



**HAL**  
open science

# L'intelligence artificielle au sein des urgences traumatologiques

Marion Calandre

► **To cite this version:**

Marion Calandre. L'intelligence artificielle au sein des urgences traumatologiques. Médecine humaine et pathologie. 2022. dumas-04176781

**HAL Id: dumas-04176781**

**<https://dumas.ccsd.cnrs.fr/dumas-04176781v1>**

Submitted on 27 Sep 2023

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

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

## AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance.

La propriété intellectuelle du document reste entièrement celle du ou des auteurs. Les utilisateurs doivent respecter le droit d'auteur selon la législation en vigueur, et sont soumis aux règles habituelles du bon usage, comme pour les publications sur papier : respect des travaux originaux, citation, interdiction du pillage intellectuel, etc.

Il est mis à disposition de toute personne intéressée par l'intermédiaire de [l'archive ouverte DUMAS](#) (Dépôt Universitaire de Mémoires Après Soutenance).

Si vous désirez contacter son ou ses auteurs, nous vous invitons à consulter en ligne les annuaires de l'ordre des médecins, des pharmaciens et des sages-femmes.

Contact à la Bibliothèque universitaire de Médecine Pharmacie de Grenoble :

[bump-theses@univ-grenoble-alpes.fr](mailto:bump-theses@univ-grenoble-alpes.fr)

Année : 2022

**L'INTELLIGENCE ARTIFICIELLE AU SEIN DES URGENCES  
TRAUMATOLOGIQUES**

THÈSE POUR LE DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE

SPÉCIALITÉ : MÉDECINE D'URGENCE

Par Mme Marion CALANDRE

*[Données à caractère personnel]*

THÈSE SOUTENUE PUBLIQUEMENT À LA FACULTÉ DE MÉDECINE DE GRENOBLE

Le 27/09/2022

DEVANT LE JURY SUIVANT :

Président du jury :

M. le Pr Gilbert FERRETTI

Membres :

M. le Pr Régis PAILHÉ

M. le Pr Damien VIGLINO

M. le Dr Pierre CHAPUIS

M. le Dr Jean-Jacques BANIHACHÉMI (directeur de thèse)

CORPS	NOM-PRENUM	Discipline universitaire
PU-PH	ALBALADEJO Pierre	Anesthésiologie-réanimation et médecine péri-opératoire
PU-PH	ARVILLIX-BARILLIUMY Catherine	Chirurgie viscérale et digestive
PU-PH	BAILLET Athan	Rhumatologie
PU-PH	BARONE ROCHETTE Gilles	Cardiologie
PU-PH	BAYAT Sam	Physiologie
MCF Ass.MG	BENDAMFEN Fouad	Médecine Générale
PU-PH	BENHAMOU Pierre-Yves	Endocrinologie, diabète et maladies métaboliques
PU-PH	BERGER François	Biologie cellulaire
MCU-PH	BETRY Cecile	Nutrition
MCU-PH	BIDARI-COULON Marie	Biologie cellulaire
PU-PH	BIOLAC-ROGIER Stéphanie	Pédopsychiatrie ; addictologie
PU-PH	BLAIN Sophie	Chirurgie vasculaire ; médecine vasculaire
PR Ass. Méd.	BOILLOT Bernard	Urologie
MCU-PH	BOISSI Sandrine	Bactériologie-virologie ; Hygiène hospitalière
PU-PH	BONAZ Bruno	Gastroentérologie ; hépatologie ; addictologie
PU-PH	BONNETIER Vincent	Médecine et santé au travail
PU-PH	BOREL Anne Laure	Nutrition
PU-PH	BOSSON Jean-Luc	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	BOTTARI Serge	Biologie cellulaire
PR Ass.MG	BOURHAUD Jacques	Médecine Générale
PU-PH	BOUGEROL Thierry	Psychiatrie d'adultes
PU-PH	BOUJILLI Laurence	Médecine interne ; gériatrie et biologie du vieillissement ; addictologie
MCU-PH	BOUSSAT Bastien	Epidémiologie, économie de la santé et prévention
PU-PH	BOULAI Pierre	Anesthésiologie-réanimation et médecine péri-opératoire
PU-PH émérite	BRAMBILLA Christian	Pneumologie
PU-PH émérite	BRAMBILLA Elisabeth	Anatomie et cytologie pathologiques
MCU-PH	BRENIER-PINCHART Marie Pierre	Parasitologie et mycologie
PU-PH	BRICAULT Ivan	Radiologie et imagerie médicale
PU-PH	BRICHON Pierre Yves	Chirurgie thoracique et cardiovasculaire
MCU-PH	BRIOT Raphaël	Thérapeutique-médecine de la douleur ; Addictologie
PU-PH émérite	CAHN Jean Yves	Hématologie
PU-PH émérite	CARPENTIER Patrick	Chirurgie vasculaire, médecine vasculaire
PR Ass.MG	CARRILLO Yannick	Médecine Générale
MCU-PH	CASPAR Yvan	Bactériologie-virologie ; hygiène hospitalière
PU-PH	CESBRON Jean-Yves	Immunologie
PU-PH	CHABARDIS Stephan	Neurochirurgie
PU-PH	CHABRE Olivier	Endocrinologie, diabète et maladies métaboliques
PU-PH	CHAILLANJON Philippe	Anatomie
MCF Ass.MG	CHAMBOREDON Benoît	Médecine Générale
PU-PH	CHABIS Julie	Dermato-vénérologie
MCF Ass.MG	CHAUVET Marion	Médecine Générale
PU-PH	CHAVANON Olivier	Chirurgie thoracique et cardiovasculaire

CORPS	NOM-PRENUM	Discipline universitaire
PU-PH	CHIQUET Christophe	Ophthalmologie
PU-PH	CHIRICA Mircea	Chirurgie viscérale et digestive
PU-PH	CINQUIN Philippe	Biostatistiques, Informatique médicale et technologies de communication
MCU-PH	CLAVARINO Giovanna	Immunologie
MCU-PH	CLINCHIFFRE Rita	Nutrition
PU-PH	COHEN Olivier	Histologie, embryologie et cytogénétique
PU-PH	COSSU NINI Charlotte	Gastroentérologie ; hépatologie ; addictologie
PU-PH	COURVOISIER Aurélien	Chirurgie infantile
PU-PH	COUILLON Charles	Génétique
PU-PH	COUTURIER Pascal	Médecine interne ; gériatrie et biologie du vieillissement ; addictologie
PU-PH	CRACOWSKI Jean-Luc	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
PU-PH	DEBATY Guillaume	Médecine d'Urgence
PU-PH	DEBILLON Thierry	Pédiatrie
PU-PH	DECAENS Thomas	Gastroentérologie ; hépatologie ; addictologie
PR Ass. Méd.	DEFFAYE Pascal	Cardiologie
PU-PH	DEGANO Bruno	Pneumologie ; addictologie
PU-PH	DEMAITIS Maurice	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
PU-PH émérite	DEMONGEOY Jacques	Biostatistiques, Informatique médicale et technologies de communication
MCU-PH	DE RANSART Colin	Physiologie
PU-PH	DESCOTES Jean-Luc	Urologie
PU-PH	DIJANI Olivier	Neurologie
MCU-PH	DIETERICH Klaus	Génétique
MCU-PH	DOUILLAUD Stéphane	Physiologie
PU-PH	EPAULARD Olivier	Maladies infectieuses ; Maladies tropicales
PU-PH	ESTIÈRE François	Biophysique et médecine nucléaire
MCU-PH	EYSSERIC Hélène	Médecine légale et droit de la santé
PU-PH émérite	FAGIOLI Daniel	Biophysique et médecine nucléaire
PU-PH	FAUCHERON Jean-Luc	Chirurgie viscérale et digestive
MCU-PH	FAURE Julien	Biochimie et biologie moléculaire
PU-PH	FERRETTI Gilbert	Radiologie et imagerie médicale
PU-PH	FONTAINI Éric	Nutrition
PU-PH	FRANCOIS Patrice	Epidémiologie, économie de la santé et prévention
MCU-MG	GAINIBALI Yoann	Médecine Générale
PU-PH	GARBAN Frédéric	Hématologie ; Transfusion
PU-PH	GAUDIN Philippe	Rhumatologie
PU-PH	GAVAZZI Guélan	Médecine interne ; gériatrie et biologie du vieillissement ; addictologie
PU-PH	GAY Emmanuel	Neurochirurgie
MCU-PH	GIAI Joris	Biostatistiques, Informatique médicale et technologies de communication
MCU-PH	GILLOIS Pierre	Biostatistiques, Informatique médicale et technologies de communication
PU-PH	GIOT Jean-Philippe	Chirurgie plastique, reconstructrice et esthétique ; Brûlologie
MCU-PH	GRANDI Sylvie	Radiologie et imagerie médicale
PU-PH émérite	GRIFFET Jacques	Chirurgie infantile
PU-PH	HAINAUT Pierre	Biochimie et biologie moléculaire
PU-PH émérite	HALIMI Serge	Nutrition
PU-PH	HENNEBICQ Sylviane	Biologie et médecine du développement et de la reproduction ; gynécologie médicale
PU-PH	HOFFMANN Pascale	Gynécologie-obstétrique ; gynécologie médicale

CORPS	NOM-PRENOM	Discipline universitaire
PU-PH émérite	HOMMEL Marc	Neurologie
PU-MG	IMHUIET Patrick	Médecine Générale
PU-PH émérite	JOUK Pierre-Simon	Génétique
PU-PH	KAHANI Philippe	Physiologie
MCU-PH	KASTLER Adrian	Radiologie et imagerie médicale
PU-PH	KRAINIK Alexandre	Radiologie et imagerie médicale
PU-PH	LABARERE José	Epidémiologie, économie de la santé et prévention
MCU-PH	LABLANCHE Sandrine	Endocrinologie, diabète et maladies métaboliques
PU-PH	LANDELLE Caroline	Bactériologie – virologie ; Hygiène hospitalière
PU-PH	LANTUEJOU Sylvie	Anatomie et cytologie pathologiques
PR Ass. Méc.	LARAMAS Mathieu	Cancérologie ; radiothérapie
MCU-PH	LAIRDY Bernard	Biochimie et biologie moléculaire
MCU-PH	LE GOUELLEC LE PISSART Audrey	Biochimie et biologie moléculaire
PU-PH	LECCIA Marie-Thérèse	Dermato-vénérologie
PR Ass.MG	LEDOUX Jean-Nicolas	Médecine Générale
PU-PH émérite	LETOUBLON Christian	Chirurgie viscérale et digestive
PU-PH	LEVY Patrick	Physiologie
PU-PH	LONG Jean-Alexandre	Urologie
MCU-PH	LUPO Julien	Bactériologie virologie ; Hygiène hospitalière
PU-PH	MAITRE Anne	Médecine et santé au travail
MCU-PH	MARLU Raphaël	Hématologie ; Transfusion
PR Ass. Méc.	MATHIEU Nicolas	Gastroentérologie ; hépatologie ; addictologie
MCU-PH	MAUBON Danièle	Parasitologie et mycologie
PU-PH	MAURIN Max	Bactériologie-virologie ; Hygiène hospitalière
MCU-PH	MCIFFER Anne	Histologie, embryologie et cytogénétique
PR Ass. Méc.	MICHY Thierry	Gynécologie-obstétrique
MCU-PH	MONDET Julie	Histologie, embryologie et cytogénétique
PU-PH	MORAND Patrice	Bactériologie-virologie ; Hygiène hospitalière
PU-PH	MOREAU GAUDRY Alexandre	Biostatistiques, Informatique médicale et technologies de communication
PU-PH	MORO Elena	Neurologie
PU-PH	MORO SIBILOT Denis	Pneumologie ; addictologie
MCU-PH	MORTAMET Guillaume	Pédiatrie
PU-PH	MOUSSEAU Mireille	Cancérologie ; radiothérapie
PU-PH émérite	MOUTET François	Chirurgie plastique, reconstructrice et esthétique ; brûlologie
MCF Ass. MG	ODDOL Christel	Médecine Générale
PR Ass. Méc.	ORMEZZANO Olivier	Cardiologie
MCU-PH	PACIFT Marie-Hélène	Biochimie et biologie moléculaire
PU-PH	PAILHE Régis	Chirurgie orthopédique et traumatologie
PU-PH	PAIDMIE Olivier	Anatomie
PU-PH	PARK Sophie	Hématologie ; Transfusion
PR Ass.MG	PAUMIER-DI SIBIURIES Françoise	Médecine Générale
PU-PH	PAYEN DE LA GARANDERIE Jean-François	Anesthésiologie-reanimation et médecine péri-opératoire
MCU-PH	PAYSANT François	Médecine légale et droit de la santé
MCU-PH	PELLETIER Laurent	Biologie cellulaire
PU-PH	PELLOUX Hervé	Parasitologie et mycologie
PU-PH	PEPIN Jean-Louis	Physiologie

CORPS	NOM-PRENOM	Discipline universitaire
PU-PH	PERARD DUMESTRE Chantal	Immunologie
PU-PH	PERENNOU Dominique	Médecine physique et de réadaptation
PU-PH	PERNOD Gilles	Chirurgie vasculaire ; Médecine vasculaire
MCF	PINSULT Nicolas	Sciences de la rééducation et de réadaptation
PU-PH	PIGIAT Christian	Chirurgie infantile
PU-PH	PISON Christophe	Pneumologie ; Addictologie
PU-PH	PLANTAZ Dominique	Pédiatrie
PU-PH	POIGNARD Pascal	Bactériologie-virologie ; Hygiène hospitalière
PU-PH émérite	POLACK Benoît	Hématologie ; Transfusion
PU-PH	POLOSAN Mircea	Psychiatrie d'adultes ; Addictologie
PU-PH émérite	RAMBEAUD Jean Jacques	Urologie
PU-PH	RAY Pierre	Biologie et médecine du développement et de la reproduction ; gynécologie médicale
PR Ass. Méd.	RECHE Fabian	Chirurgie viscérale et digestive
MCU-PH	RI NDIH John	Biochimie et biologie moléculaire
MCU-PH émérite	RIALLE Vincent	Biostatistiques, informatique médicale et technologies de communication
PU-PH	RIETHMULLER Didier	Gynécologie obstétrique ; gynécologie médicale
PU-PH	RIGHINI Christian	Oto-rhino-laryngologie
PU-PH émérite	ROMANET Jean Paul	Ophthalmologie
PU-PH	ROSTAING Lionel	Néphrologie
PU-PH	ROUSTIT Matthieu	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
MCU-PH	ROUX-BUISSON Nathalie	Biochimie et biologie moléculaire
PR Ass.MG	ROYER DE VERICOURT Guillaume	Médecine Générale
PU-PH émérite	SARAGAGLIA Dominique	Chirurgie orthopédique et traumatologie
MCU-PH	SATHI Véronique	Génétique
PU-PH	SAUDOU Frédéric	Biologie cellulaire
PU-PH	SCHEMME Sébastien	Oto-rhino-laryngologie
PU-PH	SCHWEBEL Carole	Médecine intensive-réanimation
PU-PH	SCHEIDT Virginie	Médecine légale et droit de la santé
PU-PH	SEIGNEURIN Arnaud	Epidémiologie, économie de la santé et prévention
PU-PH	SPIAR Hafaïlle	Chirurgie vasculaire ; médecine vasculaire
PU-PH émérite	STAHL Jean-Paul	Maladies infectieuses ; Maladies tropicales
PU-PH	STANKI Françoise	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
MCU-PH	STASIA Marie-José	Biochimie et biologie moléculaire
PU-PH	STURM Nathalie	Anatomie et cytologie pathologiques
PU-PH	TAMISIER Renaud	Physiologie
PU-PH	TIRZI Nicolas	Médecine intensive-réanimation
PU-PH	THEVENON Julien	Génétique
PU-PH	TOILLANT Anne-Claire	Pneumologie ; Addictologie
PU-PH	TONETTI Jérôme	Chirurgie orthopédique et traumatologie
PU-PH	TOUSSAINT Bertrand	Biochimie et biologie moléculaire
PU-PH	VALMARY-DEGANO Séverine	Anatomie et cytologie pathologiques
PU-PH	VANZULLO Gérald	Cardiologie
PU-PH	VUILLEZ Jean Philippe	Biophysique et médecine nucléaire
PU-PH	WEIL Georges	Epidémiologie, économie de la santé et prévention
PU-PH	ZAQOU Philippe	Néphrologie
PU-PH émérite	ZARSKI Jean-Pierre	Gastroentérologie ; hépatologie ; addictologie

## TABLE DES MATIÈRES

Remerciements.....	10
Résumé.....	12
Abstract.....	13
Abréviations.....	14
Introduction.....	15
Matériels et méthodes.....	21
Résultats.....	27
Discussion.....	43
Conclusion.....	47
Bibliographie.....	49
Serment d'Hippocrate.....	52

## TABLE DES TABLEAUX

**Tableau 1 :** Interprétation du coefficient Kappa selon Landis et al.

**Tableau 2 :** Résumé des interprétations des imageries en fonction du site étudié

**Tableau 3 :** Résumé des caractéristiques démographiques de la population étudiée

**Tableau 4 :** Tableau de contingence et données de performances diagnostiques relatif à l'ensemble des articulations

**Tableau 5 :** Tableau de contingence et données de performances diagnostiques relatif à l'articulation de la clavicule, de l'épaule et de l'avant bras

**Tableau 6 :** Tableau de contingence et données de performances diagnostiques relatif à l'articulation du coude et de l'avant-bras

**Tableau 7 :** Tableau de contingence et données de performances diagnostiques relatif à l'articulation du poignet, de la main et des doigts

**Tableau 8 :** Tableau de contingence et données de performances diagnostiques relatif à l'articulation du cadre obturateur, du bassin, et du fémur

**Tableau 9 :** Tableau de contingence et données de performances diagnostiques relatif à l'articulation du genou et de la jambe

**Tableau 10 :** Comparatif des interprétations radiologiques en fonction de la localisation

**Tableau 11 :** Comparatif des interprétations radiologiques en fonction de l'âge

**Tableau 12 :** Comparatif des interprétations radiologiques en fonction du sexe

**Tableau 13 :** Analyse multivariée de la bonne reconnaissance par l'IA de la lésion par localisation par rapport au membre inférieur intermédiaire

**Tableau 14 :** Analyse multivariée de la bonne reconnaissance par l'IA de la lésion par localisation par rapport au membre supérieur intermédiaire

## TABLE DES FIGURES

**Figure 1** : Sous famille de l'intelligence artificielle

**Figure 2** : Différence entre Machine Learning et Deep Learning

**Figure 3** : Réseau de Neurones Convolutif (CNN)

**Figure 4** : Les 3 étapes de développement d'une IA

**Figure 5** : Exemple de radiographies après identification par le logiciel Rayvolve® des régions d'intérêt et capture secondaire envoyée sur le PACS

**Figure 6** : Localisation de la région anatomique lorsqu'il existe un désaccord entre l'orthopédiste et l'urgentiste

**Figure 7** : région anatomique étudiée en fonction de l'âge

**Figure 8** : Diagramme de flux

**Figure 9** : Exemple d'une radiographie de clavicule droite avec comme diagnostic émis par le Gold Standard une disjonction acromio claviculaire, détectée par l'IA comme une fracture

**Figure 10** : Exemple d'une radiographie de l'épaule gauche, avec comme diagnostic émis par le Gold Standard une fracture du trochiter, détectée par le logiciel d'IA

**Figure 11** : Exemple d'une radiographie de coude gauche, avec comme diagnostic émis par le Gold Standard une simple contusion, détectée par l'IA comme une fracture

**Figure 12** : Exemple d'une radiographie du coude gauche, avec comme diagnostic émis par le Gold Standard une fracture de la tête radiale, détectée par le logiciel d'IA comme une fracture

**Figure 13** : Exemple d'une radiographie de la main droite, avec comme diagnostic émis par le Gold Standard une contusion, détectée par l'IA comme une fracture du scaphoïde

**Figure 14:** Exemple d'une radiographie du poignet gauche, avec comme diagnostic émis par le Gold Standard une simple contusion, détectée par l'IA comme une fracture

**Figure 15 :** Exemple d'une radiographie du bassin de face, avec comme diagnostic émis par le Gold Standard une fracture de la branche ischio-pubienne gauche, non détectée par l'IA

**Figure 16 :** Exemple d'une radiographie du genou droit, avec comme diagnostic émis par le Gold Standard une simple contusion, détectée à tort par l'IA comme une fracture

**Figure 17 :** Exemple d'une radiographie de la cheville gauche, avec comme diagnostic émis par le Gold Standard une entorse de la cheville, détectée par l'IA comme une fracture

**Figure 18 :** Exemple d'une radiographie du pied gauche, avec comme diagnostic émis par le Gold Standard une simple contusion, détectée à tort par l'IA comme une fracture

## REMERCIEMENTS

### Aux membres du jury :

Monsieur le Professeur Gilbert Ferretti, vous me faites l'honneur de présider ce jury. Votre expertise est primordiale. Je vous remercie profondément de l'intérêt que vous y avez porté et de vos conseils.

Monsieur le Docteur Jean-Jacques Banihachemi, directeur de thèse, je te remercie de m'avoir accompagné dans ce projet. Merci pour l'enseignement que tu m'as apporté lors de mon semestre au SAU de l'hôpital Sud du CHU de Grenoble.

Monsieur le Professeur Régis Pailhe, merci d'avoir accepté de siéger dans ce jury. Vos compétences et votre expérience dans le domaine de l'orthopédie sont précieuses pour juger ce travail.

Monsieur le Professeur Damien Viglino, je te remercie pour l'aide que tu m'as apportée dans ce travail, pour ta disponibilité et tes conseils.

Monsieur le Docteur Pierre Chapuis, c'est un honneur que vous ayez accepté de juger mon travail.

### A mes proches :

À mes parents, pour m'avoir soutenue durant toutes ces années et de m'avoir encouragée à poursuivre dans la voie que j'ai choisie malgré les difficultés. Merci aussi de m'avoir permis de faire autant de beaux voyages et de soutenir mes projets. Je vous souhaite de belles aventures en parapente et en vélo!

À mes frères et sœurs, Anne-Gaëlle, Camille et Bruno, vous avez pris des chemins bien différents du mien (et d'ailleurs je n'ai toujours pas très bien compris en quoi consistent vos métiers les filles!) mais qui me rappellent qu'il existe un monde en dehors de la médecine. Camille, mille mercis pour ton aide dans cette thèse.

À tous le reste de ma famille : j'ai beaucoup de chance de vous avoir !

À la team Poncillon ça pèse, Clément et Laetitia, pour votre bonne humeur en toutes circonstances et votre motivation sans faille. Il y aurait tant à dire et à remercier ! Encore de beaux voyages ensemble nous attendent (et qui sait, peut être même des soirées Sold Out ?)

À la bande de copains de Clermont Ferrand, pour avoir rendu ces années d'externat extraordinaires.

À mes espagnoles préférées, Laura, Ali, et le reste de la team Erasmus, quelle année folle en Pologne à vos côtés ! J'espère que l'on restera encore longtemps en contact !

À Sophie, pour cette belle amitié depuis cette fameuse nuit en Egypte, profite de ton voyage !

À mes supers copains de Grenoble : Léna, mon petit rayon de soleil, qu'elle chance de t'avoir, et qu'elle te va bien ma tenue à paillettes. Valentine, merci pour tous les moments partagés, à la montagne, à Madère, en footing ou sans skis sur une piste noire verglacée. Marion, Fannie, et Vaea, je ne pouvais pas mieux rêver comme coloc. Julien, merci d'être présent pour moi et toujours à l'écoute, tu es génial. Adrian, j'aime ton enthousiasme infailible et ta positivité en toutes circonstances. Marie-Élise, je ne te remercie pas pour les séances de fractionné !

Et à tous le reste de l'internat Grobid-19, c'est rare de rencontrer si rapidement autant de personnes avec qui bien s'entendre! Je n'oublierai jamais cette première année d'internat à vos cotés.

À toi Guillaume, je suis tellement heureuse de t'avoir à mes côtés. Tu as toujours les bonnes paroles et les bons conseils pour me faire aller de l'avant. Merci pour tout ce que tu m'apportes.

À la promo Barre des Ecrins, et à tous mes co-internes du SAU de Nord et de Sud, de gériatrie, de pédiatrie, de réanimation médicale et d'USIC.

À toutes les équipes de médecins qui m'ont accueillie chaque semestre tout au long de mon internat et auprès de qui j'ai tant appris. Merci Sara pour ta patience avec moi et de m'avoir fait autant aimer la cardiologie.

Aux équipes paramédicales des différents services où j'ai travaillé. Vous avez rempli mon internat de supers souvenirs.

## RÉSUMÉ

*Introduction* : Au sein du Service d'Accueil des Urgences (SAU), jusqu'à 40 % des diagnostics omis sont des fractures (1). Il peut en résulter de lourdes conséquences pour le patient et également pour la structure de soin. Les radiographies dans les situations d'urgence sont souvent lues, par nécessité, par des cliniciens en médecine d'urgence qui manquent d'expertise spécialisée en radiologie et en traumatologie. De plus, le flux continu élevé de patients dans les urgences cause une fatigue et une susceptibilité aux erreurs d'interprétation (2, 3). L'intelligence artificielle pourrait être utilisée comme un outil d'aide aux urgentistes afin de faire face au volume et à la charge de travail sans cesse croissante (4). Ainsi, nous souhaitons mener cette étude afin d'évaluer les performances d'un logiciel d'intelligence artificielle dans l'évaluation des fractures dans des conditions réelles d'utilisation et de définir sa place dans la démarche diagnostique.

*Matériels et méthodes* : Cette étude prospective monocentrique a été menée sur 1808 radiographies standards de 1523 patients ayant consulté aux urgences de l'hôpital Sud du CHU de Grenoble du 26 avril au 18 juin 2022. Les radiographies de patients mineurs, ainsi que les radiographies du rachis, du thorax du crâne et de la face n'ont pas été incluses.

Le Gold Standard a été défini comme le diagnostic final délivré par un orthopédiste sénior, lors du staff orthopédique journalier, lorsque toutes les radiographies réalisées aux urgences la veille sont relues. L'algorithme du logiciel d'intelligence artificielle a été évalué en comparant le Gold Standard et la prédiction délivrée par l'algorithme. Nous avons cherché à évaluer s'il existe une relation entre la précision du logiciel d'intelligence artificielle et des critères qualitatifs comme l'âge, le sexe et la région anatomique.

*Résultats* : Avec un Kappa de Cohen calculé à 0.81, il existe un accord presque parfait entre le diagnostic du chirurgien orthopédique sénior et la prédiction du logiciel d'intelligence artificielle. L'interprétation radiologique par le logiciel d'IA affiche, en référence au Gold Standard, une sensibilité de 89.5% et une spécificité de 92,5 %. Les résultats de l'étude mettent en évidence une plus faible sensibilité lors de l'interprétation par l'IA des radiographies du genou et de la jambe. Les performances du logiciel d'IA diminuent avec l'âge. Le sexe n'a pas d'impact sur les performances du logiciel.

*Conclusion* : L'intelligence artificielle est une piste d'amélioration des performances diagnostiques des médecins urgentistes dans l'interprétation des radiographies osseuses, notamment des docteurs juniors, et qui pourrait partiellement pallier au manque de spécialistes disponibles. Néanmoins, le médecin doit rester maître de son diagnostic, et l'IA doit rester un outil d'aide au diagnostic sur lequel s'appuyer en cas de besoin.

*Mots clés* : Intelligence artificielle, apprentissage profond, radiologie d'urgence et de traumatologie, radiographie.

*Filière* : médecine d'urgence, orthopédie, radiologie

## ABSTRACT

*Introduction:* Within the Emergency Department, 3.7% to 4.1% of fractures are misdiagnosed, and 70% to 80% of omitted diagnoses in the emergency room concern fractures. This can have serious consequences for the patient and also for the care structure. During emergencies, X-rays are often read out of necessity by emergency medicine clinicians who lack subspecialty expertise in radiology and orthopaedic. In addition, the continuous high flow of patients in the emergency room causes fatigue and susceptibility to misinterpretations. Artificial Intelligence could be used to help emergency physicians cope with the volume and workload that continues to increase in imaging. Thus, we will conduct this study to evaluate the performance of an Artificial Intelligence software in evaluating skeletal fractures under real conditions of use and to define its potential uses in clinical practice.

*Materials and methods:* This prospective study was conducted on 1808 X-rays of 1523 patients who consulted in the ER of the Hospital Sud de Grenoble from April 26 to June 18, 2022. X-rays of minor patients, as well as X-rays of the spine, thorax and skull, were not included. The Gold Standard was defined as the final diagnosis delivered by a senior orthopaedist, blind to the algorithm's prediction. The Artificial Intelligence software algorithm was evaluated by comparing the Gold Standard and the prediction delivered by the algorithm. We sought to assess whether there is a relationship between the accuracy of artificial intelligence software and qualitative criteria such as age, sex and anatomical region.

*Results:* With Cohen's Kappa calculated at 0.81, there is near perfect agreement between the senior orthopaedic surgeon's diagnosis and the artificial intelligence software's prediction. The radiological interpretation by the AI software shows, with reference to the Gold Standard, a sensitivity of 89.5% and a specificity of 92.5%. The results of the study further highlight a low sensitivity when interpreting knee and leg X-rays by AI. The performance of AI software declines significantly with age. Gender has no significant impact on software performance.

*Conclusion:* Artificial intelligence is a way to improve the diagnostic performance of emergency doctors in the interpretation of x-rays of the skeleton, especially junior doctors, which could partially compensate for the lack of available specialists. Nevertheless, the doctor must remain in control of his diagnosis, and AI must remain a diagnostic tool to rely on when needed.

*Keywords:* Artificial intelligence, deep learning, emergency and trauma radiology, radiography.

*Sector:* emergency medicine, orthopaedics, radiology

## LISTE DES ABRÉVIATIONS

IA : Intelligence Artificielle

CNN : Convolutional neural Network (en Français, réseau de neurones à convolution)

DL : Deep Learning (en Français, apprentissage profond)

IRM : Imagerie par Résonance Magnétique

SAU : Service d'Accueil des Urgences

CHU : Centre Hospitalier Universitaire

CE : conformité Européenne

ISO : Organisation Internationale de Normalisation

PACS : Picture Archiving and Communication System (en Français : système d'archivage et de transmission d'images)

RGPD : Règlement Général sur la Protection des Données

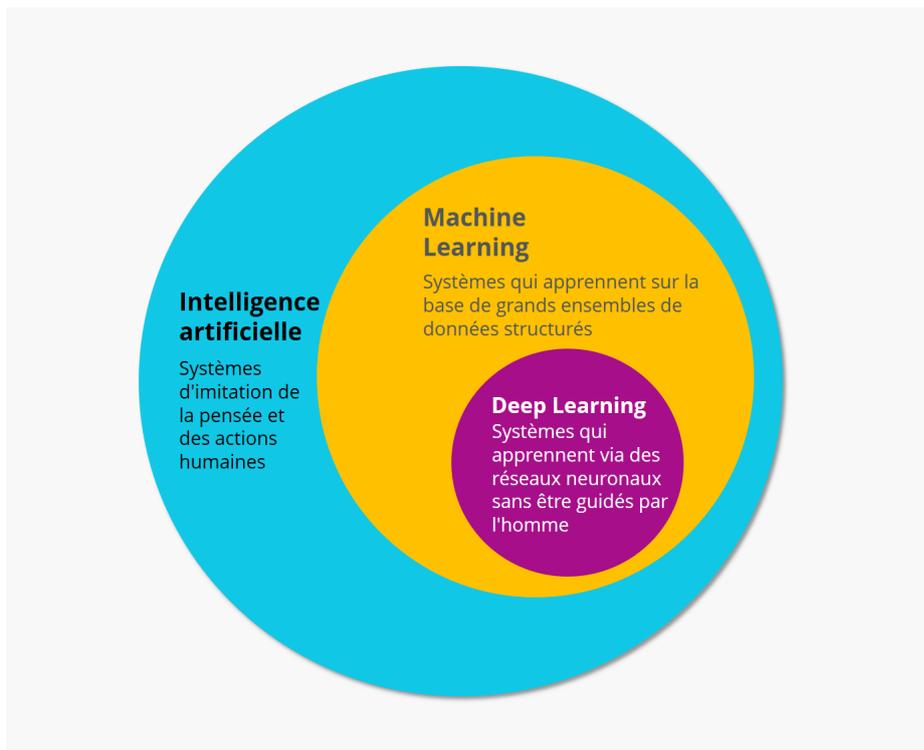
## INTRODUCTION

### *Définition de l'intelligence artificielle*

Créée dans les années 50 par John McCarthy, l'intelligence artificielle (IA) est un terme général qui implique l'utilisation d'un ordinateur pour modéliser un comportement intelligent avec une intervention humaine minimale (5).

Comme l'intelligence humaine, l'IA a besoin d'apprendre pour se développer et réaliser une tâche de façon autonome. Deux principales techniques d'apprentissage sont utilisées : le « Machine Learning » (apprentissage automatique) et le « Deep Learning » (apprentissage profond) (6).

**Figure 1** : Les sous famille de l'intelligence artificielle (7).



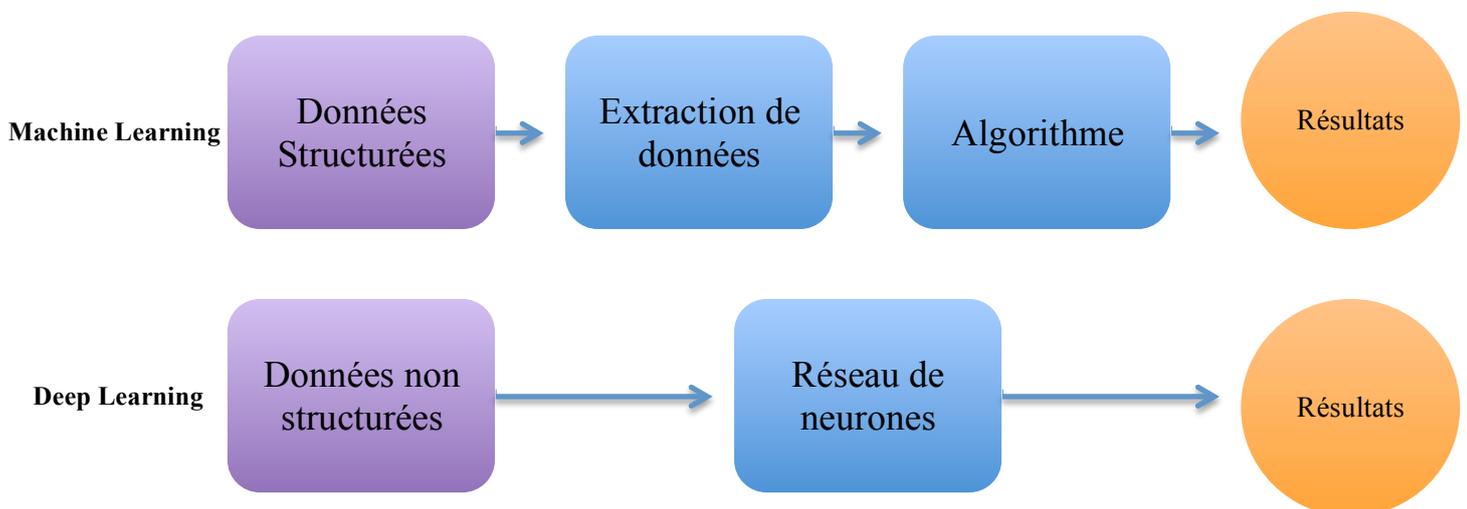
Dans le « Machine Learning », des algorithmes reçoivent des données d'entrée quantitatives et structurées. Un « Data Scientist » va extraire les données qui vont influencer la prédiction : les « variables déterminantes » (par exemple, les

variables peuvent être l'âge, le sexe, le poids...). En fonction de ces données, des algorithmes vont prédire une donnée de sortie selon des analyses statistiques. On parle d'apprentissage supervisé puisque l'intervention humaine est nécessaire.

Le « Deep Learning » est un sous-type du « Machine Learning ». Le système n'a pas besoin d'être guidé par l'homme, les « variables déterminantes » sont calculées automatiquement par l'algorithme. Il fonctionne à partir de plusieurs couches de réseaux neuronaux. Ces réseaux neuronaux fonctionnent de façon similaire à ce que font nos réseaux de neurones en réponse aux signaux nerveux qui leur sont destinés. Selon le type et la fréquence des messages reçus, certains réseaux de neurones se développent quantitativement et qualitativement alors que d'autres régressent (7).

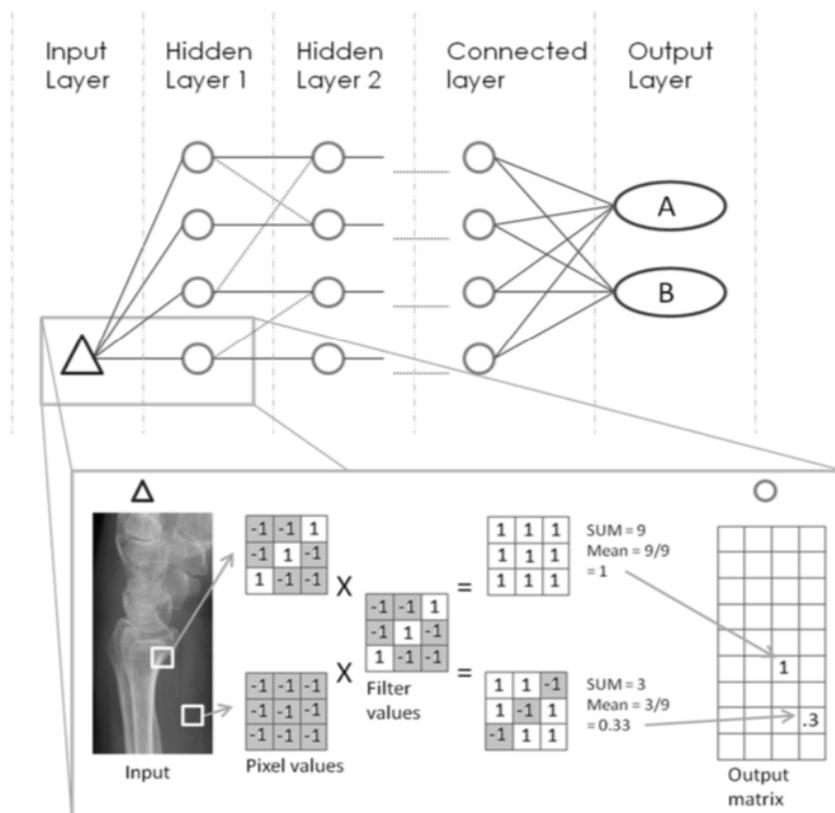
C'est le « Deep Learning » qui est utilisé dans les logiciels d'IA de détection de fracture sur les radiographies.

**Figure 2** : Différence entre Machine Learning et Deep Learning



Parmi les nombreuses architectures existantes de réseaux de neurones, le Réseau de Neurones Convolutif (CNN) est un type de réseau inspiré du cortex visuel. La convolution consiste à appliquer successivement des filtres de tailles définies sur une image. Dans le cas de la radiologie, pour une zone de la radiographie définie par la taille du filtre, les caractéristiques numériques propres à chaque pixel sont extraites afin d'obtenir une valeur représentative de la zone de pixels. Cela permet de générer une image de dimension plus petite afin de réduire la quantité d'information. Ces caractéristiques seront ensuite intégrées à un réseau de neurones, qui, en fonction des valeurs des différentes zones de pixels, localise une zone fracturée sur la radiologie (8).

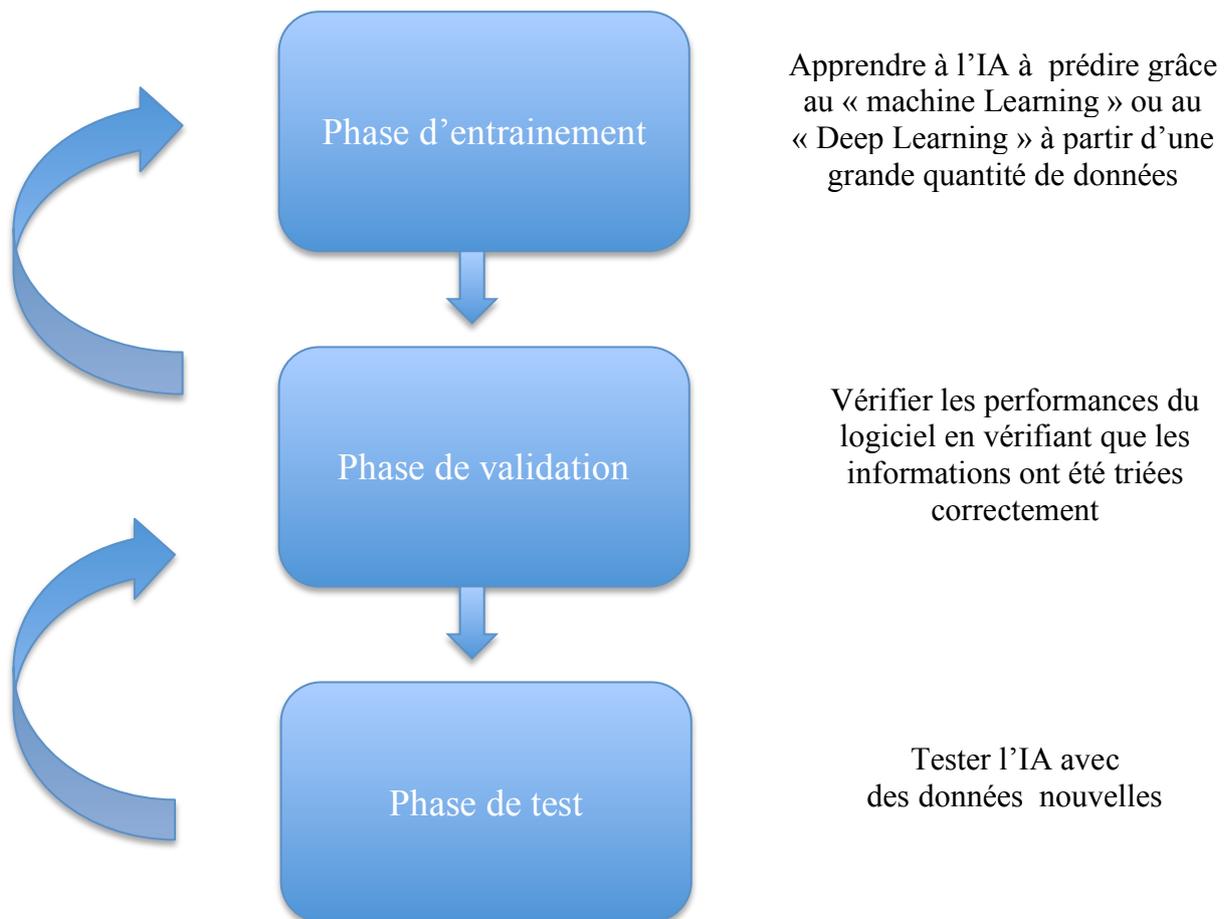
**Figure 3** : Réseau de Neurones Convolutif (CNN) (15).



La seconde étape dans le développement d'une IA après la phase d'entraînement consiste à déterminer si les informations reçues à la phase initiale ont été triées correctement afin de vérifier les performances du logiciel. Si l'IA réussit le test,

elle passe à l'étape suivante. Sinon, elle revient à la case de départ pour un ajustement de ses défauts. La dernière étape consiste à tester l'IA avec des données qui lui sont nouvelles. Si l'IA réussit l'épreuve, cela veut dire qu'elle est opérationnelle.

**Figure 4 :** Les 3 étapes de développement d'une IA



### ***Intérêt et justification de l'étude***

L'intelligence artificielle est un domaine de recherche en plein essor, et c'est un outil clé en radiologie diagnostique. Depuis 2012, le Deep Learning est rapidement devenu une méthode permettant d'améliorer les performances en analyse d'images médicales avec l'utilisation des réseaux de neurones convolutifs (CNN).

Des publications ont démontré que les logiciels d'IA peuvent produire des résultats égaux aux performances humaines dans la détection de la rétinopathie chez les

personnes diabétiques via des photographies du fond d'œil rétinien, la détection de lésions mammographiques et la reconnaissance du cancer de la peau (9,10,11).

Des niveaux similaires de précision diagnostique n'ont pas encore été rapportés dans l'analyse de radiographies osseuses; cependant, quelques études prometteuses ont été publiées.

L'étude d'Olczak et al. est une des premières études à déterminer la faisabilité de l'utilisation du Deep Learning pour les radiographies du squelette. Dans cette étude, la performance d'un logiciel d'IA est comparée à un Gold standard dans 400 radiographies du poignet, de la main et de la cheville. La concordance globale entre le logiciel d'IA et le Gold Standard dans la détection des fractures est de 83 % (12).

L'utilisation du DL permet de détecter les fractures de l'humérus proximal sur les radiographies standards de face de l'épaule avec une sensibilité de 99 % (13).

L'étude de Yoon et al. (14) montre que sur 11 838 radiographies du scaphoïde, les réseaux de neurones formés pour distinguer les fractures du scaphoïde ont atteint une sensibilité et une spécificité globales de 87,1 % et 92,1 %, respectivement.

Des résultats satisfaisants ont aussi été obtenus sur l'utilisation du DL pour l'analyse de radiographies latérales du poignet (15) ainsi que les fractures de hanche (16,17).

Enfin, la littérature actuelle indique des performances d'IA similaires à celles des humains pour la détection des lésions cartilagineuses sur l'IRM et moins de variabilité que les humains dans le classement de la gravité de l'arthrose du genou sur les radiographies (18).

Cependant, très peu d'études ont cherché à évaluer les performances du Deep Learning dans le cadre de lésions traumatiques dans un service d'urgence en utilisant une cohorte indépendante, sans aucun prétraitement ni tri des données d'entrée.

Au sein du Service d'Accueil des Urgences (SAU), jusqu'à 40 % des diagnostics omis sont des fractures (1). Les conséquences pour le patient sont multiples : douleurs, perte de fonction et d'autonomie, risque de pseudarthrose, moins bons résultats à long terme ; ces erreurs peuvent nécessiter à terme une chirurgie. Il existe un impact pour le patient d'une part mais également pour la structure de soin. Outre une morbidité augmentée, les erreurs de diagnostic relatives aux fractures représentent une cause fréquente de dépôt de plainte contre les établissements hospitaliers et les professionnels de santé. Par exemple, aux États-Unis, ces erreurs de diagnostic représentent le deuxième motif d'action en justice, après les néoplasies mammaires (19).

Les radiographies au SAU sont souvent lues, par nécessité, par des cliniciens en médecine d'urgence qui manquent d'expertise surspécialisée en radiologie osseuse. Le flux continu élevé de patients dans les urgences cause une fatigue et une susceptibilité aux erreurs d'interprétation (2,3). De plus, ces erreurs augmentent en proportion en nuit profonde (20), du fait du surmenage et du manque de spécialistes disponibles. L'IA pourrait être utilisée pour aider les urgentistes à faire face au volume et à la charge de travail sans cesse croissante (4).

Enfin, le plus grand défi est peut-être l'intégration des technologies d'IA dans la pratique. Malgré des succès de recherche importants, il reste très peu d'exemples d'algorithmes d'IA intégrés avec succès dans la pratique clinique quotidienne dans des systèmes de santé (21,22). Le manque d'information vis à vis des logiciels d'IA (23,24), la méfiance des cliniciens ainsi que celle des patients étant liés (21).

L'objectif principal de cette étude est d'évaluer les performances d'un logiciel d'intelligence artificielle dans l'évaluation des fractures dans des conditions réelles d'utilisation et de définir sa place dans la démarche diagnostique.

L'objectif secondaire est de déterminer si les performances du logiciel d'intelligence artificielle varient selon l'âge, le sexe et la région anatomique.

## **MATÉRIEL ET MÉTHODE**

### ***Conflits d'intérêts et affiliation à l'entreprise Azmed***

Notre étude est non rémunérée et non affiliée à l'entreprise Azmed. Cette entreprise, développant le logiciel d'IA, n'a financé par aucun moyen les différentes étapes de l'étude. En outre, les professionnels de santé n'ont aucun conflit d'intérêt à déclarer en lien avec le sujet d'étude.

### ***Autorisation de l'étude et protection des données personnelles***

Cette étude a été réalisée dans le cadre réglementaire de la « recherche n'impliquant pas la personne humaine » avec autorisation du CHU Grenoble-Alpes, et information aux patients de leur possibilité de s'opposer à tout moment à l'utilisation de leurs données médicales et des données recueillies. Dans ce cadre, la non-opposition était recherchée. Un affichage des études en cours et de l'utilisation des données patients est réalisé dans le service des urgences Sud du CHU Grenoble Alpes à titre d'information avec possibilité d'opposition au recueil des données de la part des patients par renseignement d'un formulaire dédié. Aucune opposition à l'utilisation des données n'a été relevée.

### ***Logiciel d'intelligence artificielle d'aide à la détection des fractures : Rayvolve***

Rayvolve est le logiciel d'intelligence artificielle utilisé aux urgences de Grenoble depuis juillet 2021. Il a été créé par une startup française, AZmed.

C'est un dispositif médical certifié CE répondant aux exigences réglementaires européennes et conçu conformément à la norme ISO 13485.

Il a été formé sur plus de 21 138 radiographies dont plus de 10 000 radiographies sans fracture.

### ***Population d'étude***

Parmi les 3606 patients ayant consulté aux urgences de l'hôpital Sud de Grenoble du 26 avril au 18 juin 2022, la population d'étude a été sélectionnée prospectivement et incluse dans la base de données cliniques informatisée au regard des critères d'inclusion suivants : patient s'étant présenté aux urgences de l'Hôpital Sud de Grenoble du 26 avril au 18 juin 2022, pour lequel des radiographies ont été réalisées et pour qui un diagnostic médical a été émis.

Les critères d'exclusion étaient : consultation aux urgences pour toute cause non traumatique, historique clinique flou, patient mineur, consultation pour traumatisme rachidien, thoracique ou crânien, pas de radiographie réalisée, pas d'interprétation réalisée par le logiciel d'IA.

De même, les radiographies de contrôle « post-réduction d'une luxation » ou « réfection de plâtre » ainsi que les radiographies avec utilisation du Telos<sup>®1</sup> ont également été exclues.

Au total, nous avons analysé 1808 radiographies.

### ***Données recueillies***

Les données enregistrées ont été les suivantes : Le sexe et l'âge du patient au moment du passage aux urgences, la ou les régions anatomiques explorées, la prédiction du logiciel d'intelligence artificielle et les conclusions du sénior orthopédique.

Ces données ont été recueillies sous forme de variables binaires (concordance entre l'IA et l'orthopédiste ou non) et de chaîne de texte (décrivant le diagnostic délivré).

Lorsque, pour un seul patient, il existait plusieurs radiographies de différentes régions anatomiques, les ensembles radiographiques par région ont été traités un par un.

---

<sup>1</sup> Il s'agit d'un examen radiographique standard durant lequel une contrainte de 150 à 250 Newton est exercée sur une articulation (cheville ou genou) dans le plan latéral et antéro-postérieur à l'aide d'une machine spécifiquement conçue dans ce but afin d'évaluer la laxité.

Par conséquent, un patient peut apparaître plusieurs fois dans le tableau, pour chaque région anatomique.

Toutes les images radiographiques ont été incluses, quelle que soit leur qualité. Aucun prétraitement n'a été effectué. Ces données sont représentées par un diagramme de flux (figure 8).

### ***Anonymisation des données à caractère personnel***

Afin de garantir une anonymisation des données relatives aux patients, un numéro a été attribué aléatoirement à chaque patient. Cela a permis une interprétation anonymisée de l'ensemble des examens inclus dans l'étude tel que le prévoit la loi sur la protection des données personnelles.

### ***Lecture et interprétation des imageries médicales***

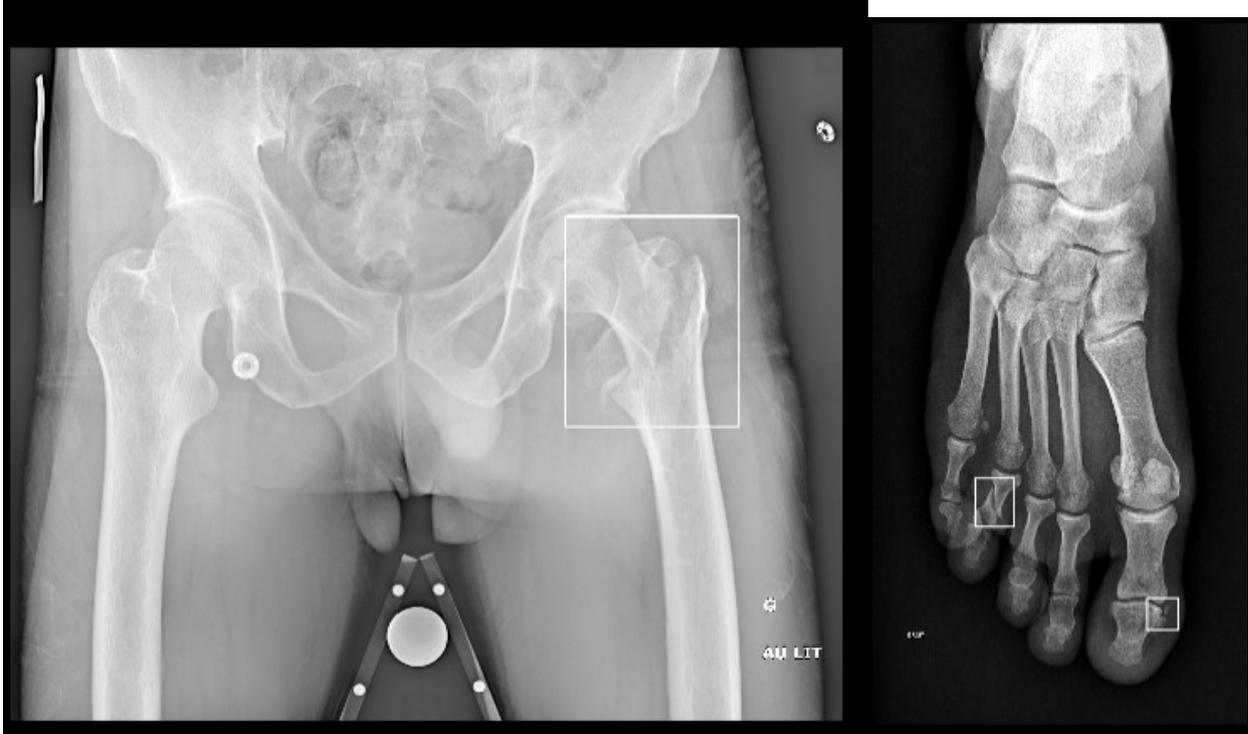
L'ensemble des interprétations ont été réalisées de façon prospective pendant 54 jours, par un orthopédiste senior, lors du staff orthopédique journalier, quand toutes les radiographies de la veille sont relues.

Les données cliniques, ainsi que le diagnostic émis par le médecin urgentiste après la consultation aux urgences du patient, étaient accessibles à l'orthopédiste.

L'orthopédiste sénior pouvait être différent selon les jours de la semaine.

Le médecin urgentiste et l'orthopédiste étaient en aveugle du diagnostic formulé par l'IA lors de la lecture de la radiographie.

**Figure 5** : Exemple de radiographies après identification par le logiciel Rayvolve® des régions d'intérêt et capture secondaire envoyée sur le PACS



### *Analyse statistique*

Le Gold Standard a été défini comme le diagnostic final délivré par un orthopédiste sénior, lors d'une relecture des radiographies au cours du staff orthopédique journalier. Durant cette seconde relecture, l'orthopédiste confirmait le diagnostic émis par le médecin urgentiste. Si il y avait un désaccord entre l'orthopédiste et l'urgentiste, c'est le diagnostic de l'orthopédiste qui était retenu. Les données cliniques du patient étaient accessibles à l'orthopédiste, et l'urgentiste ayant fait le diagnostic était présent lors du staff. L'orthopédiste et le médecin urgentiste étaient à l'aveugle de la prédiction de l'algorithme lors de l'interprétation de la radiographie.

L'algorithme du logiciel d'intelligence artificielle a ensuite été évalué en comparant le Gold Standard avec la prédiction délivrée par l'algorithme.

Les données d'interprétation ont été analysées par un test de concordance diagnostique dit test non paramétrique Kappa de Cohen (K). Ce test permet de chiffrer l'accord entre deux ou plusieurs observateurs lorsque les jugements sont qualitatifs. L'estimation de l'accord entre les jugements catégoriels appliqués aux mêmes objets, fournis par deux ou plusieurs observateurs, prend en compte la concordance aléatoire et permet de s'en affranchir. La « concordance » signifie la proportion de sujets pour lesquels il y a accord entre les observateurs (25). On évalue la concordance grâce au calcul d'un coefficient nommé kappa. Ce coefficient varie de 0 à 1 en fonction du niveau de concordance observé.

Le calcul du K se fait de la manière suivante :

$$K = \frac{Pr(a) - Pr(e)}{1 - Pr(e)}$$

où  $Pr(a)$  est la proportion de l'accord entre codeurs et  $Pr(e)$  la probabilité d'un accord aléatoire.

Le tableau, ci-après, expose l'interprétation du coefficient kappa. Cette interprétation est basée sur une étude unique et subjective (26) et reste discutable.

**Tableau 1** : Interprétation du coefficient Kappa selon Landis et al. (26)

Coefficient Kappa	Estimation du degré de concordance
0	mauvais
0 à 0.2	négligeable
0.2 à 0.4	faible
0.4 à 0.6	moyen
0.6 à 0.8	bon
0.8 à 1	excellent

Les paramètres diagnostiques du logiciel ont également été calculés tel que :

- La sensibilité, mesurant la proportion de fractures correctement identifiées par le logiciel d'intelligence artificielle. Les patients avec une fracture correctement identifiée sont considérés comme de vrais positifs (VP), tandis que les patients avec une fracture, mais qui n'est pas identifiée par l'algorithme, sont considérés comme des faux négatifs (FN).
- La spécificité, mesurant la proportion d'absence de fracture correctement identifiée. Les patients n'ayant pas de fracture qui sont identifiés par l'algorithme comme ayant une fracture sont considérés comme des faux positifs (FP).
- La valeur prédictive positive (VPP) correspond à la probabilité que la fracture soit présente lorsque le test est positif.
- La valeur prédictive négative (VPN) est, à contrario, la probabilité que la fracture ne soit pas présente lorsque le test est négatif.

Nous avons utilisé une méthode de régression logistique afin de contrôler l'effet de caractéristiques quantitatives et qualitatives sur l'impact étudié, ici la concordance.

## RÉSULTATS

### *Caractéristiques de la population étudiée*

Du 26 avril au 18 juin 2022, 1808 radiographies de 1523 patients ont été incluses.

Les résultats sont présentés dans le tableau 2, et mettent en évidence une forte disparité du nombre de radiographies incluses en fonction de la localisation.

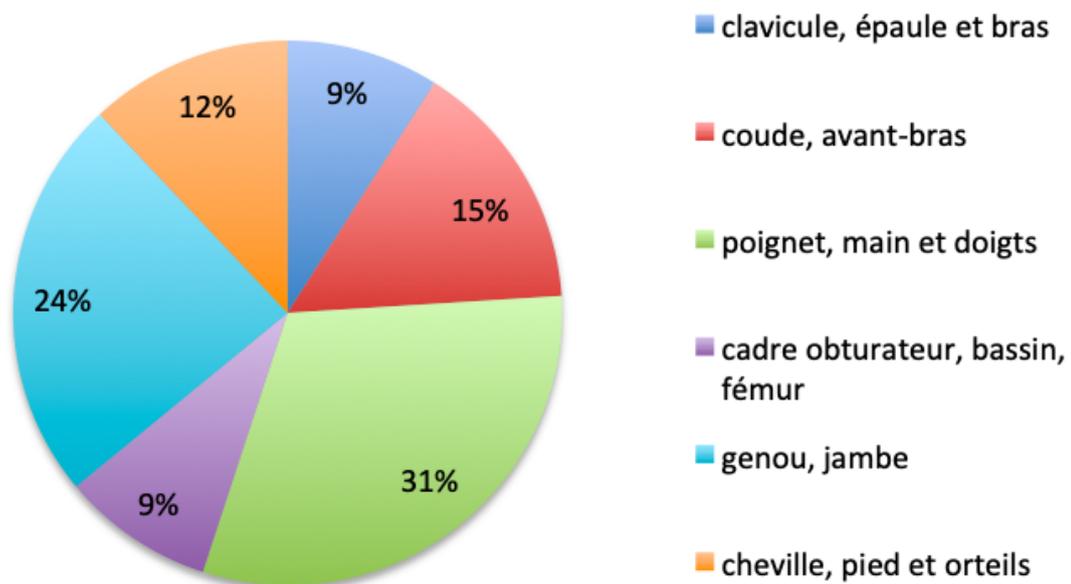
Au cours de l'interprétation des examens radiographiques par le Gold Standard, représenté par le senior orthopédique, 708 présentent une fracture (39.1%) et 1100 examens ne présentent pas de fracture (60.9%). Il existe une forte disparité des taux d'examens présentant une fracture. Seuls 13% des examens intéressants le genou et la jambe sont pathologiques contre 53% concernant le coude et l'avant-bras et 59% le cadre obturateur, bassin et fémur.

**Tableau 2 :** Résumé des interprétations des imageries selon le diagnostic du Gold Standard en fonction du site étudié

Localisation	Nombre d'examens	% du nombre d'examens	Fractures	Pas de fracture	% de fracture
membres supérieurs proximaux (clavicule, épaule et bras),	236	13,1 %	84	152	35%
membres supérieurs intermédiaires (coude, avant-bras)	135	<b>7,5 %</b>	72	63	53%
les membres supérieurs distaux (poignet, main et doigts),	507	28%	235	272	46%
membres inférieurs proximaux (cadre obturateur, bassin, fémur),	136	<b>7,5%</b>	80	56	59%
membres inférieurs intermédiaires (genou, jambe)	320	17,7%	42	278	<b>13%</b>
membres inférieurs distaux (cheville, pied et orteils).	474	26,2%	195	279	41%
Total	1808		708	1100	
% Total	100%	100%	<b>39.1%</b>	<b>60.9%</b>	

Le pourcentage de désaccord entre l'orthopédiste et l'urgentiste est de 3.6% et concerne en plus grande proportion les radiographies du poignet et de la main (31%) (Figure 6).

**Figure 6** : Localisation de la région anatomique lorsqu'il existe un désaccord entre l'orthopédiste et l'urgentiste



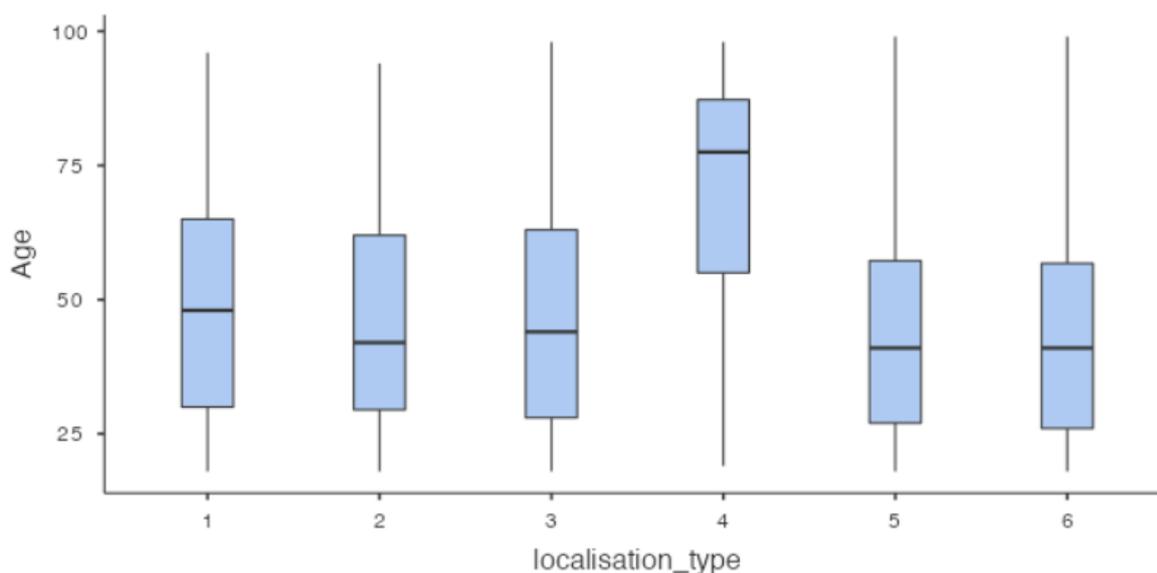
L'âge moyen est de 47 ans, la fourchette d'âge entre 18 et 99 ans, et la valeur médiane est de 45 ans. L'âge moyen chez l'homme est de 53 ans, et de 42 ans chez la femme. Concernant le sexe ratio, 45% sont des hommes et 55% sont des femmes (tableau 3).

L'âge moyen est plus élevé pour les radiographies des membres inférieurs proximaux (cadre obturateur, bassin, fémur) (figure 7).

**Tableau 3** : Résumé des caractéristiques démographiques de la population étudiée

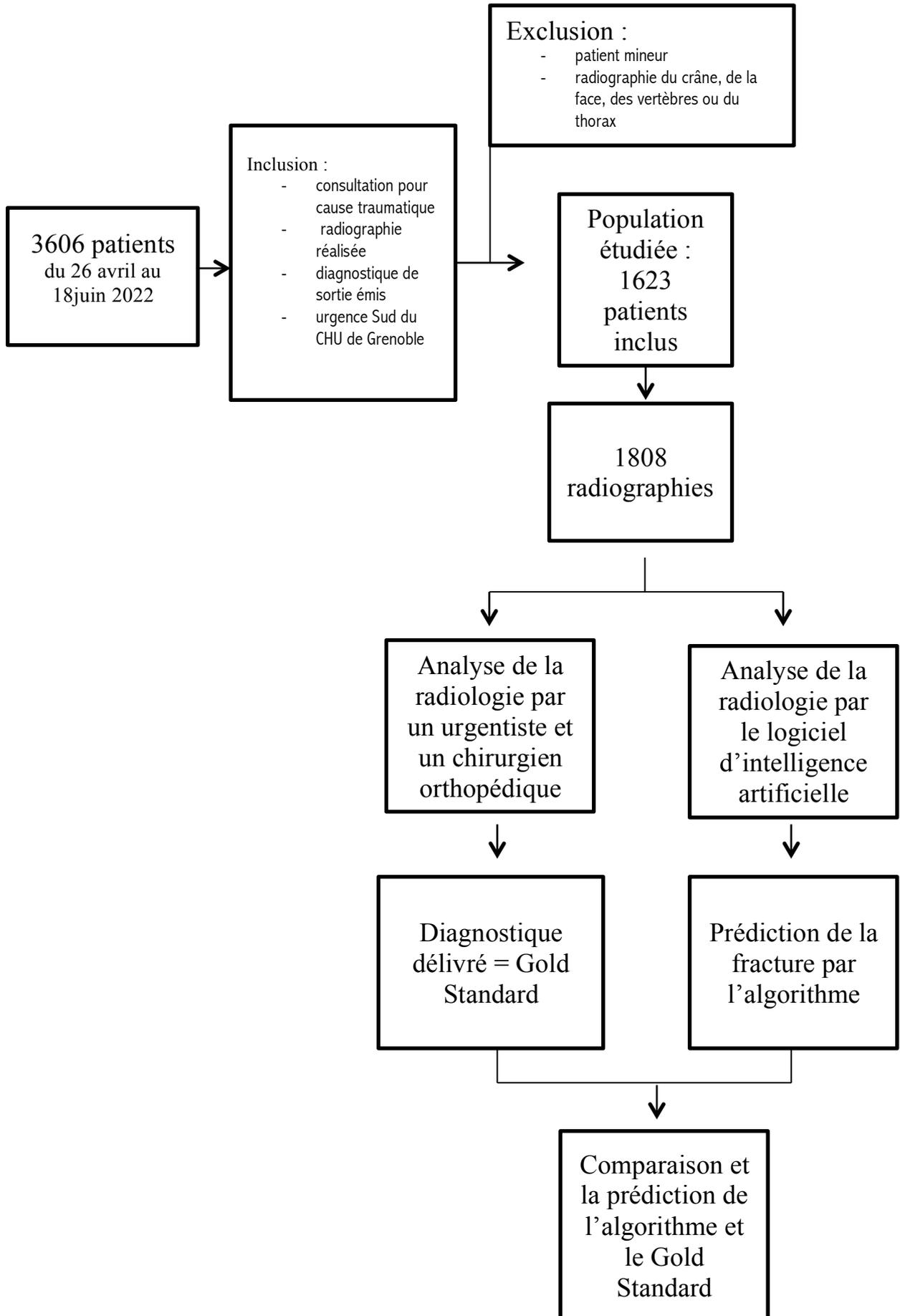
Localisation	Moyenne d'âge	Moyenne d'âge femme	Moyenne d'âge homme	Médiane d'âge	Fourchette d'âge	Nombre de femmes	Nombre d'hommes
membres supérieurs proximaux	48.5	42.6	56.7	48	18-96	137	99
membres supérieurs intermédiaires	46.6	44.0	49.7	24	18-94	74	61
membres supérieurs distaux	46.9	42.3	51.9	44	18-98	265	241
membres inférieurs proximaux	69.8	62.1	74.7	77.5	19-98	53	83
membres inférieurs intermédiaires	44	40.5	49.6	41	18-99	196	124
membres inférieurs distaux	42.6	39.2	47.0	41	18-99	269	205
Total	47.2	42.3	53.1	45	18-99	995 (55%)	813 (45%)

**Figure 7** : régions anatomiques étudiées en fonction de l'âge



1 : membres supérieurs proximaux (clavicule, épaule et bras), 2 : membres supérieurs intermédiaires (coude, avant-bras), 3 : membres supérieurs distaux (poignet, main et doigts), 4 : membres inférieurs proximaux (cadre obturateur, bassin, fémur), 5 : membres inférieurs intermédiaires (genou, jambe), 6 : membres inférieurs distaux (cheville, pied et orteils).

Figure 8 : Diagramme de flux



## *Analyse univariée*

### ❖ Analyse globale

Sur les 708 fractures identifiées par le Gold Standard, 74 fractures sont non visualisées (FN) par l'IA et sur les 1100 radiographies sans anomalies identifiées par le Gold Standard, 82 anomalies sont décelées à tort (FP) par l'IA. L'interprétation radiologique par le logiciel d'IA affiche, en référence au Gold Standard, une sensibilité de 89.5% et une spécificité de 92,5 %.

Avec un nombre de faux positifs de 82 contre 634 véritables fractures identifiées (VP), le logiciel d'IA présente une valeur prédictive positive (VPP) de 88,6 %.

Avec 74 faux négatifs, et 1018 radiographies sans fracture correctement identifiée (VN), la valeur prédictive négative (VPN) est de 93,2 %.

La concordance K est quant à elle égale de 0.81.

**Tableau 4** : Tableau de contingence et données de performances diagnostiques relatives à l'ensemble des articulations

Performances diagnostiques du logiciel	
Sensibilité	89.5 %
Spécificité	92.5 %
Valeur prédictive positive	88.5%
Valeur prédictive négative	93.2%
Kappa de Cohen	0.81

### ❖ Analyse par localisation

#### ● Membre supérieur proximal (clavicule, épaule et bras)

Avec 8 fractures non visualisées et 9 fractures décelées à tort, l'interprétation radiologique par le logiciel d'IA affiche, en référence au Gold Standard, une sensibilité de 90.5% et une spécificité de 92,8 %. Avec un nombre de faux positifs de 9 contre 76 véritables fractures identifiées, ce logiciel d'aide au diagnostic présente une valeur prédictive positive de 89,4 %. Avec 8 faux négatifs, et 143

radiographies sans fractures correctement identifiées, la valeur prédictive négative est de 94.7 %.

La concordance K est quant à elle égale à 0.84.

**Tableau 5** : Tableau de contingence et données de performances diagnostiques relatives à l'articulation de la clavicule, de l'épaule et du bras

Performances diagnostiques	
Sensibilité	90.5 %
Spécificité	94.1 %
Valeur prédictive positive	89.4 %
Valeur prédictive négative	94.7 %

Kappa de Cohen	0.84
----------------	------

**Figure 9** : Exemple d'une radiographie de clavicule droite avec comme diagnostic émis par le Gold Standard une disjonction acromio claviculaire, détectée par l'IA comme une fracture



**Figure 10** : Exemple d'une radiographie de l'épaule gauche, avec comme diagnostic émis par le Gold Standard une fracture du trochiter, détectée par le logiciel d'IA



- **Membre supérieur intermédiaire (coude, avant-bras)**

Avec 7 fractures non visualisées et 2 fractures décelées à tort, l'interprétation radiologique par le logiciel d'IA affiche, en référence au Gold Standard, une sensibilité de 90.3% et une spécificité de 96,8 %.

Avec un nombre de faux positifs de 2 contre 65 véritables fractures identifiées, ce logiciel d'aide au diagnostic présente une valeur prédictive positive de 97%. Avec 87 faux négatifs, et 61 radiographies sans fracture correctement identifiées, le logiciel d'IA présente une valeur prédictive négative de 89.7 %.

La concordance K est quant à elle égale à 0.86.

**Tableau 6** : Tableau de contingence et données de performances diagnostiques relatives à l'articulation du coude et de l'avant-bras

Performances diagnostiques	
Sensibilité	90.3 %
Spécificité	96.8 %
Valeur prédictive positive	97.0 %
Valeur prédictive négative	89.7 %

Kappa de Cohen	0.86
----------------	------

**Figure 11** : Exemple d'une radiographie de coude gauche, avec comme diagnostic émis par le Gold Standard une simple contusion, détectée par l'IA comme une fracture



**Figure 12** : Exemple d'une radiographie du coude gauche, avec comme diagnostic émis par le Gold Standard une fracture de la tête radiale, détectée par le logiciel d'IA comme une fracture



- **Membre supérieur distal (poignet, main et doigts)**

Avec 25 fractures non visualisées et 40 fractures décelées à tort, l'interprétation radiologique par le logiciel d'IA affiche, en référence au Gold Standard, une sensibilité de 89.4% et une spécificité de 85,3 %. Avec un nombre de faux positifs de 40 contre 210 véritables fractures identifiées, ce logiciel d'aide au diagnostic présente une valeur prédictive positive de 85.3 %. Avec 25 faux négatifs, et 232 radiographies sans fracture correctement identifiées, le logiciel d'IA présente une valeur prédictive négative de 90.2 %.

La concordance K est quant à elle égale à 0.74.

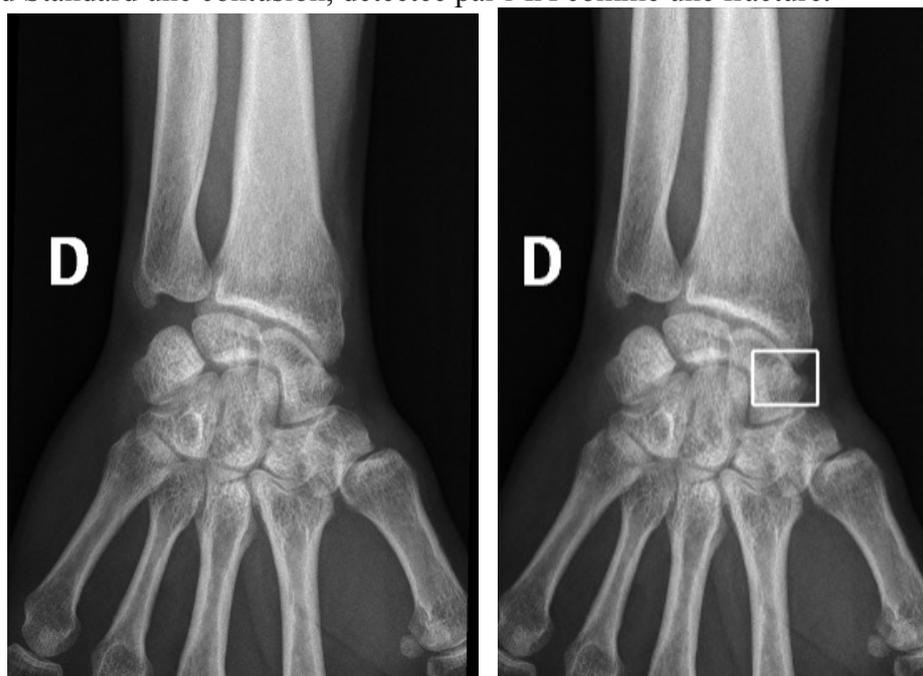
**Tableau 7** : Tableau de contingence et données de performances diagnostiques relatives à l'articulation du poignet, de la main et des doigts

Performances diagnostiques	
Sensibilité	89.4 %
Spécificité	85.3 %
Valeur prédictive positive	85.3%
Valeur prédictive négative	90.2 %

Kappa de Cohen	0.74
----------------	------

**Figure 13** : Exemple d'une radiographie de la main droite, avec comme diagnostic émis par le Gold Standard une contusion, détectée par l'IA comme une fracture.



**Figure 14** : Exemple d'une radiographie du poignet gauche, avec comme diagnostic émis par le Gold Standard une simple contusion, détectée par l'IA comme une fracture



- **Membre inférieur proximal (cadre obturateur, bassin, fémur)**

Avec 8 fractures non visualisées et 5 fractures décelées à tort, l'interprétation radiologique par le logiciel d'IA affiche, en référence au Gold Standard, une sensibilité de 90.0% et une spécificité de 91,1 %. Avec un nombre de faux positifs de 5 contre 72 véritables fractures identifiées, ce logiciel d'aide au diagnostic présente une valeur prédictive positive de 93.5 %. Avec 8 faux négatifs, et 51 radiographies sans fracture correctement identifiées, le logiciel d'IA présente une valeur prédictive négative de 86.4 %.

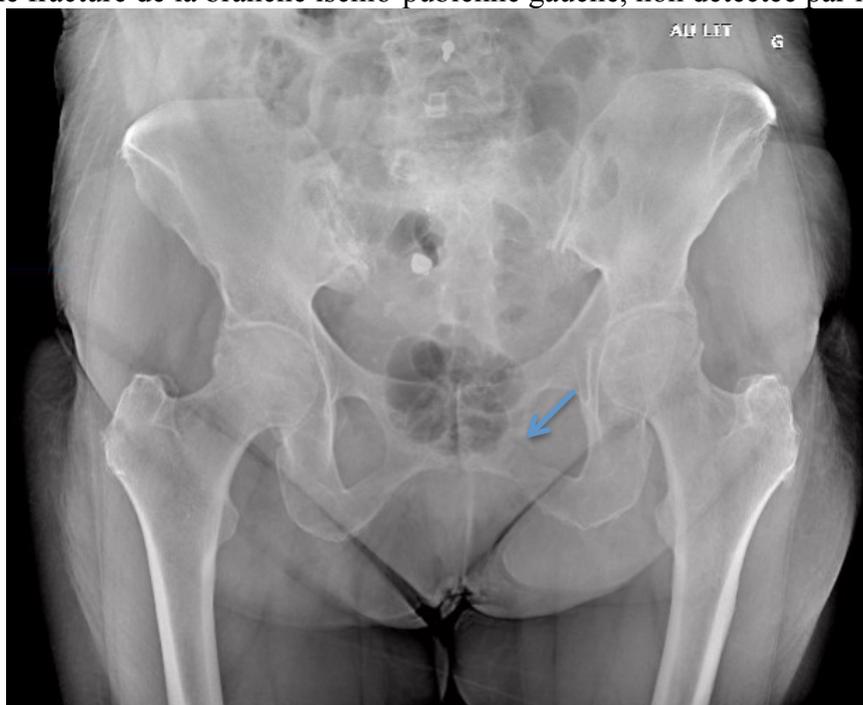
La concordance K est quant à elle égale à 0.80.

**Tableau 8** : Tableau de contingence et données de performances diagnostiques relatives à l'articulation du cadre obturateur, du bassin, et du fémur

Performances diagnostiques	
Sensibilité	90.0 %
Spécificité	91.1 %
Valeur prédictive positive	93.5 %
Valeur prédictive négative	86.4 %

Kappa de Cohen	0.80
----------------	------

**Figure 15** : Exemple d'une radiographie du bassin de face, avec comme diagnostic émis par le Gold Standard une fracture de la branche ischio-pubienne gauche, non détectée par l'IA



- **Membre inférieur intermédiaire (genou, jambe)**

Avec 7 fractures non visualisées et 6 fractures décelées à tort, l'interprétation radiologique par le logiciel d'IA affiche, en référence au Gold Standard, une sensibilité de 83.3% et une spécificité de 97,5 %. Avec un nombre de faux positifs de 6 contre 35 véritables fractures identifiées, ce logiciel d'aide au diagnostic présente une valeur prédictive positive de 85.4 %. Avec 7 faux négatifs, et 272 radiographies sans fracture correctement identifiées, le logiciel d'IA présente une valeur prédictive négative de 97.2 %.

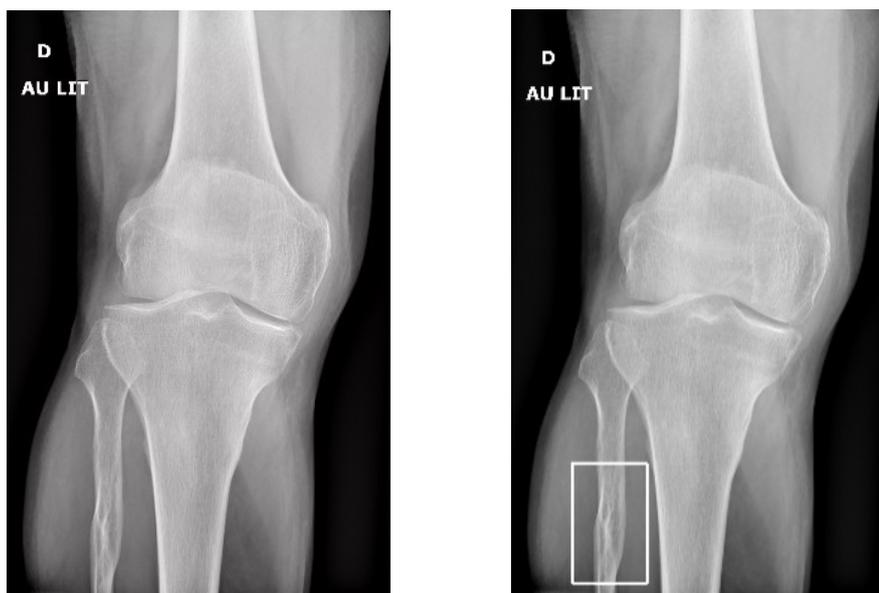
La concordance K est quant à elle égale à 0.82.

**Tableau 9** : Tableau de contingence et données de performances diagnostiques relatives à l'articulation du genou et de la jambe

Performances diagnostiques	
Sensibilité	83.3 %
Spécificité	97.5 %
Valeur prédictive positive	85.4 %
Valeur prédictive négative	97.2 %

Kappa de Cohen	0.82
----------------	------

**Figure 16** : Exemple d'une radiographie du genou droit, avec comme diagnostic émis par le Gold Standard une simple contusion, détectée à tort par l'IA comme une fracture



- **Membre inférieur distal (cheville, pied et orteils)**

Avec 19 fractures non visualisées et 20 fractures décelées à tort, l'interprétation radiologique par le logiciel d'IA affiche, en référence au Gold Standard, une sensibilité de 90.3% et une spécificité de 92,8 %. Avec un nombre de faux positifs de 19 contre 176 véritables fractures identifiées, ce logiciel d'aide au diagnostic présente une valeur prédictive positive de 89.8 %. Avec 19 faux négatifs, et 259 radiographies sans fracture correctement identifiées, le logiciel d'IA présente une valeur prédictive négative de 93.2 %.

La concordance K est quant à elle égale à 0.83.

**Tableau 10** : Tableau de contingence et données de performances diagnostiques relatives à l'articulation de la cheville, du pied et des orteils

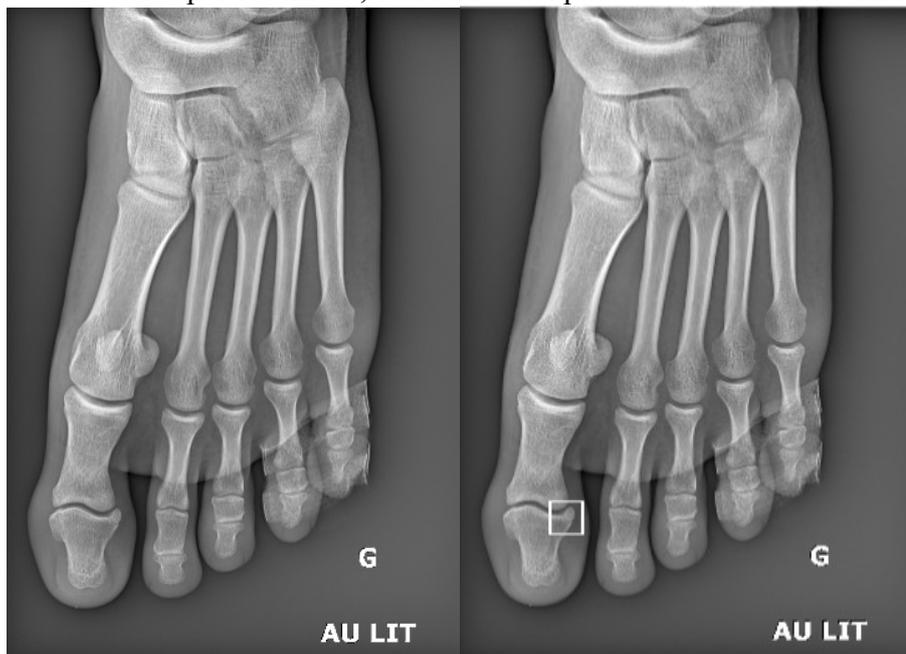
Performances diagnostiques	
Sensibilité	90.3 %
Spécificité	92.8 %
Valeur prédictive positive	89.8 %
Valeur prédictive négative	93.2 %

Kappa de Cohen	0.83
----------------	------

**Figure 17** : Exemple d'une radiographie de la cheville gauche, avec comme diagnostic émis par le Gold Standard une entorse de la cheville, détectée par l'IA comme une fracture



**Figure 18** : Exemple d'une radiographie du pied gauche, avec comme diagnostic émis par le Gold Standard une simple contusion, détectée à tort par l'IA comme une fracture



- **Comparaison entre sous-groupe**

La mesure de l'accord moyen entre le Gold Standard et l'IA est de 0,81, avec une faible disparité en fonction des différentes articulations étudiées. Si l'on se reporte au tableau de l'interprétation du coefficient kappa, il en résulte une estimation du degré d'accord qualifié comme excellent.

Bien que la disparité des résultats soit faible, la concordance diagnostique présente un meilleur score lorsqu'il s'agit de l'articulation du coude et de l'avant- avec un résultat de 0,86. Les interprétations des imageries relatives à l'articulation du poignet, de la main et des doigts sont celles présentant une concordance diagnostique la plus faible avec un score atteignant 0.74 (tableau 10).

Concernant les performances diagnostiques, en comparaison avec les performances globales de l'IA, les résultats de l'étude mettent en évidence une baisse de la sensibilité concernant l'interprétation par l'IA des radiographies du genou et de la jambe. A contrario, on constate une excellente spécificité et valeur prédictive négative (97.5% et 97.2% respectivement).

**Tableau 10** : Comparatif des interprétations radiologiques en fonction de la localisation

	Global	clavicule, épaule et bras	coude, avant-bras	poignet, main et doigts	cadre obturateur, bassin, fémur	genou, jambe	cheville, pied et orteils
Sensibilité	89.5 % (87.2-91.9)	90.5 % (83.9-97.1)	90.3 % (83.4-97.5)	87.1% (85.2-93.5)	90.0 % (83.1-96.9)	<b>83.3 %</b> <b>(71.0-95.5)</b>	90.3 % (85.9-94.6)
Spécificité	92.5 % (90.9-94.2)	94.1 % (90.2-97.9)	96.8 % (92.5-99.2)	89.4 % (81.1-89.5)	91.1 % (83.6-98.5)	97.5 % (96.1-99.5)	92.8 % (89.8-95.9)
VPP	88.5% (86.1-91.0)	89.4 % (82.9-96.0)	97.0 % (92.9-99.6)	85.3 % (79.5-88.5)	93.5 % (88.0-99.0)	<b>85.4 %</b> <b>(74.5-96.2)</b>	89.8 % (85.6-94.0)
VPN	93.2% (91.7-94.8)	94.7 % (91.0-98.4)	89.7 % (82.1-97.3)	90.2 % (86.5-94.1)	86.4 % (77.0-95.8)	97.2 % (95.6-99.3)	93.2 % (90.1-96.2)
Kappa	0.81	0.84	<b>0.86</b>	<b>0.74</b>	0.80	0.82	0.83

Les nombres entre parenthèses correspondent à l'intervalle de confiance à 95%

#### ❖ Analyse par âge

Selon l'âge, en comparaison avec les performances globales de l'IA, les résultats de l'étude mettent en évidence une baisse de la spécificité des interprétations effectuées par l'IA après 59ans.

On constate également dans le tableau ci-dessous une diminution de la valeur de la concordance diagnostique Kappa en avançant dans l'âge.

**Tableau 11** : Comparatif des interprétations radiologiques en fonction de l'âge

Age	18-38ans	39-58ans	59-78ans	79-99ans
Sensibilité	87.9% (83.6-92.2)	90.7% (86.3-95.1)	<b>89.8% (85.4-94.3)</b>	<b>90.5% (84.6-97.4)</b>
Spécificité	<b>95.9% (94.2-97.7)</b>	91.7% (88.6-94.8)	<b>87.2% (82.0-92.3)</b>	<b>85.5% (76.5-94.4)</b>
VPP	90.3% (86.2-94.5)	86.0% (80.7-91.3)	88.1% (83.3-92.8)	90.5% (84.6-96.4)
VPN	94.8% (92.9-96.8)	94.6% (92.1-97.2)	89.1% (84.3-93.9)	85.5% (76.5-94.5)
Kappa	0.84	0.81	<b>0.77</b>	<b>0.75</b>

Les nombres entre parenthèses correspondent à l'intervalle de confiance à 95%

#### ❖ Analyse par sexe

Comme le montre le tableau ci-dessous, le sexe n'a pas d'impact sur la sensibilité. La spécificité, la VPP et le coefficient Kappa sont plus élevés chez la femme. La VPN est plus élevée chez l'homme.

**Tableau 12** : Comparatif des interprétations radiologiques en fonction du sexe

Sexe	Homme	Femme
Sensibilité	89.3% (86.4-93.1)	89.8% (85.9-92.8)
Spécificité	92.1% (90.8-95.5)	93.1% (90.1-94.2)
Valeur prédictive positive	85.9% (88.3-94.2)	91.3% (82.3-89.5)
Valeur prédictive négative	94.2% (89.3-94.5)	91.9% (92.3-96.0)
Kappa	0.80	0.83

Les nombres entre parenthèses correspondent à l'intervalle de confiance à 95%

#### *Analyse multivariée*

L'utilisation d'un modèle de régression logistique permet de contrôler l'effet de caractéristiques quantitatives et qualitatives sur l'impact étudié, ici la concordance. Par exemple, les statistiques descriptives de distribution de l'âge par localisation de fracture montrent un âge moyen supérieur pour les fractures des membres inférieurs proximaux. En ne prenant pas en compte cette inégale distribution, nous sommes en incapacité de déterminer si une différence significative vient de la localisation de la fracture ou bien de la différence d'âge moyen par type de fracture. L'inclusion de variables de contrôle dans la régression logistique permet donc d'obtenir l'impact d'une variable en particulier, toute chose étant égale par ailleurs.

A âge et sexe équivalents, la concordance entre le logiciel d'IA et le Gold Standard est moins bonne lors de l'interprétation des radiographies du genou et de la jambe (membre inférieur intermédiaire) par rapport à toutes les autres régions, avec une différence significative en comparaison avec le poignet et la main ( $p < 0.01$ ) (cf. tableau 13). La concordance entre le logiciel d'IA et le Gold Standard est meilleure lors de l'interprétation des radiographies du poignet et de la main par rapport à toutes les autres régions anatomiques à âge et sexe équivalents, avec une différence significative en comparaison avec le genou et la jambe ( $p < 0.01$ ) (cf. tableau 14). Enfin, la concordance entre le logiciel d'IA et le Gold Standard est diminuée de façon significative ( $p < 0.01$ ) avec l'augmentation de l'âge des patients.

**Tableau 13 : Analyse multivariée de la bonne reconnaissance par l'IA de la lésion par localisation par rapport au membre inférieur intermédiaire**

	Odds ratio	Intervalle de confiance		P-value
Age	1.0141	1.0056	1.0227	0.001
sexe	0.7638	0.5398	1.0808	0.128
Localisation :				
1 – 5	1.7424	0.8273	3.6699	0.144
2 – 5	1.4613	0.5901	3.6184	0.412
3 – 5	3.4399	1.8594	6.3636	<b>&lt; .001</b>
4 – 5	1.8529	0.8119	4.2288	0.143
6 – 5	2.3336	1.2272	4.4372	0.010

IC : intervalle de confiance à 95%

1 : membres supérieurs proximaux (clavicule, épaule et bras), 2 : membres supérieurs intermédiaires (coude, avant-bras), 3 : membres supérieurs distaux (poignet, main et doigts), 4 : membres inférieurs proximaux (cadre obturateur, bassin, fémur), 5 : membres inférieurs intermédiaires (genou, jambe), 6 : membres inférieurs distaux (cheville, pied et orteils).

**Tableau 14 : Analyse multivariée de la bonne reconnaissance par l'IA de la lésion par localisation par rapport au membre supérieur intermédiaire**

	Odds ratio	Intervalle de confiance		P-value
Age	1.0141	1.0056	1.023	0.001
sexe	0.7638	0.5398	1.081	0.128
Localisation				
1 – 3	0.5065	0.2892	0.887	0.017
2 – 3	0.4248	0.1982	0.911	0.028
4 – 3	0.5386	0.2790	1.040	0.065
5 – 3	0.2907	0.1571	0.538	<b>&lt; .001</b>
6 – 3	0.6784	0.4473	1.029	0.068

## DISCUSSION

Au cours des 5 dernières années, de nombreux articles traitaient de l'intégration de l'intelligence artificielle dans le domaine de la radiologie et pourtant, dans la pratique quotidienne, peu de choses ont changé (21,22).

Une des principales faiblesses des systèmes d'IA actuels est le manque d'expérience lorsqu'il est utilisé dans la pratique clinique courante aux urgences (27).

Les analyses statistiques que nous avons effectuées visent à mesurer les performances diagnostiques d'un logiciel d'IA sur des radiographies réalisées aux urgences. Dans la plupart des études précédemment citées, les images sont recadrées, réduites ou présélectionnées. Dans notre étude, le système d'IA traite des images en pleine résolution avec plusieurs radiographies par patient et directement intégrées dans le PACS<sup>2</sup>. Nous avons conservé l'ensemble des examens radiographiques de mauvaise qualité.

De plus, peu d'études ont évalué les performances de l'IA dans la recherche de fractures sur de multiples articulations. La plupart des études se sont concentrées sur des articulations précises telles que les fractures de l'humérus proximal (13), du scaphoïde (14), du poignet (15), ainsi que les fractures de hanche (16,17).

Avec un Kappa de Cohen calculé à 0.81, il existe un accord presque parfait entre le diagnostic du Gold Standard et la prédiction du logiciel d'intelligence artificielle.

Concernant les paramètres diagnostiques du logiciel, on attend surtout d'un logiciel d'IA une forte sensibilité, c'est-à-dire l'identification correcte des fractures la plus élevée possible.

Dans la littérature, il est décrit qu'un médecin urgentiste non spécialiste en traumatologie présente une sensibilité minimale d'environ 80 % (28) pour la

---

<sup>2</sup> Le *Picture Archiving and Communication System* (PACS) est le système de gestion électronique des images médicales avec des fonctions d'archivage, de stockage et de communication rapide.

détection des fractures. Avec une sensibilité moyenne de 89.5%, on peut dire que notre algorithme est pertinent en situation clinique.

Les résultats de l'étude mettent en évidence une moins bonne sensibilité ainsi qu'une moins bonne concordance diagnostic lors de l'interprétation par le logiciel d'IA des radiographies du genou et de la jambe par rapport à toutes les autres régions anatomiques à âge et sexe équivalent.

Il existe une forte disparité des taux d'examens présentant une anomalie dite pathologique. Seuls 13 % des examens intéressants le genou et la jambe sont pathologiques contre 59 % concernant le bassin, les hanches et le fémur, 46 % concernant le poignet et la main, et 53 % concernant le coude et l'avant-bras.

Les fractures du genou sont donc plus rares mais plus difficilement détectables par le logiciel d'IA. L'interprétation de l'IA dans les radiographies du genou doit donc être analysée avec prudence. Par transposition, il s'agit probablement aussi de fractures plus difficilement détectables par les médecins, notamment aux urgences, et le recours à un spécialiste ou au scanner/à l'IRM devrait être facilement envisagé.

Les performances du logiciel d'IA tendent à diminuer avec l'âge : on constate une diminution de la sensibilité et de la spécificité après 59 ans, probablement en lien avec l'apparition d'arthrose.

Notre étude comporte des limites.

Le Gold Standard a été défini comme le diagnostic final délivré par un orthopédiste sénior lors du staff orthopédique journalier, lorsque toutes les radiographies réalisées aux urgences la veille sont relues. Une consultation de suivi systématique par un orthopédiste avec des radiographies de contrôle, ou encore la réalisation systématique d'un scanner ou d'une IRM permettrait la consolidation du diagnostic.

Il s'agit d'une étude monocentrique sur une période relativement courte. Selon la période hivernale ou estivale, certaines fractures sont plus ou moins fréquentes.

Une étude prospective multicentrique sur une période plus longue permettrait d'inclure un plus large panel de radiographies et de renforcer les résultats de notre étude.

Les différentes régions anatomiques ont été regroupées en sous-groupes. De ce fait, l'étude ne permet pas d'obtenir la précision du logiciel d'intelligence artificielle pour une région anatomique précise. Hors, certaines fractures sont davantage source d'erreur d'interprétation et il serait intéressant d'étudier les performances du logiciel d'IA au sein des urgences pour ce type de fractures.

Nous n'avons pas inclus de radiographies pédiatriques, hors leur interprétation peut-être parfois plus complexes que chez les adultes. Une étude Française récente a montré des résultats satisfaisants pour ce même logiciel d'IA dans la détection de fractures chez les enfants, avec une sensibilité de 95,7 %, et une spécificité de 91,2 % (29).

Néanmoins, notre étude montre des résultats concluants. L'aide d'un logiciel d'intelligence artificielle est une piste d'amélioration des performances diagnostiques aux urgences, concernant notamment les centres dans lesquels il n'existe pas de seconde relecture par un spécialiste.

Il serait intéressant d'évaluer si l'utilisation d'un logiciel d'IA permet à un interne, en fin de cursus, ou à un urgentiste sénior, d'améliorer ses performances diagnostiques concernant la détection de fractures grâce à une amélioration de sa sensibilité de détection.

En plus du manque d'expérience dans la pratique clinique quotidienne, les difficultés d'intégration de l'IA dans la pratique clinique quotidienne sont liées à d'autres problématiques.

Il est démontré qu'il existe un manque de formation et d'information sur l'IA chez les radiologues français, malgré la multiplication des publications scientifiques (24). Or, il est difficile d'accepter des technologies que l'on ne comprend pas, surtout s'il s'agit des soins de santé.

Une des préoccupations actuelles est de protéger la sécurité des patients et la confidentialité des données (30). Et si cette technologie permet une amélioration des prises en charges médicales, elle provoque aussi la méfiance des professionnels de santé ainsi que des patients et soulève des questions d'ordre éthique. Pour apprendre et fonctionner, les systèmes d'IA doivent être alimentés par une importante quantité de données. Même si la RGPD<sup>3</sup> fait office de cadre juridique pour la collecte et le traitement des données personnelles, une des craintes principales de la population s'avère être l'utilisation malveillante de leurs données personnelles (31).

La question de responsabilité est également débattue. En cas de problème, est-ce que le concepteur de l'algorithme pourrait-être remis en cause ? Même si la question reste entière, une supervision humaine est nécessaire dans toutes les décisions établies par les algorithmes d'IA.

Le coût, difficilement supportable pour certaines structures hospitalières lié à l'acquisition de ces logiciels, reste une limite à leur emploi.

Enfin, un des plus grands dangers de l'IA est que le clinicien accorde une confiance exagérée à l'IA, au détriment de son propre jugement clinique et de son raisonnement diagnostic.

Les médecins doivent pondérer ce qu'ils apprennent des données de l'IA par rapport à leur examen clinique. Seul le médecin peut émettre une réflexion entre les différents symptômes de la personne et sa radiographie, afin d'émettre un diagnostic. L'IA doit rester un outil d'aide au diagnostic permettant au médecin de s'appuyer dessus en cas de besoin, pour confirmer son diagnostic et/ou le compléter. La présence humaine sera toujours indispensable pour émettre un diagnostic.

---

<sup>3</sup> Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données).

## CONCLUSION

Le logiciel d'IA d'aide à la détection de fractures, lorsqu'il est utilisé dans la pratique clinique courante aux urgences, montre des résultats satisfaisants. Avec un Kappa de Cohen calculé à 0.81, il existe un accord presque parfait entre le diagnostic du Gold Standard et la prédiction du logiciel d'intelligence artificielle. Les résultats montrent aussi une bonne performance diagnostique, avec une sensibilité de 89.5% et une spécificité de 92,5 %. Les médecins doivent cependant être conscients de certaines limites concernant la détection des fractures du genou et de la jambe.

Une des singularités de la présente étude réside dans le fait d'étudier tous types de lésions traumatiques, sur un nombre important de patients, tout en étant proche d'une activité « en vie réelle », sans prétraitement des images, avec l'utilisation d'un logiciel d'IA directement intégré au PACS du centre hospitalier.

L'utilisation d'un logiciel d'intelligence artificielle est une piste d'amélioration des performances diagnostiques des médecins urgentistes, notamment des docteurs juniors, et qui pourrait partiellement pallier au manque de spécialistes disponibles.

Néanmoins, le médecin doit rester maître de son diagnostic, l'IA étant un outil d'aide au diagnostic sur lequel s'appuyer en cas de besoin. La présence humaine sera toujours indispensable pour émettre un diagnostic.

THESE SOUTENUE PAR : Marion Calandre

TITRE :

L'INTELLIGENCE ARTIFICIELLE AU SEIN DES URGENCES TRAUMATOLOGIQUES

CONCLUSION :

Le logiciel d'IA d'aide à la détection de fracture, lorsqu'il est utilisé dans la pratique clinique courante aux urgences, montre des résultats satisfaisants. Avec un Kappa de Cohen calculé à 0.81, il existe un accord presque parfait entre le diagnostic du Gold Standard et la prédiction du logiciel d'intelligence artificielle. Les résultats montrent aussi une bonne performance diagnostique, avec une sensibilité de 89.5% et une spécificité de 92,5 %. Les médecins doivent cependant être conscients de certaines limites concernant la détection des fractures du genou et de la jambe.

Une des singularité de la présente étude réside dans le fait d'étudier tous types de lésions traumatiques, sur un nombre importants de patient, tout en étant proche d'une activité « en vie réelle », sans prétraitement des images, avec l'utilisation d'un logiciel d'IA directement intégré au PACS du centre hospitalier.

L'utilisation d'un logiciel d'intelligence artificielle est une piste d'amélioration des performances diagnostiques des médecins urgentistes, notamment des docteurs juniors, et qui pourrait partiellement palier au manque de spécialistes disponibles.

Néanmoins, le médecin doit rester maître de son diagnostic, l'IA étant un outil d'aide au diagnostic sur lequel s'appuyer en cas de besoin. La présence humaine sera toujours indispensable pour émettre un diagnostic.

VU ET PERMIS D'IMPRIMER

Grenoble, le : 17/08/2022

LE DOYEN DE LA FACULTE DE MEDECINE

LE PRESIDENT / LA PRESIDENTE DU JURY

  
Pr. Patrice MORAND  
Pour le Président  
et par délégation  
Le Doyen de Médecine  
Pr. Patrice MORAND

  
Pr Gilbert FERRETTI

## BIBLIOGRAPHIE

1. Hussain, Faris et al. 2019. « Diagnostic Error in the Emergency Department: Learning from National Patient Safety Incident Report Analysis ». *BMC Emergency Medicine* 19(1): 77.
2. Hirschmann, Anna et al. 2019. « Artificial Intelligence in Musculoskeletal Imaging: Review of Current Literature, Challenges, and Trends ». *Seminars in Musculoskeletal Radiology* 23(03): 304-11.
3. Kachalia, Allen et al. 2007. « Missed and Delayed Diagnoses in the Emergency Department: A Study of Closed Malpractice Claims From 4 Liability Insurers ». *Annals of Emergency Medicine* 49(2): 196-205.
4. Jalal, Sabeena, William Parker, Duncan Ferguson, et Savvas Nicolaou. 2021. « Exploring the Role of Artificial Intelligence in an Emergency and Trauma Radiology Department ». *Canadian Association of Radiologists Journal* 72(1): 167-74.
5. Ramesh, An, C Kambhampati, Jrt Monson, et Pj Drew. 2004. « Artificial Intelligence in Medicine ». *Annals of The Royal College of Surgeons of England* 86(5): 334-38.
6. Basha, C. M. A. K. Zeelan, Maruthi Padmaja, et G. N. Balaji. 2018. « Computer Aided Fracture Detection System ». *Journal of Medical Imaging and Health Informatics* 8(3): 526-31.
7. Kalmet, Pishtiwan H S et al. 2020. « Deep Learning in Fracture Detection: A Narrative Review ». *Acta Orthopaedica* 91(2): 215-20.
8. Nakaura, Takeshi et al. 2020. « A Primer for Understanding Radiology Articles about Machine Learning and Deep Learning ». *Diagnostic and Interventional Imaging* 101(12): 765-70.
9. Esteva, Andre et al. 2017. « Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks ». *Nature* 542(7639): 115-18.
10. Kooi, Thijs et al. 2017. « Large Scale Deep Learning for Computer Aided Detection of Mammographic Lesions ». *Medical Image Analysis* 35: 303-12.
11. Gulshan, Varun et al. 2016. « Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs ». *JAMA* 316(22): 2402.

12. Olczak, Jakub et al. 2017. « Artificial Intelligence for Analyzing Orthopedic Trauma Radiographs: Deep Learning Algorithms—Are They on Par with Humans for Diagnosing Fractures? » *Acta Orthopaedica* 88(6): 581-86.
13. Chung, Seok Won et al. 2018. « Automated Detection and Classification of the Proximal Humerus Fracture by Using Deep Learning Algorithm ». *Acta Orthopaedica* 89(4): 468-73.
14. Yoon, Alfred P. et al. 2021. « Development and Validation of a Deep Learning Model Using Convolutional Neural Networks to Identify Scaphoid Fractures in Radiographs ». *JAMA Network Open* 4(5): e216096.
15. Kim, D.H., et T. MacKinnon. 2018. « Artificial Intelligence in Fracture Detection: Transfer Learning from Deep Convolutional Neural Networks ». *Clinical Radiology* 73(5): 439-45.
16. Langerhuizen, David W. G. et al. 2019. « What Are the Applications and Limitations of Artificial Intelligence for Fracture Detection and Classification in Orthopaedic Trauma Imaging? A Systematic Review ». *Clinical Orthopaedics & Related Research* 477(11): 2482-91.
17. Urakawa, Takaaki et al. 2019. « Detecting Intertrochanteric Hip Fractures with Orthopedist-Level Accuracy Using a Deep Convolutional Neural Network ». *Skeletal Radiology* 48(2): 239-44.
18. Garwood, Elisabeth R., Ryan Tai, Ganesh Joshi, et George J. Watts V. 2020. « The Use of Artificial Intelligence in the Evaluation of Knee Pathology ». *Seminars in Musculoskeletal Radiology* 24(01): 021-029.
19. Whang JS, Baker SR, Patel R, Luk L, Castro A. « The causes of medical malpractice suits against radiologists in the United States ». *Radiology*. févr 2013;266(2): 548-54.
20. Hallas, Peter, et Trond Ellingsen. 2006. « Errors in Fracture Diagnoses in the Emergency Department – Characteristics of Patients and Diurnal Variation ». *BMC Emergency Medicine* 6(1): 4.
21. Stewart, Jonathon, Peter Sprivulis, et Girish Dwivedi. 2018. « Artificial Intelligence and Machine Learning in Emergency Medicine ». *Emergency Medicine Australasia* 30(6): 870-74.

22. Chong, Jaron, et Michael N. Patlas. 2021. « Radiology Artificial Intelligence: Bringing Theory to Clinical Practice ». *Canadian Association of Radiologists Journal* 72(1): 6-6.
23. Jalal, Sabeena, William Parker, Duncan Ferguson, et Savvas Nicolaou. 2021. « Exploring the Role of Artificial Intelligence in an Emergency and Trauma Radiology Department ». *Canadian Association of Radiologists Journal* 72(1): 167-74.
24. Waymel, Q. et al. 2019. « Impact of the Rise of Artificial Intelligence in Radiology: What Do Radiologists Think? » *Diagnostic and Interventional Imaging* 100(6): 327-36.
25. Zhang-Yin JT. « Évaluation de la concordance dans les études d'imagerie diagnostique : une étude de la qualité des données publiées ». 18 juin 2015;69.
26. Landis JR, Koch GG. « The Measurement of Observer Agreement for Categorical Data. *Biometrics*. » 1977;33(1):159-74.
27. Pesapane, Filippo et al. 2020. « Myths and Facts about Artificial Intelligence: Why Machine- and Deep-Learning Will Not Replace Interventional Radiologists ». *Medical Oncology* 37(5): 40.
28. Lindsey, Robert et al. 2018. « Deep Neural Network Improves Fracture Detection by Clinicians ». *Proceedings of the National Academy of Sciences* 115(45): 11591-96.
29. Dupuis, Michel, Léo Delbos, Raphael Veil, et Catherine Adamsbaum. 2022. « External Validation of a Commercially Available Deep Learning Algorithm for Fracture Detection in Children ». *Diagnostic and Interventional Imaging* 103(3): 151-59.
30. Liew, Charlene. 2018. « The Future of Radiology Augmented with Artificial Intelligence: A Strategy for Success ». *European Journal of Radiology* 102: 152-56.
31. Omoumi, Patrick et al. 2021. « To Buy or Not to Buy—Evaluating Commercial AI Solutions in Radiology (the ECLAIR Guidelines) ». *European Radiology* 31(6): 3786-96.

# **L**e serment d'Hippocrate

Texte revu par l'Ordre des médecins en 2012

**“ Au moment d'être admis(e) à exercer la médecine, je promets et je jure d'être fidèle aux lois de l'honneur et de la probité.**

Mon premier souci sera de rétablir, de préserver ou de promouvoir la santé dans tous ses éléments, physiques et mentaux, individuels et sociaux.

Je respecterai toutes les personnes, leur autonomie et leur volonté, sans aucune discrimination selon leur état ou leurs convictions. J'interviendrai pour les protéger si elles sont affaiblies, vulnérables ou menacées dans leur intégrité ou leur dignité. Même sous la contrainte, je ne ferai pas usage de mes connaissances contre les lois de l'humanité.

J'informerai les patients des décisions envisagées, de leurs raisons et de leurs conséquences.

Je ne tromperai jamais leur confiance et n'exploiterai pas le pouvoir hérité des circonstances pour forcer les consciences.

Je donnerai mes soins à l'indigent et à quiconque me les demandera. Je ne me laisserai pas influencer par la soif du gain ou la recherche de la gloire.

Admis(e) dans l'intimité des personnes, je tairai les secrets qui me seront confiés. Reçu(e) à l'intérieur des maisons, je respecterai les secrets des foyers et ma conduite ne servira pas à corrompre les mœurs.

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

Je préserverai l'indépendance nécessaire à l'accomplissement de ma mission. Je n'entreprendrai rien qui dépasse mes compétences. Je les entretiendrai et les perfectionnerai pour assurer au mieux les services qui me seront demandés.

J'apporterai mon aide à mes confrères ainsi qu'à leurs familles dans l'adversité. Que les hommes et mes confrères m'accordent leur estime si je suis fidèle à mes promesses ; que je sois déshonoré(e) et méprisé(e) si j'y manque. ”