 1an
- **Externe** 4ème année 5ème année 6ème année

2) SIEGE DE LA LESION



Merci d'entourer la zone correspondante à la lésion sur le schéma.

Présence de matériel d'ostéosynthèse : OUI NON

3) **ARBRE DIAGNOSTIC**

- Voyez-vous une lésion ?

OUI, de manière certaine OUI, mais j'ai un doute NON

- L'Intelligence artificielle (IA) montre-t-elle une lésion ?

OUI Doute NON

- La lésion ou le doute détecté par l'IA correspond-il à votre diagnostic initial ?

OUI NON

Si non, pourquoi ?.....

- Si une lésion est mise en évidence, quelle est sa nature ?

Fracture Luxation Epanchement

4) **POUR CONCLURE**

- Suivez-vous l'avis de l'IA pour prendre votre décision ? OUI NON

- Sollicitez-vous un avis complémentaire ? OUI NON

Si oui, le(s)quel(s) : Radiologue Sénior des urgences Chirurgien

- Ces avis sont-ils en accord avec l'IA ? OUI NON

- Demandez-vous un examen complémentaire ? OUI NON

Si oui, lequel ? TDM IRM ECHOGRAPHIE

Et pourquoi ?

Discordance humain/IA

Aide à l'orientation thérapeutique

Autre

=> Finalement l'intervention de l'IA a t'elle modifié votre prise en charge ?

OUI NON

Si oui : Attelle Plâtre Chirurgie

=> Quel est votre ressenti sur l'intérêt de l'intelligence artificielle dans le diagnostic de la lésion ostéoarticulaire dans ce cas précis ? (0 : intérêt nul - 10 : très fort intérêt)

0 1 2 3 4 5 6 7 8 9 10

Annexe 2 : Attestation d'avis favorable du CERCL

Commission Ethique de la Recherche Clinique (CERCL)



Présidence

Dr Laurie Ferret *

Membres

Dr Lidvine Godaert*

Dr Clotilde Fontier

Dr Fabien Lambiotte

Mme Stéphanie Pigla*

Dr Morgane Plancon

Me Catherine Fache

Dr Audrey Parent

* membres permanents

Avis de la commission Ethique de la Recherche Clinique du Centre Hospitalier de Valenciennes

La Commission Ethique de la Recherche Clinique (CERCL) du CH Valenciennes s'est réunie pour examiner le projet intitulé :

Intérêt des logiciels d'intelligence artificielle et spécifiquement dans le service des urgences adultes du CHV dans le cadre du diagnostic des lésions ostéo articulaires.

Présenté par le Dr Erwan Dupire et Agathe Vinot (interne)

Et a donné un **avis favorable (Ref CHV-2024-019)** après examen de sa conformité aux règles en vigueur régissant la recherche clinique.

Avis rendu le 20/08/2024

Pour les membres de la commission, la présidence,

Dr Laurie Ferret

PharmD- PhD

Annexe 3 : Formulaire internet d'information aux patients.

Intelligence artificielle (IA) et le diagnostic des lésions lors d'une radiographie réalisée aux urgences dans le cadre d'un traumatisme de membre

Le Centre hospitalier de Valenciennes souhaite réaliser une recherche n'impliquant pas la personne humaine ayant pour objet *l'intelligence artificielle (IA) et le diagnostic des lésions lors d'une radiographie réalisée aux urgences dans le cadre d'un traumatisme de membre.*

Cette recherche sera développée au Centre hospitalier de Valenciennes du *1^{er} décembre 2023* au *1^{er} mai 2024*.

Quel est l'objectif de la recherche ?

Cette recherche a pour objectif d'évaluer le taux de différence de diagnostic entre l'IA et l'urgentiste, lorsque l'IA a raison, c'est-à-dire lorsque l'IA est en accord avec l'avis du radiologue.

Cette étude est menée pour des motifs d'intérêt public dans le domaine de la santé, par son objectif d'améliorer la qualité de prise en charge des patients. A ce titre, l'étude est conforme aux dispositions particulières relatives aux traitements à des fins de recherche, d'étude ou d'évaluation dans le domaine de la santé autorisées par la Commission Nationale Informatique et libertés en application de la loi Informatique et Libertés.

Comment la recherche va-t-elle se dérouler ?

Cette recherche ne demande aucune participation de votre part ni aucune consultation ou examen supplémentaire. En effet, cette recherche repose exclusivement sur la réutilisation et le traitement de vos données de santé contenues dans votre dossier médical. Cette recherche ne présente donc aucun risque, il s'agit uniquement de l'analyse de vos données de santé.

Les professionnels de santé du Centre hospitalier de Valenciennes, soumis au secret médical, recueilleront, dans votre dossier médical, les seules informations relatives à votre séjour hospitalier strictement pertinentes et adéquates à l'objectif scientifique poursuivi.

Ensuite, vos données de santé seront pseudonymisées. Cela signifie que vos données seront identifiées par un numéro de code confidentiel afin que vous ne puissiez pas être directement identifiable.

Finalement, vos données pseudonymisées seront extraites et analysées conformément au protocole scientifique encadrant la recherche.

Pourquoi vous propose-t-on de participer à cette recherche ?

Vous avez été hospitalisé au Centre hospitalier de Valenciennes pour un traumatisme de membre et une radiologie a été réalisée du *1^{er} décembre 2023* au *1^{er} mai 2024*. Vos données de santé présentent donc un intérêt pour la finalité de la recherche.

Quelle est la nature des données traitées ?

Les catégories de données pseudonymisées nécessaires à la réalisation de cette recherche concernent la démarche diagnostique réalisée par le médecin lors de votre prise en charge, et le compte rendu radiologique de votre radiographie réalisée.

Les données utilisées sont recueillies dans votre dossier médical. Elles ont été générées dans le cadre du soin, indépendamment de la recherche.

Quels sont les destinataires des données ?

L'accès à vos données de santé identifiantes est restreint aux professionnels de santé du Centre hospitalier de Valenciennes ayant participé à votre prise en charge médicale et à leurs collaborateurs de santé intervenant dans la recherche.

Un accord de confidentialité étant établie entre le Centre hospitalier de Valenciennes et l'investigateur, vos données resteront strictement confidentielles.

Conservation des données et publication des résultats de la recherche

Vos données pseudonymisées seront conservées pour toute la durée de la recherche, de manière sécurisée.

Ensuite, les données de la recherche seront conservées jusqu'à deux ans après la dernière publication des résultats de la recherche ou, en cas d'absence de publication, jusqu'à la signature du rapport final de la recherche. Elles feront ensuite l'objet d'un archivage sur papier ou support informatique pour une durée maximale de vingt ans.

Les résultats de cette recherche pourront faire l'objet de communications lors de congrès scientifiques et/ou être publiés dans une revue scientifique. Dans tous les cas, les publications des résultats des recherches et des études seront réalisées à partir de données anonymes ne permettant en aucun cas de vous identifier.

Vous pourrez être informés des résultats globaux de cette recherche par l'intermédiaire du service des urgences adulte du Centre Hospitalier de Valenciennes dans lequel vous avez été pris en charge.

Quels sont vos droits en tant que participant à la recherche ?

La réutilisation de vos données de santé s'effectue dans le respect du cadre juridique relatif à la protection des données (Règlement général sur la protection des données-RGPD et loi Informatique et Libertés modifiée) et du respect de vos droits détaillés dans notre politique de protection des données personnelles : <https://www.ch-valenciennes.fr/protection-des-donnees-personnelles>. A ce titre, vous disposez d'un droit d'accès, de rectification, d'effacement et de limitation de vos données de santé personnelles.

Vous disposez également d'un droit d'opposition à participer à cette recherche. Vous pouvez donc, à tout moment et sans vous justifier, vous opposer à la réutilisation de vos informations médicales dans le cadre de cette recherche. Une opposition n'entraînera aucune conséquence sur la qualité des soins qui vous seront prodigués dans la structure que vous consultez. En l'absence d'opposition, nous commencerons l'analyse des données.

Vous pouvez exercer l'ensemble de vos droits, dont le droit d'opposition, en écrivant au délégué à la protection des données du Centre hospitalier de Valenciennes :

- ✓ Par mail à l'adresse : dpo@ch-valenciennes.fr ;
- ✓ Par courrier au Centre hospitalier de Valenciennes - Délégué à la protection des données, 114 avenue Désandrouin, 59322 Valenciennes.

Pour toute question, vous pouvez également contacter le responsable scientifique de la recherche, le Dr *DUPIRE*, au 03.27.14.33.33

Identité et coordonnées du Responsable de traitement et du Délégué à la protection des données :

- ✓ Responsable du traitement : Centre Hospitalier de Valenciennes, Avenue Desandrouin, S 50479, 59322 Valenciennes Cedex.

Nous vous remercions pour votre contribution à cette recherche.

<p>AUTEUR : Nom : VINOT Prénom : Agathe</p> <p>Date de soutenance : 4 octobre 2024</p> <p>Titre de la thèse : Apport de l'Intelligence Artificielle dans la prise en charge des lésions ostéoarticulaires du squelette appendiculaire aux urgences.</p> <p>Thèse - Médecine - Lille 2024</p> <p>Cadre de classement : Médecine d'urgence et Radiologie</p> <p>DES : Médecine d'urgence</p> <p>Mots-clés : : intelligence artificielle, radiologie, urgence, lésions ostéoarticulaires</p> <p>Contexte : L'Intelligence Artificielle prend une place de plus en plus importante dans le domaine médical et spécialement dans l'aide diagnostique. Les traumatismes de membre sont un motif courant de consultation aux urgences et de nombreuses radiographies sont réalisées. Depuis 2022, les urgences adultes du Centre Hospitalier de Valenciennes sont dotées d'un logiciel d'intelligence artificielle (logiciel SmartUrgences® par la société MILVUE®) pour l'aide au diagnostic des lésions pulmonaires et ostéoarticulaires. Ce logiciel a déjà fait ses preuves et ses tests de performance ont été analysés. L'impact de ce logiciel en pratique clinique n'a pas été évalué à ce jour. L'intérêt de cette étude est d'évaluer ce logiciel comme aide diagnostique pour les urgentistes en traumatologie au moment de la prise en charge des patients et d'étudier les variables qui exercent une influence sur cette aide diagnostique.</p> <p>Méthode : Il s'agit d'une étude monocentrique, observationnelle et descriptive, réalisée aux urgences adultes du Centre Hospitalier de Valenciennes. Elle a été réalisée par recueil de données rétrospectif, du 1er décembre 2023 au 1er mai 2024. Le <i>Gold Standard</i> est le compte rendu du radiologue édité à postériori et récupéré à distance de la prise en charge.</p> <p>Résultats : 8,5% des diagnostics posés par l'urgentiste étaient erronés alors que l'IA était en accord avec le radiologue. L'IA est suivie lorsqu'elle a raison dans 96,1% des cas. Les facteurs associés au suivi de l'IA sont : l'absence de diagnostic de lésion par l'urgentiste, tout type de lésion et lorsque l'IA pose le bon diagnostic. Le taux de changement d'avis de l'urgentiste quand l'IA a raison est de 8,8%. La prise en charge a été modifiée vers celle adéquate dans 6,2%. Les facteurs associés à l'erreur de l'urgentiste sont : les lésions au niveau du pied, de la main, du genou et de la cheville, le doute initial, la demande d'avis complémentaire, et le fait d'être un urgentiste thésé. Le ressenti moyen des médecins vis-à-vis de l'IA est de 5,08 (+/- 2,98), pour des notes allant de 0 à 10.</p> <p>Conclusion : L'IA constitue une aide diagnostique non négligeable pour le médecin urgentiste, et permet des modifications de prise en charge ainsi que des rectifications de diagnostic. Cependant, l'IA ne reste pas la seule décisionnaire face au diagnostic final. Elle a un impact dans la prise en charge des patients, en apportant au médecin une assistance supplémentaire. L'alliance IA et urgentiste reste la meilleure association pour diminuer les erreurs de diagnostic.</p> <p>Composition du Jury : Président : Pr Eric WIEL Assesseurs : Dr Jean Marie RENARD Dr Jérôme MIZON Directeur de thèse : Dr Erwan DUPIRE</p>
