



HAL
open science

Performance diagnostique d'un modèle de deep learning pour la prédiction de cancer du sein en mammographie : expérience du système Transpara® sur 89 lésions frontières (B3) au CHU Grenoble Alpes

Kévin Guttin

► **To cite this version:**

Kévin Guttin. Performance diagnostique d'un modèle de deep learning pour la prédiction de cancer du sein en mammographie : expérience du système Transpara® sur 89 lésions frontières (B3) au CHU Grenoble Alpes. Médecine humaine et pathologie. 2022. dumas-04278425

HAL Id: dumas-04278425

<https://dumas.ccsd.cnrs.fr/dumas-04278425v1>

Submitted on 10 Nov 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

Année : 2022

**« PERFORMANCE DIAGNOSTIQUE D'UN MODÈLE DE DEEP LEARNING POUR
LA PREDICTION DE CANCER DU SEIN EN MAMMOGRAPHIE : EXPÉRIENCE DU
SYSTÈME TRANSPARA® SUR 89 LÉSIONS FRONTIÈRES (B3)
AU CHU GRENOBLE ALPES »**

THÈSE

POUR LE DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE

SPÉCIALITÉ : RADIOLOGIE ET IMAGERIE MÉDICALE

Par M. Kévin GUTTIN

[Données à caractère personnel]

THÈSE SOUTENUE PUBLIQUEMENT À LA FACULTÉ DE MÉDECINE DE GRENOBLE

Le 03/10/22

DEVANT LE JURY COMPOSÉ DE :

Président du jury :

M. le Pr Gilbert FERRETTI

Membres :

M. le Pr Ivan BRICAULT

Mme le Dr Bénédicte GUILLAUME

Mme le Dr Anne-Cécile PHILIPPE

Mme le Dr Aurélie DELOUCHE (directrice de thèse)

L'UFR de Médecine de Grenoble n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions sont considérées comme propres à leurs auteurs.

CORPS	NOM Prénom	DISCIPLINE UNIVERSITAIRE
PU-PH	ALBALADEJO Pierre	Anesthésiologie-réanimation et médecine péri-opératoire
PU-PH	ARVIEUX-BARTHÉLÉMY Catherine	Chirurgie viscérale et digestive
PU-PH	BAILLET Athan	Rhumatologie
PU-PH	BARONE-ROCHETTE Gilles	Cardiologie
PU-PH	BAYAT Sam	Physiologie
MCF Ass.MG	BENDAMENE Farouk	Médecine Générale
PU-PH	BENHAMOU Pierre-Yves	Endocrinologie, diabète et maladies métaboliques
PU-PH	BERGER François	Biologie cellulaire
MCU-PH	BÉTRY Cécile	Nutrition
MCU-PH	BIDART-COUTTON Marie	Biologie cellulaire
PU-PH	BIOULAC-ROGIER Stéphanie	Pédopsychiatrie ; addictologie
PU-PH	BLAISE Sophie	Chirurgie vasculaire ; médecine vasculaire
PR Ass. Méd.	BOILLOT Bernard	Urologie
MCU-PH	BOISSET Sandrine	Bactériologie-virologie ; Hygiène hospitalière
PU-PH	BONAZ Bruno	Gastroentérologie ; hépatologie
PU-PH	BONNETERRE 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	BOUCHAUD Jacques	Médecine Générale
PU-PH	BOUGEROL Thierry	Psychiatrie d'adultes
PU-PH	BOUILLET Laurence	Médecine interne ; gériatrie et biologie du vieillissement ; addictologie
MCU-PH	BOUSSAT Bastien	Épidémiologie, économie de la santé et prévention
PU-PH	BOUZAT 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 émérite	CESBRON Jean-Yves	Immunologie
PU-PH	CHABARDÈS Stephan	Neurochirurgie
PU-PH	CHABRE Olivier	Endocrinologie, diabète et maladies métaboliques
PU-PH	CHAFFANJON Philippe	Anatomie
MCF Ass.MG	CHAMBOREDON Benoît	Médecine Générale

CORPS	NOM Prénom	DISCIPLINE UNIVERSITAIRE
PU-PH	CHARLES Julie	Dermato-vénérologie
MCF Ass.MG	CHAUVET Marion	Médecine Générale
PU-PH	CHAVANON Olivier	Chirurgie thoracique et cardiovasculaire
MCU-PH	CHEVALLIER Marie	Pédiatrie
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	CLIN CHERPEC Rita	Nutrition
PU-PH	COHEN Olivier	Histologie, embryologie et cytogénétique
PU-PH	COURVOISIER Aurélien	Chirurgie infantile
PU-PH	COUTTON 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.	DEFAYE Pascal	Cardiologie
PU-PH	DEGANO Bruno	Pneumologie ; addictologie
PU-PH	DEMATTEIS Maurice	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
PU-PH émérite	DEMONGEOT Jacques	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	DERANSART Colin	Physiologie
PU-PH	DESCOTES Jean-Luc	Urologie
PU-PH	DETANTE Olivier	Neurologie
MCU-PH	DIETERICH Klaus	Génétique
MCU-PH	DOUTRELEAU Stéphane	Physiologie
PU-PH	DUMESTRE PÉRARD Chantal	Immunologie
PU-PH	ÉPAULARD Olivier	Maladies infectieuses ; Maladies tropicales
MCU-PH	EVAIN Jean-Noël	Anesthésiologie-réanimation et médecine périopératoire
MCU-PH	EYSSERIC Hélène	Médecine légale et droit de la santé
PU-PH émérite	FAGRET Daniel	Biophysique et médecine nucléaire
PU-PH	FAUCHERON Jean-Luc	Chirurgie viscérale et digestive
PU-PH	FAURE Julien	Biochimie et biologie moléculaire
PU-PH	FERRETTI Gilbert	Radiologie et imagerie médicale
PU-PH	FIARD Gaëlle	Urologie
PU-PH	FONTAINE Éric	Nutrition
PU-PH émérite	FRANÇOIS Patrice	Épidémiologie, économie de la santé et prévention
PR Ass. Méd.	FREY Gil	Chirurgie thoracique et cardiovasculaire
MCU-MG	GABOREAU Yoann	Médecine Générale
PU-PH	GARBAN Frédéric	Hématologie ; Transfusion
PU-PH	GAUDIN Philippe	Rhumatologie
MCU-PH	GAUTIER-VEYRET Elodie	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
PU-PH	GAVAZZI Gaétan	Médecine interne ; gériatrie et biologie du vieillissement ; addictologie

CORPS	NOM Prénom	DISCIPLINE UNIVERSITAIRE
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 ; Brûlogie
MCU-PH	GRAND Sylvie	Radiologie et imagerie médicale
PU-PH émérite	GRIFFET Jacques	Chirurgie infantile
PU-PH	HAINAUT Pierre	Biochimie et biologie moléculaire
PU-PH	HENNEBICQ Sylviane	Biologie et médecine du développement et de la reproduction ; gynécologie médicale
PU-PH	HOFFMANN Pascale	Gynécologie-obstétrique ; gynécologie médicale
PU-PH émérite	HOMMEL Marc	Neurologie
PU-MG	IMBERT Patrick	Médecine Générale
PU-PH émérite	JOUK Pierre-Simon	Génétique
PU-PH	KAHANE Philippe	Physiologie
PU-PH	KASTLER Adrian	Radiologie et imagerie médicale
MCU-PH	KHERRAF Zine-Eddine	Biologie et médecine du développement et de la reproduction ; gynécologie médicale
PU-PH	KRAINIK Alexandre	Radiologie et imagerie médicale
PU-PH	LABARÈRE José	Épidémiologie, économie de la santé et prévention
PU-PH	LABLANCHE (CORNALI) Sandrine	Endocrinologie, diabète et maladies métaboliques
PU-PH	LANDELLE Caroline	Bactériologie – virologie ; Hygiène hospitalière
PU-PH	LANTUEJOL Sylvie	Anatomie et cytologie pathologiques
PR Ass. Méd.	LARAMAS Mathieu	Cancérologie ; radiothérapie
MCU-PH	LARDY Bernard	Biochimie et biologie moléculaire
PU-PH	LAURENT-COSTENTIN Charlotte	Gastroentérologie ; hépatologie ; addictologie
MCU-PH	LE GOUËLLEC LE PISSART Audrey	Biochimie et biologie moléculaire
MCU-PH	LE MARÉCHAL Marion	Maladies infectieuses ; maladies tropicales
PU-PH	LECCIA Marie-Thérèse	Dermato-vénérologie
PR Ass.MG	LEDOUX Jean-Nicolas	Médecine Générale
PU-PH émérite	LÉTOUBLON Christian	Chirurgie viscérale et digestive
PU-PH émérite	LÉVY Patrick	Physiologie
PU-PH	LONG Jean-Alexandre	Urologie
MCU-PH	LUPO Julien	Bactériologie-virologie ; Hygiène hospitalière
PU-PH	MAÎTRE Anne	Médecine et santé au travail
MCU-PH	MARLU Raphaël	Hématologie ; Transfusion
PR Ass. Méd.	MATHIEU Nicolas	Gastroentérologie ; hépatologie ; addictologie
MCU-PH	MAUBON Danièle	Parasitologie et mycologie
PU-PH	MAURIN Max	Bactériologie-virologie ; Hygiène hospitalière
MCU-PH	MC LEER Anne	Histologie, embryologie et cytogénétique
MCU-PH	MEONI Sara	Neurologie
PR Ass. Méd.	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

CORPS	NOM Prénom	DISCIPLINE UNIVERSITAIRE
PU-PH	MORO-SIBILOT Denis	Pneumologie ; addictologie
MCU-PH	MORTAMET Guillaume	Pédiatrie
PU-PH émérite	MOUSSEAU Mireille	Cancérologie ; radiothérapie
PU-PH émérite	MOUTET François	Chirurgie plastique, reconstructrice et esthétique ; brûlologie
PR Ass.MG	ODDOU Christel	Médecine Générale
PR Ass. Méd.	ORMEZZANO Olivier	Cardiologie
MCU-PH	PACLET Marie-Hélène	Biochimie et biologie moléculaire
PU-PH	PAILHÉ Régis	Chirurgie orthopédique et traumatologie
PU-PH	PALOMBI Olivier	Anatomie
PU-PH	PARK Sophie	Hématologie ; Transfusion
PR Ass.MG	PAUMIER-DESBRIÈRES Françoise	Médecine Générale
PU-PH	PAYEN DE LA GARANDERIE Jean-François	Anesthésiologie-réanimation 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	PÉPIN Jean-Louis	Physiologie
PU-PH	PÉRENNOU Dominique	Médecine physique et de réadaptation
PU-PH	PERNOD Gilles	Chirurgie vasculaire ; Médecine vasculaire
MCF	PINSAULT Nicolas	Sciences de la rééducation et de réadaptation
PU-PH	PIOLAT Christian	Chirurgie infantile
PU-PH émérite	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	RENDU 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 VÉRICOURT Guillaume	Médecine Générale
PU-PH émérite	SARAGAGLIA Dominique	Chirurgie orthopédique et traumatologie
MCU-PH	SATRE Véronique	Génétique
PU-PH	SAUDOU Frédéric	Biologie cellulaire
PU-PH	SCHMERBER Sébastien	Oto-rhino-laryngologie
PU-PH	SCHWEBEL Carole	Médecine intensive-réanimation
PU-PH	SCOLAN Virginie	Médecine légale et droit de la santé

CORPS	NOM Prénom	DISCIPLINE UNIVERSITAIRE
PU-PH	SEIGNEURIN Arnaud	Épidémiologie, économie de la santé et prévention
PU-PH	SPEAR Rafaëlle	Chirurgie vasculaire ; médecine vasculaire
PU-PH émérite	STAHL Jean-Paul	Maladies infectieuses ; Maladies tropicales
PU-PH	STANKE 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	THEVENON Julien	Génétique
PU-PH	TOFFART 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	VANZETTO Gérard	Cardiologie
PU-PH	VIGLINO Damien	Médecine d'urgence
PU-PH	VUILLEZ Jean-Philippe	Biophysique et médecine nucléaire
PU-PH	WEIL Georges	Épidémiologie, économie de la santé et prévention
PU-PH	ZAOUI Philippe	Néphrologie
PU-PH émérite	ZARSKI Jean-Pierre	Gastroentérologie ; hépatologie ; addictologie

PU-PH	Professeur des universités - Praticien hospitalier
MCU-PH	Maître de conférences des universités - Praticien hospitalier
PU-PH émérite*	Professeur des universités - Praticien hospitalier émérite
MCU-PH émérite*	Maître de conférences des universités - Praticien hospitalier émérite
PU-MG	Professeur des universités de médecine générale
MCU-MG	Maître de conférences des universités de médecine générale
MCF	Maître de conférences des universités
PR Ass. Med.	Professeur des universités associé (à mi-temps)
PR Ass. MG	Professeur des universités de médecine générale associé (à mi-temps)
MCF Ass.MG	Maître de conférences des universités de médecine générale associé (à mi-temps)

Remerciements

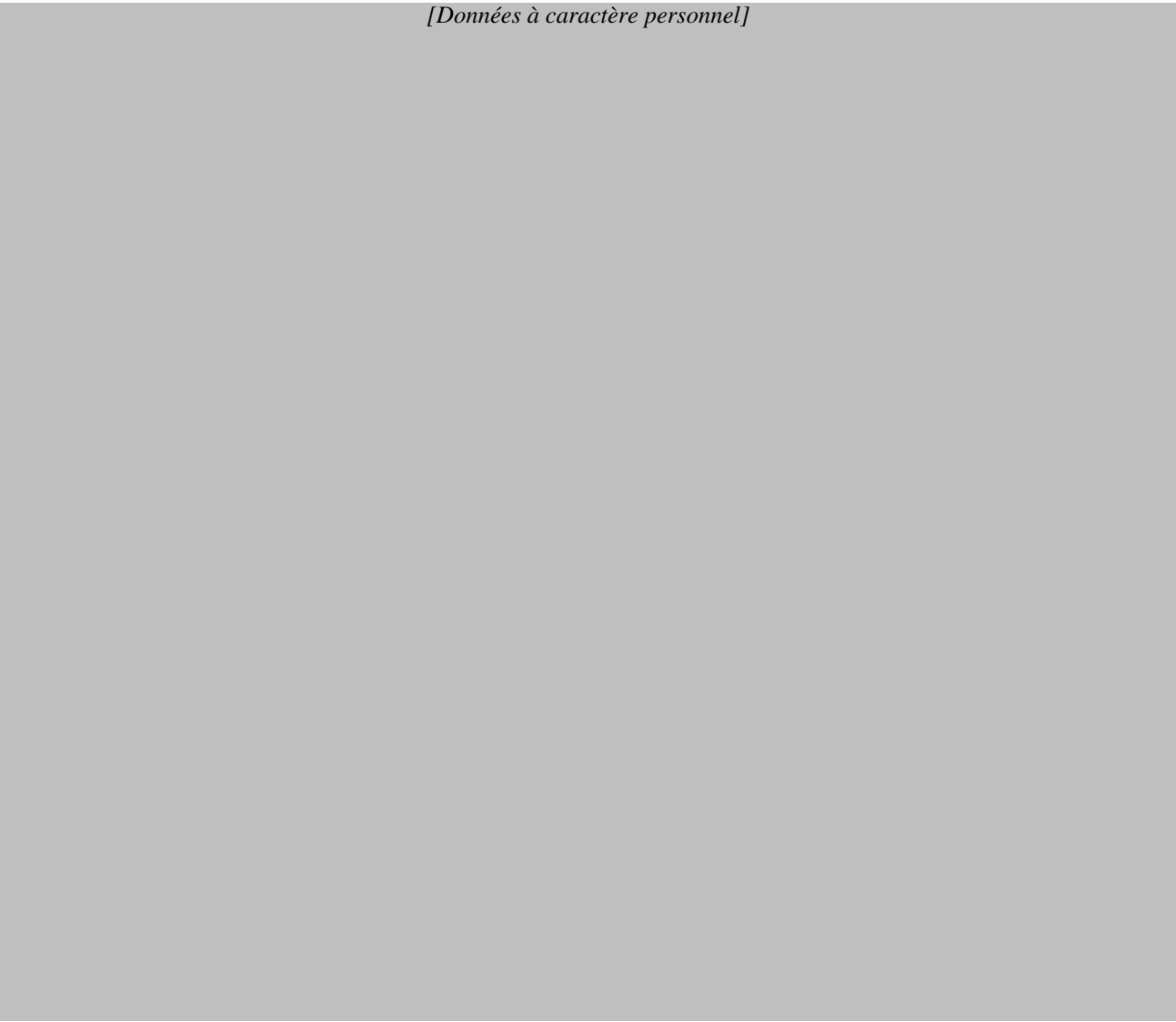
[Données à caractère personnel]

[Données à caractère personnel]

[Données à caractère personnel]

[Données à caractère personnel]

[Données à caractère personnel]



« PERFORMANCE DIAGNOSTIQUE D'UN MODÈLE DE DEEP LEARNING POUR LA PRÉDICTION DE CANCER DU SEIN EN MAMMOGRAPHIE : EXPÉRIENCE DU SYSTÈME TRANSPARA® SUR 89 LÉSIONS FRONTIÈRES (B3) AU CHU GRENOBLE ALPES »

RÉSUMÉ

Introduction : L'amélioration des techniques d'imageries pour le dépistage du cancer du sein est responsable depuis plusieurs années d'une augmentation du nombre de biopsies mammaires et de la découverte de lésions frontières (B3). Elles représentent environ 7 à 9% des biopsies et sont un groupe hétérogène d'anomalies histologiques marqueurs d'un tissu mammaire à risque de cancer du sein. Leur prise en charge a longtemps reposé sur la chirurgie devant leur risque de sous-estimation néoplasique de 9,9% à 35,1%. Si certains critères cliniques, radiologiques ou histologiques sont connus pour majorer le risque de sous-estimation de ces lésions frontières, il n'existe pas à l'heure actuelle de recommandation officielle sur leur prise en charge qui reste pays et centre dépendant. Notre objectif est d'étudier les performances d'un modèle d'intelligence artificielle (Transpara®) pour la prédiction de malignité des pièces opératoires d'anomalies mammographiques macrobiopsiées et correspondant à des lésions frontières B3.

Méthodes : Cette étude rétrospective monocentrique a inclus 89 lésions frontières macrobiopsiées correspondant à des anomalies mammographiques chez 86 patientes, de mars 2016 à février 2022 au CHUGA. L'ensemble des patientes a été opéré avec des résultats histologiques définitifs : 11 cancers, 15 lésions bénignes et 63 lésions frontières. Le modèle Transpara® a été appliqué à posteriori (79 lésions) et à priori (10 lésions) sur les mammographies 2D. Pour chaque anomalie mammographique, le modèle attribuait un score de risque de cancer de 0 à 100 (plus le score est élevé, plus le risque est important). Les performances de l'algorithme ont été analysées sur les lésions B2 définitives (absence de prise en charge) et les lésions B5 définitives (chirurgie complémentaire). Elles ont été illustrées sous la forme d'une courbe ROC et d'une aire sous la courbe.

Résultats : Les performances du modèle avaient une aire sous la courbe à 0,521 avec une sensibilité à 0,36 et une spécificité à 0,87. Les scores de l'IA étaient croissants dans le sens du BIRADS ($p=0,004$) alors qu'il n'existait pas de différence statistique selon la morphologie ou la distribution des microcalcifications ($p=0,052$). Les cancers infiltrants étaient soit non vus par le modèle (1 carcinome infiltrant de type non spécifique), soit vus avec un score intermédiaire (1 carcinome lobulaire infiltrant).

Conclusion : Le modèle Transpara® semble non performant dans notre population constituée uniquement de lésions frontières dont certaines patientes sont en dehors du cadre du dépistage. Ces résultats sont en marge des études actuelles qui se basent uniquement sur des mammographies de dépistage.

MOTS CLÉS : Cancer du sein, lésions frontières, B3, Intelligence artificielle, Transpara

FILIÈRE : Médecine, radiodiagnostic et imagerie médicale

Liste des acronymes

ABUS : *Automated breast ultrasound screening* ou examen échographique mammaire automatisé

ACR : American College of Radiology

BI-RADS : Breast Imaging-Reporting And Data System

CAD : *Computer aided diagnoses* ou système d'aide au diagnostique

CC : Cranio-caudale

CHUGA : Centre hospitalier universitaire Grenoble Alpes

CLIS : Carcinome lobulaire in situ

CR : Cicatrice radiaire

DL : *Deep Learning* ou apprentissage profond

DRCI : Direction de la Recherche Clinique et de l'Innovation

DBT : *Digital breast tomosynthesis* ou tomosynthèse

EUREF : European Reference Organisation for Quality Assured Breast Screening and Diagnostic Services

EWGBSP : European Working Group on Breast Screening Pathology

GPS : General Problem Solver

GPU : Graphic Processing Units

HCA : Hyperplasie canalaire atypique

HLA : Hyperplasie lobulaire atypique

IA : Intelligence artificielle

IBM : International Business Machines Corporation

INLSVRC : Imagenet Large Scale Visual Recognition Challenge

LM : Latéro-médiale

MCA : Métaplasie cylindro-cubique atypique

ML : Médio-latérale

ML : *Machine Learning* ou apprentissage automatique

MLO : Médio-latérale oblique

My PEBS : My Personal Breast Screening

PA : Papillome atypique

PS : Papillome simple

SM : *Synthetic mammography* ou Mammographie synthétique

TP : Tumeur phyllode

VEGF : Vascular Endothelial Growth Factor ou facteur de croissance de l'endothélium vasculaire

WISDOM : Women Informed to Screen Depending on Measures of Risk

TABLE DES MATIÈRES

Partie 1 : Introduction

A) Lésions frontières.....	16
1. Définition	16
2. Diagnostic	17
3. Problématique.....	21
B) Intelligence Artificielle	22
1. Définition	22
a. Les prémices.....	23
b. Machine Learning.....	24
c. Deep Learning	25
2. IA en sénologie en France et ailleurs.....	27
a. Acquisition de l'image	27
b. Détection automatique de la densité mammaire	28
c. CAD ou système d'aide à la décision	29
d. Analyse radiomique tumorale.....	33
e. Dépistage personnalisé.....	34
3. IA en sénologie au CHUGA : Transpara®	35
a. Marques fixes.....	35
b. Score régional	36
c. Score d'examen	36
d. Marques interactives	39
e. Aide à la navigation en tomosynthèse.....	39
f. Résumé des études cliniques	39

Partie 2 : Article

A) Introduction	40
B) Matériels et méthodes	42
1. Design de l'étude	43
2. Algorithme Transpara®	43
3. Analyses statistiques	44
C) Résultats	45
1. Analyse descriptive de la population	45
a. Analyse de la population B3 biopsiée	45
b. Analyse des pièces opératoires	46
2. Analyse du score Transpara®	48
a. Analyse descriptive du score régional	48
b. Foyers additionnels	54
c. Performance du modèle entre les groupes B2 et B5	58
d. Analyse descriptive de la distribution du score du modèle des 81 foyers de calcifications de notre population B3 biopsiée	59
e. Analyse descriptive des masses dans le groupe B5	60
D) Discussion	61
1. Description de la population B3	61
2. Analyse du score Transpara®	61
3. Limites de l'étude	63
E) Conclusion	65

PARTIE 1 : Introduction

A) Lésions Frontières

1. Définitions

Le cancer du sein est le premier cancer par incidence de la femme en France (58 458 en 2018) et la première cause de décès par cancer chez la femme [1].

Depuis 2004, la France a mis en place un dépistage organisé tous les deux ans chez les femmes de 50 à 74 ans sans risque particulier et sans anomalie clinique [2].

Il consiste en un examen clinique, la réalisation de mammographies (clichés obliques externes et faces bilatéraux) et d'une échographie complémentaire pour les seins denses.

L'amélioration des techniques d'imagerie pour le dépistage du cancer du sein est responsable depuis plusieurs années d'une augmentation du nombre de biopsies mammaires. Pour chaque prélèvement, dans le but d'avoir un langage commun la lésion biopsiée est analysée et classée d'un point de vue histologique selon la classification européenne du European Working Group on Breast Screening Pathology (EWGBSP) [3].

B1 : prélèvement intéressant du tissu normal ou non représentatif des aspects décrits en imagerie, à renouveler.

B2 : lésion bénigne, résultats concordants avec l'imagerie.

B3 : lésion bénigne mais à potentiel évolutif incertain ou fréquemment associée à des lésions plus péjoratives. Peut nécessiter une exérèse percutanée sous échographie, une reprise chirurgicale ou une surveillance rapprochée.

B4 : prélèvement partiel d'une lésion suspecte nécessitant des biopsies complémentaires.

B5 : lésion maligne nécessitant une prise en charge adaptée.

Les lésions classées B3, dites « frontières », représentent environ 7 à 9% [4] des biopsies. Elles constituent un groupe hétérogène de lésions mammaires épithéliales

et fibro-épithéliales avec ou sans atypies nucléaires, associées à un risque de malignité très différent après résection (2 à 40%) [5-7].

Elles font l'objet d'un grand intérêt depuis plusieurs années car elles représentent un sur-risque de néoplasie mammaire pour différentes raisons :

- Risque de sous-estimation (risque de lésion néoplasique sur la pièce opératoire non retrouvée sur la biopsie) dû au défaut d'échantillonnage de la lésion en raison de son hétérogénéité histologique allant de 0 à 30%.
- Marqueur d'un tissu glandulaire à risque (risque élevé de développer un cancer du sein dans le futur).
- Risque de dégénérescence (lésions précurseurs).

Cette catégorie possède deux sous-groupes :

- **B3a**, lésions épithéliales **sans atypies**, ayant un risque de sous-estimations locales.
- **B3b**, lésions épithéliales **avec atypies**, marqueurs de risque et précurseurs non obligatoires de malignité.

Les principales lésions classées B3 sont, de la plus péjorative à la moins selon la littérature [7], le papillome atypique (PA), l'hyperplasie canalaire atypique (HCA), le carcinome lobulaire in situ (CLIS), l'hyperplasie lobulaire atypique (HLA), la métaplasie cylindro-cubique atypique (MCA), la cicatrice radiaire (CR), le papillome simple (PS) ; la tumeur phyllode (TP) étant à part.

2. Diagnostic en imagerie

Trois modalités d'imageries s'offrent aux radiologues pour la détection des lésions mammaires : la mammographie, l'échographie et l'IRM.

En mammographie, une lésion peut être une asymétrie de densité focale (zone glandulaire de densité plus élevée localement), une masse (visible sur plusieurs incidences), une distorsion architecturale (rupture focale de l'architecture) (*figure 1*) ou un foyer de microcalcifications (> 5 microcalcifications mesurant chacune moins de 1mm sur 1cm²) (*figure 2*). Ce dernier est visible dans la très grande majorité des cas exclusivement sur les mammographies, indispensables dans leur analyse (quelques foyers de microcalcifications étant également visibles en échographie).

L'échographie et l'IRM viennent en complément selon les cas.

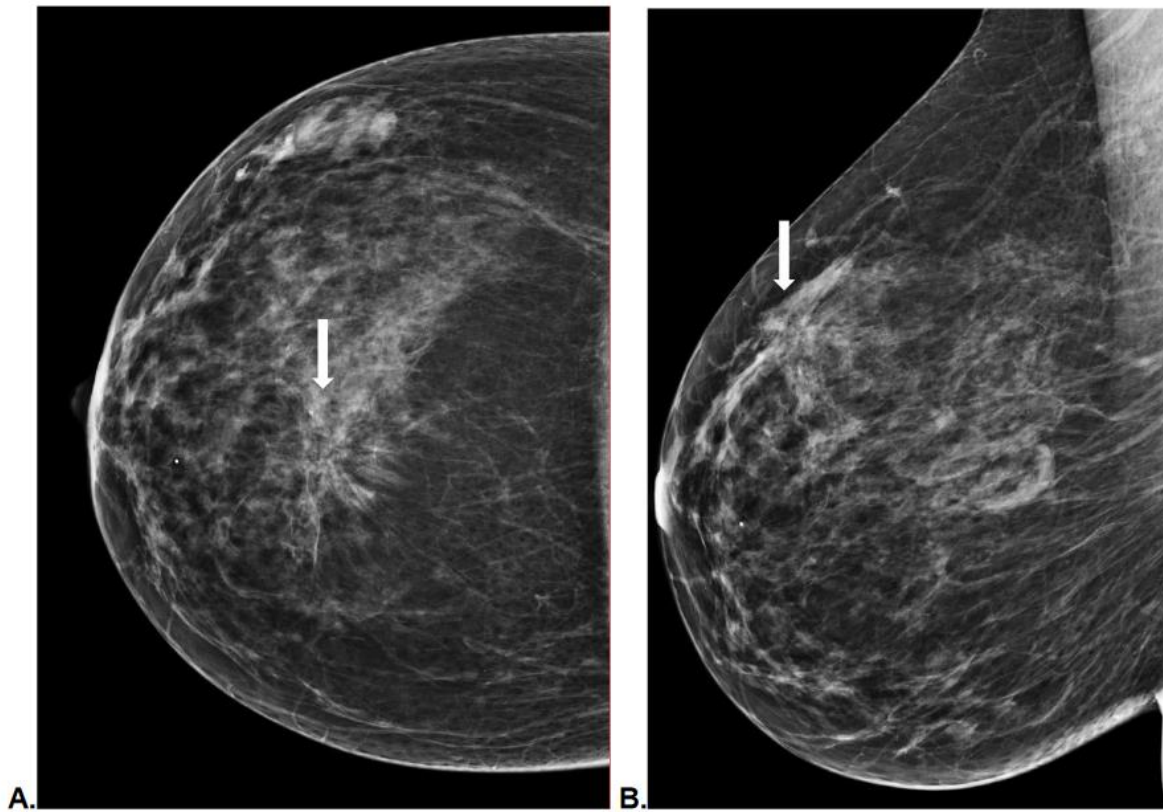


Figure 1. Images de distorsion architecturale de l'union des quadrants supérieurs (flèche) sur des mammographies du sein droit de face (A) et de profil (B).

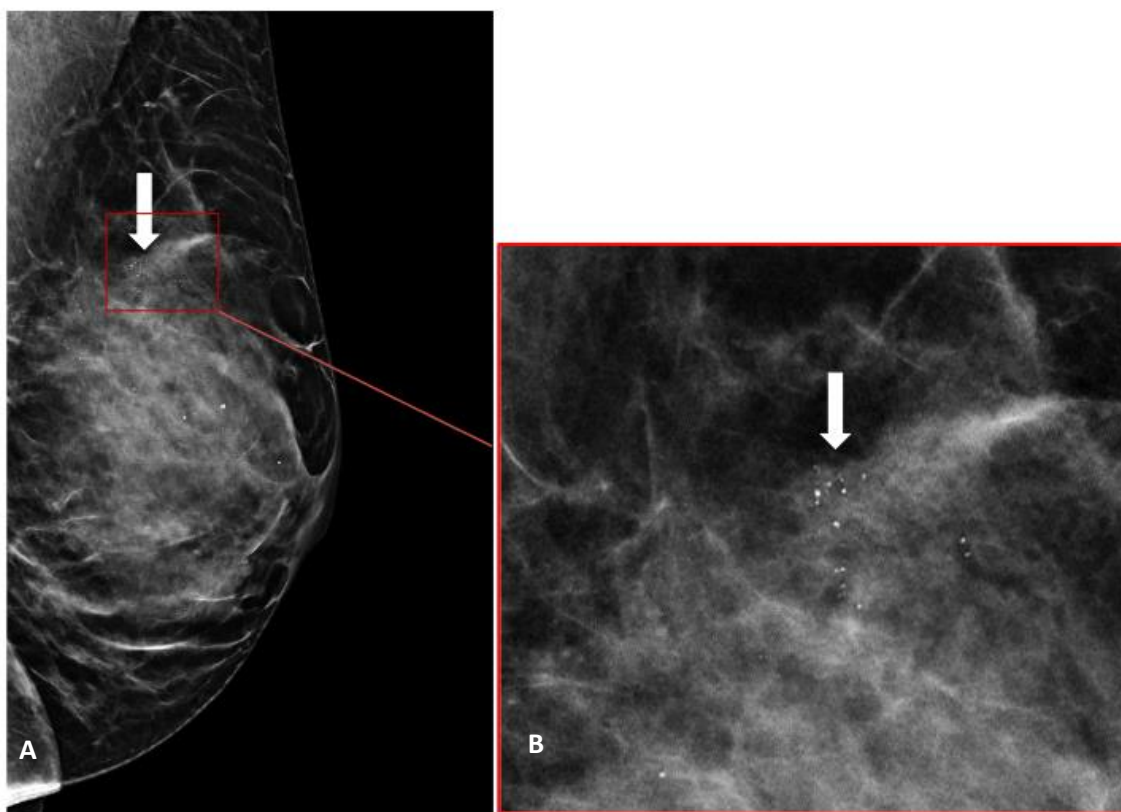


Figure 2. (A) Mammographie et zoom sur le cliché agrandi de profil gauche. (B) Foyer de microcalcifications rondes, fines et polymorphes (flèches), ACR 4A.

C'est devant le besoin de standardiser les comptes-rendus pour homogénéiser les pratiques qu'en 1990 a été mise au point la classification internationale de l'ACR (American College of Radiology) : le lexique BI-RADS (Breast Imaging-Reporting And Data System).

Il détermine un langage commun devant une lésion mammaire détectée selon les différentes modalités d'imageries disponibles et permet d'établir une conduite à tenir standardisée. Chaque lésion est classée selon un degré de suspicion radiologique de cancer mammaire en 5 catégories :

ACR 1 : Pas d'anomalie.

ACR 2 : Anomalie bénigne ne nécessitant ni surveillance ni examen complémentaire (VPP de cancer = 0%), ex : kyste simple, microcalcifications typiques de « lait calcique ».

ACR 3 : Anomalie probablement bénigne ; une surveillance à court terme est conseillée (VPP < 2 %), ex : microcalcifications rondes en amas.

ACR 4 : Anomalie indéterminée ou suspecte qui indique une vérification histologique (2 % < VPP < 95 %).

4A : faiblement suspecte : > 2 % à ≤ 10 %

4B : modérément suspecte : > 10 % à ≤ 50 %

4C : très suspecte : > 50 % à < 95 %

ACR 5 : Anomalie considérée comme maligne (VPP de cancer ≥ 95 %)

La haute autorité de santé (HAS) recommande la réalisation de prélèvements per cutanés pour les lésions classées ACR 4 ou 5 pour analyse anatomopathologique [2]. Dans certaines circonstances, les lésions ACR 3 sont également biopsiées (cancer homolatéral concomitant, patiente à très haut risque génétique, évolution d'une lésion ACR 3, difficultés de surveillance ou anxiété de la patiente...).

Les lésions frontières (B3) se traduisent en grande partie par des anomalies mammographiques, notamment à type de foyers de microcalcifications ou de distorsion architecturale, en dehors du papillome surtout retrouvé en échographie.

Les prélèvements de ces lésions s'effectueront majoritairement par guidage stéréotaxique ou sous tomosynthèse (*Digital breast tomosynthesis, DBT*). La tomosynthèse consiste en plusieurs acquisitions mammographiques le long d'un angle de -15 à 60 degrés avec reconstitutions 3D par un algorithme mathématique. Les aiguilles utilisées seront de gros calibre (macrobiopsie) car le taux de sous-estimation

de lésions in situ est élevé (sous-diagnostic de carcinome canalaire in situ, CCIS) de 10,7% pour des microbiopsies de 14G contre 1,6% avec des macrobiopsies en 11G) [8]. Le nombre de prélèvements recommandé est de 12 prélèvements en 11G et de 8 prélèvements en 7G, chaque prélèvement pesant environ 4 g [7].

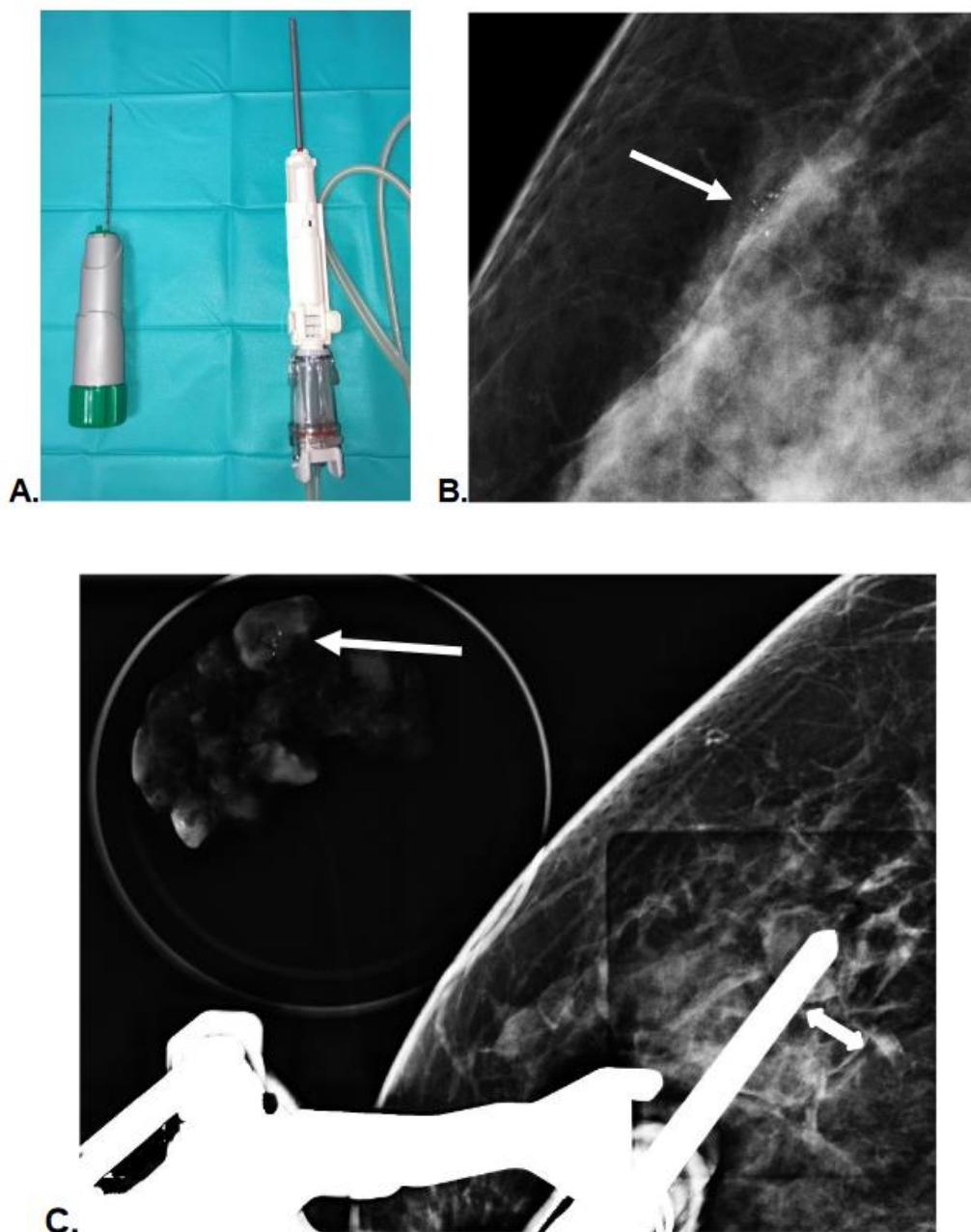


Figure 3. (A) Aiguilles de biopsies : à gauche, microbiopsie en 14G ; à droite, macrobiopsie en 7G. (B) Foyer de 12mm de microcalcifications grossières et hétérogènes (flèche) du sein droit, ACR4C. (C) Macrobiopsie sous tomosynthèse du foyer de microcalcifications (double flèche) et radiographie des prélèvements (flèche simple). Les résultats histologiques montrent des foyers de métaplasie cylindrique simple et atypique et d'hyperplasie canalaire atypique. Après RCP radio-histologique, une exérèse complémentaire par chirurgie est recommandée du fait de la coexistence de deux lésions à risque (HCA et MCA) et de l'exérèse incomplète du foyer. Indication de surveillance mammographique annuelle pendant 10 ans puis tous les 2 ans [9].

3. Problématique

En généralisant le dépistage du cancer du sein, le nombre de lésions infra-cliniques n'a cessé d'augmenter et par conséquent le nombre de biopsies ainsi que le nombre de lésions frontières.

Ces dernières représentent environ 7% des biopsies avec une très grande variabilité selon les centres (4 à 22%) [4]. Leur prise en charge a longtemps reposé sur la chirurgie devant leur risque de sous-estimation de 9,9% à 35,1% [6].

Cependant, depuis plusieurs années, la tendance est à la désescalade chirurgicale avec notamment le développement de la macrobiopsie d'exérèse comme alternative à la chirurgie. Si certains critères cliniques, radiologiques ou histologiques sont connus pour majorer le risque de sous-estimation de ces lésions frontières, il n'existe pas à l'heure actuelle de recommandation officielle sur leur prise en charge qui reste pays et centre dépendante.

L'étude rétrospective monocentrique observationnelle descriptive sur une période de 6 ans, de novembre 2015 à juillet 2021, récemment réalisée au CHU Grenoble Alpes (thèse pour l'obtention du DES radiodiagnostic et imagerie médicale de Mme Shirin BENHAMICHE, dirigée par le Dr DELOUCHE) a permis l'analyse de 153 lésions B3 (11,26% de l'ensemble des biopsies mammaires réalisées en 6 ans (256/2272)) [9].

Cette dernière, dont les résultats sont proches des données de la littérature, met en évidence que 60% des lésions B3 ont été opérées. Cela montre que les décisions prises en RCP au CHUGA sont encore plutôt favorables à la chirurgie qu'à une surveillance.

Le taux de sous-estimation global est de 13,72% (21/153), concordant avec la littérature (15,6%) [10] qui décrit une grande variabilité allant de 9,9% à 35,1% pour l'ensemble des lésions frontières [7]. Le taux de sous-estimation par sous-type histologique est également semblable à celui d'autres équipes : les deux sous-types histologiques les plus sous-estimés sont le papillome atypique (33,33%) et l'HCA (25%).

Il a également été cherché à déterminer quel facteur était significativement associé à la découverte de lésions cancéreuses sur pièce opératoire, notamment si la classification radiologique (ACR) permettait de prédire cette malignité. Cliniquement, seul l'antécédent personnel de cancer du sein était significativement plus important dans le groupe B5 comparativement aux groupes B2 et B3 (p-value 0,022). Aucun critère radiologique ne se distinguait dans le groupe B5. La cible était majoritairement

des microcalcifications en amas, fines et polymorphes, sans que ce soit significativement différent des autres groupes (p-values 0,48 et 0,54).

À partir de ce constat, nous nous sommes posé la question de savoir si un des algorithmes d'intelligence artificielle (IA) disponible sur le marché en sénologie pourrait être une aide supplémentaire à la décision thérapeutique des lésions frontières B3. La problématique posée était : peut-on s'appuyer des résultats de l'IA pour prédire la malignité des pièces opératoires de lésions frontières trouvées sur macrobiopsies ? Ainsi au décours des prises de décisions thérapeutiques des lésions frontières discutées en RCP (abstention, chirurgie, surveillance ou exérèse par macro-biopsie), en plus des données radiologiques et médicales classiques permettant d'influencer le choix de traitement, les données de l'IA pourraient également faire partie intégrante de la décision médicale.

B) Intelligence artificielle

1. Définition

La dernière décennie peut être marquée comme l'ère de l'intelligence artificielle en raison des avancées technologiques massives qui ont permis un accès, un traitement et un stockage facile d'énormes quantités de data (données).

L'IA est un terme général faisant partie d'une branche de l'informatique correspondant à l'ensemble des théories et techniques via l'utilisation d'ordinateurs et d'algorithmes permettant de modéliser des tâches traditionnellement associées à l'intelligence humaine (compréhension, perception ou décision) [11].

Elle a besoin d'énormément de données pour apprendre, communément appelées « big data », dont le volume dépasse la capacité d'analyse de l'homme et des outils informatiques classiques, le tout rendu possible par la numérisation et le stockage informatique dématérialisé, le « cloud ».

Il faut bien comprendre que la notion d'apprentissage profond ou Deep learning (DL) ayant une importante couverture médiatique actuelle est un sous type d'apprentissage automatique ou Machine Learning (ML), elle-même un sous type de l'IA globale (*figure 4*).

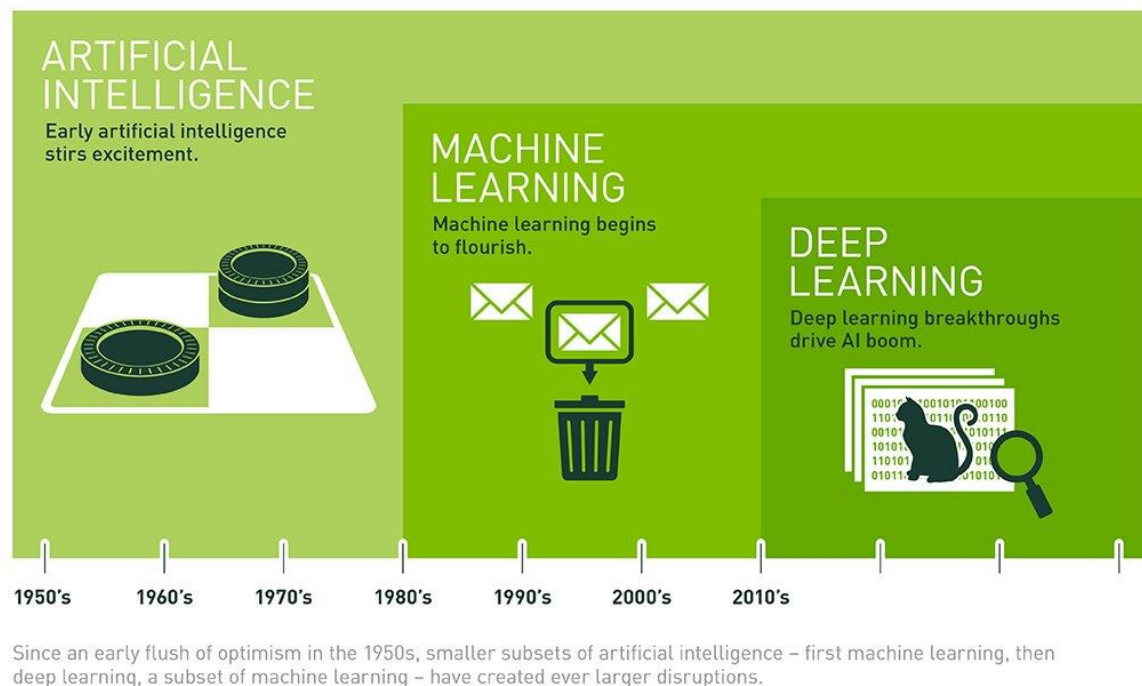


Figure 4. Intelligence artificielle, Machine Learning et Deep Learning
 (Source : <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>, consultée le 12/06/22)

a. Les prémices de l'IA (années 1950)

Un des théoriciens précurseurs de l'informatique, le mathématicien britannique Alan M. Turing, lança le concept d'IA en 1950, lorsqu'il décrit le « jeu de l'imitation » ou test de Turing dans un article resté célèbre, *Computing Machinery & Intelligence* [12]. La question qu'il posait est la suivante : un homme relié par télé-imprimante à ce qu'il ignore être une machine, disposée dans une pièce voisine, peut-il être berné et manipulé par la machine avec une efficacité comparable à celle d'un être humain ? Pour Turing, l'IA consistait donc en un simulacre de psychologie humaine aussi abouti que possible.

S'en suivit dans les dizaines d'années suivantes, plusieurs programmations et algorithmes jusqu'à l'élaboration du GPS (*general problem solver*) en 1959 qui marqua la fin de la première période de l'IA.

L'ensemble de ces programmes étaient des algorithmes non dotés de sens et se concentraient uniquement sur la tâche pour laquelle ils étaient programmés. Ils n'évoluaient pas et ne se basaient pas sur une quantité de donnée.

b. Machine learning ou apprentissage automatique (années 1980)

Arthur Samuel, informaticien américain pionnier dans le secteur de l'intelligence artificielle, est le premier à faire usage de l'expression de Machine Learning en 1959 à la suite de la création de son programme pour IBM. Le programme jouait au Jeu de Dames et s'améliorait en jouant. À terme, il parvint à battre le 4^e meilleur joueur des États-Unis.

Le ML ou apprentissage automatique est un sous-domaine de l'IA qui s'intéresse aux capacités d'apprentissages évolutifs à partir de données sans programme explicite. Il s'appuie sur le développement d'algorithmes capables de découvrir des patterns et de réaliser des prédictions, tout en améliorant leurs performances de manière autonome face à des situations pour lesquelles ils n'ont pas été programmés [13]. Le ML permet de surmonter l'explosion combinatoire, c'est-à-dire la complexité à décrire l'ensemble des comportements possibles résultant de toutes les entrées possibles.

Il existe plusieurs types d'apprentissage :

- **Supervisé** (supervised learning) : l'algorithme connaît le résultat à trouver (cible ou target), et va s'entraîner sur des données labélisées (training dataset). Le modèle va trier ces données et leurs caractéristiques en acquérant des compétences, via une exécution de la tâche de manière itérative, jusqu'aux combinaisons idéales permettant de prédire la sortie recherchée (*output*, ex. nodule bénin ou malin). Une fois que le modèle a suffisamment appris via des données d'apprentissages (nettement moindre par rapport au DL), on applique de nouvelles données de validation et de testing (données brutes de la vie réelle) pour vérifier qu'il prédit bien ce que l'on attend, avant son utilisation en routine. C'est actuellement la méthode la plus utilisée.
- **Non supervisé** (unsupervised learning) : l'algorithme va s'entraîner sur des données non étiquetées, à la recherche de caractéristiques latentes (features) pour explorer la structure des données. Cette structure pourra ensuite être réutiliser pour faire du clustering (regroupement) [14]. C'est le principe de la classification moléculaire initiale des cancers du sein (Luminal A, Luminal B, HER surexprimé et Triple négatif).

- **Semi-supervisé** : il combine une petite quantité de données étiquetées avec une grande quantité de données non étiquetées.
- **Par renforcement** : dans ce cas de figure, l'algorithme apprend en essayant encore et encore d'atteindre un objectif précis. Il pourra essayer toutes sortes de techniques pour y parvenir. Le modèle est récompensé s'il s'approche du but, ou pénalisé s'il échoue.

Dans ce sous-type d'IA, on a connaissance du fonctionnement du modèle de l'entrée à la sortie, contrairement au Deep Learning.

En application médicale, ce sont les premiers CAD (computer aided diagnoses) qui ont dominé initialement, permettant la détection et la classification des lésions, la segmentation automatique ou le triage de données. Ils ont été en partie abandonnés notamment en sénologie du fait de leur manque de spécificité [15] (cf. *IA en sénologie*).

c. Deep learning ou apprentissage profond (figure 5)

En 2012, l'équipe de *Krizhevsky et al.* a remporté le INLSVRC 2012 (image net large scale visual recognition challenge) [16], grand « data challenge » de reconnaissance d'image, par une notion nouvelle par rapport au machine learning : le deep learning ou apprentissage profond, connu sous le nom de réseaux de neurones. Grâce à ces réseaux de neurones et à l'amélioration de la puissance de calcul par les GPU (Graphic Processing Units), la performance qui était stable à 74% depuis plusieurs années est passée à 84%, pour ne cesser d'augmenter d'année en année depuis.

Le DL est défini comme des outils statistiques dans lequel des réseaux de neurones à plusieurs couches (analogie avec le neurone biologique) sont utilisés pour apprendre un modèle complexe à partir de données d'entrée (*input*). L'aspect « profond » (*deep*) du DL fait référence à l'architecture multicouche des neurones artificiels (*multilayers perceptrons*) de la couche intermédiaire, où chaque neurone est connecté à tous les neurones en amont et en aval. Cette dernière est nommée couche cachée (*hidden layers*), car l'humain ne comprend pas exactement son fonctionnement (« boîte noire »). La caractéristique principale du DL est l'apprentissage itératif sans intervention humaine dans le but d'obtenir une prédiction claire (ex. cancer ou non) à

partir de données. Contrairement au ML, l'humain n'invente pas de descripteurs car le réseau le fait tout seul, ce qui est un avantage mais aussi un inconvénient.

La dernière décennie a consacré l'apprentissage profond comme le sous-type le plus fréquent d'apprentissage machine dans le domaine de la santé. L'apprentissage profond a montré des performances parfois proches de consensus d'expert dans la détection d'anomalies radiologiques [17]. Les applications potentielles sont multiples en recherche clinique tandis que la commercialisation et les marquages par les autorités sanitaires concernent essentiellement les applications de screening dans des données médicales, basées sur des algorithmes de segmentation automatique.

Néanmoins, ces applications de l'apprentissage profond se sont faites au détriment de l'explicabilité (capacité de mettre en relation et de rendre compréhensible les éléments pris en compte par le système d'IA pour la production de résultats) avec la dénonciation d'un effet « boîte noire ». Il est difficile actuellement de savoir comment l'algorithme arrive à un résultat donné, et il n'y a pas de possibilité de décomposer le modèle ou de le simuler [17]. Problématique également, les tentatives d'explicabilité par des analyses à posteriori à l'aide de « heat maps » ont montré que les algorithmes d'apprentissage profond pouvaient arriver à la bonne décision à partir d'informations non médicales et que l'apprentissage erroné de ces raccourcis pouvait être étendu à une base de validation externe sans que l'expert ne s'en aperçoive. Ces limites ont amené des auteurs à dénoncer les faux espoirs de l'explicabilité de l'apprentissage profond dans la santé [18].

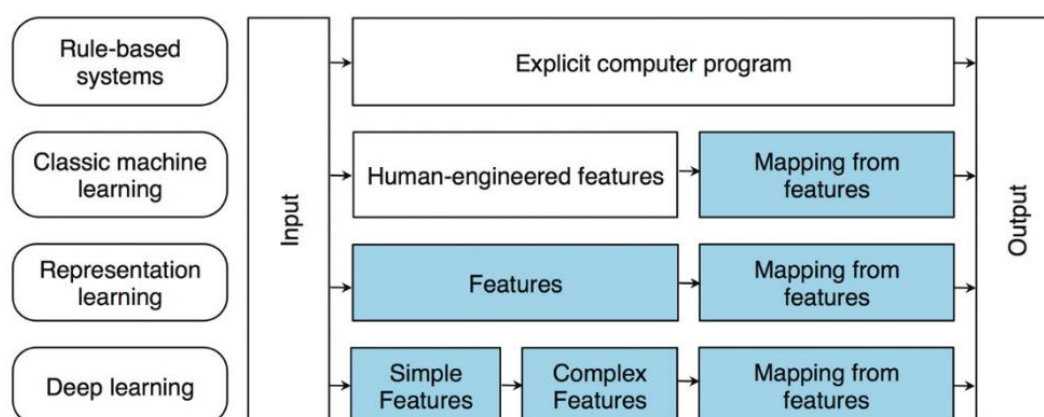


Figure 5. Fonctionnement schématique des différents types d'IA (d'après Chartrand et al. 2017) [19]. Le **Machine Learning** dépend de la segmentation sémantique des caractéristiques de l'image en descripteurs (*feature engineering*). Elle doit être rigoureuse et nécessite une expertise humaine. Le **Deep Learning** contourne le *feature engineering* en tirant parti des grandes quantités de données et de la flexibilité des modèles. Les items en bleu représentent les composants appris en ajustant un modèle à des exemples de données ; le Deep Learning permet d'apprendre un algorithme de bout en bout, de l'entrée jusqu'à la sortie, sans intervention humaine, en utilisant des données complexes de réseaux de neurones.

2. IA en sénologie

Le dépistage du cancer du sein est une application considérable de l'IA en imagerie médicale étant donné l'incidence majeure du cancer mammaire et l'importante variabilité de détections des cancers selon les radiologues [20].

Il existe cinq principaux champs d'application de l'IA en imagerie mammaire à ce jour, avec des algorithmes très différents selon le type d'utilisation.

a. Acquisition de l'image [21]

- **Optimisation de la dose d'irradiation** à travers une augmentation de la rapidité d'acquisition associée à une diminution du bruit et de la dose. Le premier sur le marché a été l'option PRIME (Progressive Reconstruction Intelligently Minimizing Exposure) proposée par le constructeur Siemens, qui permet de délivrer une dose adaptée à l'épaisseur du sein sans altérer la qualité du cliché. D'autres permettent une optimisation de la compression par un bon positionnement à travers un rétrocontrôle sur la qualité du cliché (General Electric, Siemens).
- **Contrôle qualité des mammographies digitales**, à l'exemple du National Health Service Breast Screening Programme au Royaume-Uni qui a développé un modèle permettant le contrôle qualité des clichés de tomosynthèse à l'échelle nationale, en référence aux recommandations européennes (EUREF, European Reference Organisation for Quality Assured Breast Screening and Diagnostic Services). À noter une grande difficulté d'implantation de ces logiciels en France du fait d'une importante variabilité des doses et de la qualité des clichés selon les centres ainsi qu'un nombre important de constructeurs différents sur l'ensemble du territoire.
- **Mammographie synthétique** (*Synthetic mammography (SM)*) (figure 6). La tomosynthèse ne rentre toujours pas dans les recommandations officielles du dépistage en France malgré son aide incontestable et un récent avis favorable de la commission Européenne spéciale de lutte contre le cancer du sein (réalisation de clichés 2D ou d'une tomosynthèse en première lecture) [22]. Cela implique en routine clinique une augmentation de la dose d'irradiation (clichés 2D puis tomosynthèse sur une ou deux incidences). Des modèles sont apparus permettant de créer des clichés de mammographies 2D à partir de la

tomosynthèse, diminuant ainsi la dose d'irradiation (constructeur Hologic notamment). Il a été démontré la non-infériorité de la combinaison tomosynthèse et mammographies synthétiques versus mammographies 2D natives [23]. Une augmentation de la détection de cancer a également été décrite tout en diminuant la dose d'irradiation par deux. La combinaison DBT et SM semble également permettre une réduction du nombre de biopsie par rapport à l'association actuelle (mammographies 2D et tomosynthèse) [24].

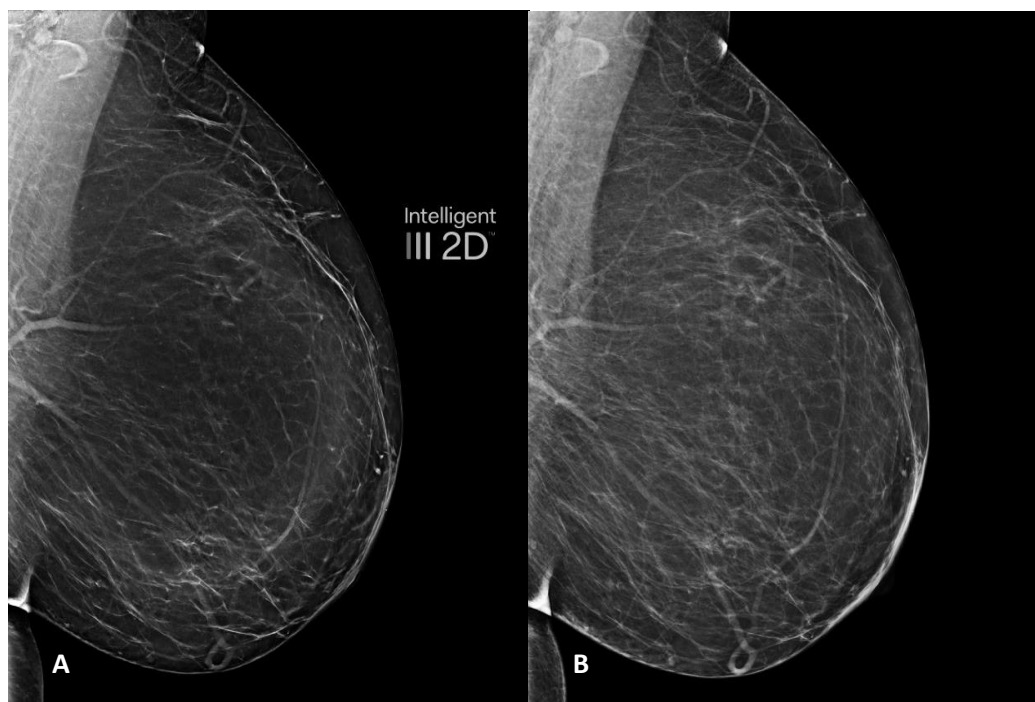


Figure 6. (A) Cliché 2D « synthétique » oblique externe gauche élaboré à partir des clichés de tomosynthèse (3D) à l'aide d'un algorithme proposé par le constructeur Hologic. Le cliché est systématiquement annoté : « intelligent III 2D ». **(B)** Cliché 2D natif oblique externe gauche.

b. Détection automatique de la densité mammaire

La densité mammaire élevée (C ou D), qui est une indication d'échographie complémentaire, est un facteur de risque indépendant de cancer du sein associé à un nombre plus élevé de cancer de l'intervalle [25]. Étant donné la faible reproductibilité inter et intra-radiologue, notamment entre les densités B et C (17% des femmes ont des densités variables selon les radiologues [26]), plusieurs modèles ont été développés avec des performances proches du score BI RADS [27], à l'exemple du logiciel QUANTRA 2.2 du constructeur Hologic (*figure 7*), fonctionnant à l'aide de modèle de machine learning.

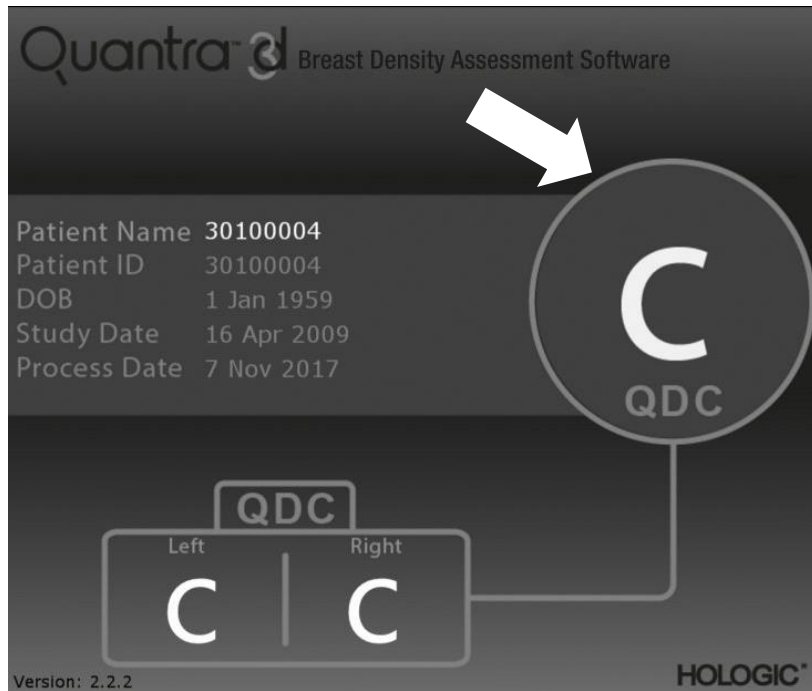


Figure 7. Classification d'un bilan mammographie selon le lexique BIRADS par QUANTRA d'Hologic, disponible sur la console d'interprétation. Chaque sein est étudié indépendamment et l'algorithme donne une densité globale pour les deux seins (flèche). Dans cet exemple, la densité des seins gauche et droit est de type C. La densité de l'ensemble du bilan mammographique est de type C également. L'échographie mammaire complémentaire est indiquée.

c. CAD ou système d'aide à la décision

- **CAD initiaux :**

Devant l'important flux de travail du dépistage et la variabilité intra et inter individuelle de la détection des cancers mammaires (10 à 30%) [28] et pour diminuer les faux négatifs, des systèmes d'aide informatisés à la détection et l'interprétation dérivés du ML ont vu le jour. Lors de l'introduction de ces algorithmes, une grande promesse d'amélioration des performances était prévue, avec des résultats initiaux encourageants [20]. Après l'approbation du modèle *Image Checker M1000* aux Etats-Unis (détection des masses et microcalcifications avec prédiction de risque de cancer), les services de santé américains ont augmenté le remboursement des CAD pour l'interprétation des mammographies, ce qui a conduit à une augmentation de leur utilisation de 5 % en 2003 à 83 % en 2012 [28].

Après des années d'utilisation clinique, plusieurs études à grande échelle montraient qu'il n'y avait pas de différence dans le taux de détection des cancers (4,1 cancer pour 1000 femmes dépistées avec ou sans CAD) [29], voire que les performances avec CAD étaient inférieures à celles du radiologue sans CAD (sensibilité à 83% avec,

contre 89% sans). Il y avait également une augmentation de 20% de biopsies réalisées ainsi qu'un taux de rappel plus élevé [30]. En ce qui concerne la réduction du nombre de cancers non initialement diagnostiqués par le radiologue, plus de 1000 faux positifs devaient être pris en compte pour qu'un cancer supplémentaire soit détecté [31]. Les résultats étaient similaires avec la tomosynthèse [32].

Dans l'ensemble, l'utilisation de ces CAD était responsable d'un grand nombre de faux positifs, ce qui diminuait la spécificité et la valeur prédictive positive sans augmentation significative de la sensibilité [33].

- **AI-CAD ou nouvelle génération d'aide à la décision dérivé du DL**

Au cours des dernières années, des produits commerciaux plus matures ont été développés, faisant émerger une myriade d'études.

- Interprétation autonome par l'IA

Initialement, les études se sont intéressées à la comparaison d'interprétation des mammographies par les modèles seuls versus radiologue seul, avec des résultats initiaux intéressants. Les performances des modèles étaient supérieures aux radiologues à l'exception des radiologues expérimentés [34].

La revue systématique de *Freeman and al.* a freiné cet engouement. Il en ressortait que 96% des 36 logiciels utilisés avaient des performances inférieures aux radiologues, notamment sur la spécificité [35]. L'analyse récente *INSPIREED 003 (Pfof and al.)* montrait qu'à ce jour, l'étude multimodalité du dépistage avec l'échographie mammaire restait supérieure aux algorithmes commercialisés [36].

- Interprétation par le radiologue aidé de l'IA

D'autres études, plus encourageantes, se sont penchées sur un modèle du radiologue augmenté par l'IA. Il en ressort schématiquement que la performance du radiologue aidé de l'IA apparaît supérieure au radiologue seul [37] avec sensibilité (86% vs 83%) et spécificité augmentées (79% vs 77%), notamment sur les masses et distorsions [38].

- Dépistage et triage des mammographies

Rodriguez et al. a montré que l'on pouvait réduire de 20% le nombre de mammographies de dépistage, en excluant les mammographies ayant un score de 0 à 2 (score allant jusqu'à 10, risque maximal de cancer) sur le logiciel Transpara®, sans perte de chance (0 cancer non diagnostiqué dans l'étude) [39]. *Povedano et al.* a considéré qu'avec le même logiciel et si l'on tolérait une perte de 5% de sensibilité (absence de lecture par un radiologue des mammographies ayant un score de 0 à 7 par Transpara®), on diminuait la volumétrie de 70% en augmentant la sensibilité et avec un temps d'interprétation globalement identique [40]. Une autre étude a estimé que l'analyse de la tomosynthèse seule par un modèle était non-inférieure au dépistage organisé actuel (double lecture des clichés 2D par des radiologues différents) [41].

Ce gain de temps permettrait d'intégrer au dépistage la tomosynthèse en double lecture, ce qui serait une valeur ajoutée à la simple lecture actuelle. Cependant, compte tenu du nombre de radiologues disponibles et du temps de lecture actuel, son intégration n'est pas envisagée à court terme.

- Cancer de l'intervalle

Deux études sur des petites séries ont montré que l'on pouvait réduire d'environ de moitié le taux de cancer de l'intervalle (47%, Transpara® [42] et 48% iCAD® [41]). Cette performance est à approfondir compte tenu de l'importance pronostique de la taille lésionnelle à la survie des patientes. *Wanders et al.* est allé plus loin en étudiant un modèle associant le score d'un logiciel IA à la densité mammaire. Ce modèle était plus performant que l'étude de la densité seule ou du modèle d'IA seul, en détectant d'avantage de cancers de l'intervalle invasifs [43].

D'autres modèles existent pour l'étude remnographique et échographique. Par exemple, les systèmes d'échographies mammaires automatisées (ABUS, automated breast ultrasound sreening), avec reconstruction 3D et système d'aide au diagnostic (CAD) semblent réduire la variabilité inter-individuelle tout en en diminuant le temps de lecture sans compromis sur le diagnostic [44].

Plusieurs facteurs de variation de la performance des modèles en pratique courante ont été décrits :

- Population training différente de celle testée [45],
- Analyse d'un seul cliché par le modèle contre une analyse des deux seins par le radiologue, et donc une absence d'étude de l'asymétrie de densité, diminuant la spécificité [41],
- Absence d'analyse concomitante des antériorités,
- Type d'anomalie : meilleure étude des masses et distorsions architecturales par rapport aux foyers de microcalcifications [46].

L'étude méthodologique de *Freeman et al.* portant sur les biais des études antérieures et l'applicabilité de l'IA, est venue rappeler que l'intégration des modèles au dépistage organisé du cancer du sein était un long processus pour acquérir les qualités optimales à leur intégration en routine clinique (*figure 8*).

Study reference	Risk of bias					Applicability concerns					
	Patient selection	Index test	Comparator test	Reference standard	Flow and timing	Patient selection	Index test	Comparator test	Reference standard		
Standalone AI systems (5 studies)											
Lotter 2021 ²⁸	High	High	High	Unclear	Unclear	High	High	High	High		
McKinney 2020 ²⁹	High	High	Low	High	High	High	High	High	Low		
Rodriguez-Ruiz 2019 ³³	High	High	High	Unclear	Unclear	High	High	High	Unclear		
Salim 2020 ³⁵	High	High	Low	High	High	High	High	Low*	High*		
Schaffter 2020 ³⁶	Low	High	Low	High	High	Unclear	High	Low*	High*		
AI as reader aid (3 studies)											
Pacilè 2020 ³⁰	High	High	High	High	Unclear	High	High	High	High		
Rodriguez-Ruiz 2019 ^{32,34}	High	High	High	High	Unclear	High	High	High	High		
Watanabe 2019 ³⁷	High	High	High	Unclear	Unclear	High	High	High	Low		
AI for triage (4 studies)											
Balta 2020 ²⁵	Low	High	None	High	High	Low	High	None	High		
Dembrower 2020 ²⁶	High	High	None	Low†	High†	High	High	High	None	Low†	High†
Lång 2020 ²⁷	Low	High	None	High	High	Low	High	None	High		
Raya-Povedano 2021 ³¹	Low	Low	Low	Low	High	Low	High	High	Low		

Figure 8. Aperçu des risques de biais et d'applicabilité des études récentes portant sur l'utilisation des modèles d'IA dans le dépistage du cancer du sein (*selon Freeman et al. 2021*) [35]. **Bas*** si lecture par plusieurs radiologues, **Haut*** pour une lecture médicale seule par rapport au modèle. **Bas †** pour le précédent dépistage (cancer prouvé sur biopsie ou suivi pendant 2 ans pour les test négatifs), **Haut †** pour le dépistage actuel (cancer prouvé sur biopsie sans suivie pour les test négatifs). Transpara® était le modèle analysé pour les études encadrées en jaune.

d. Analyse radiomique de la tumeur (figure 9)

En imagerie, les *inputs* peuvent être des pixels ou des paramètres quantitatifs tels que les radiomiques. Cela consiste à extraire des données pertinentes depuis les examens d'imageries pour aboutir à des diagnostics de précision et notamment des informations prédictives ou pronostiques de pathologies cancéreuses. On parlera ainsi de "biomarqueurs d'imagerie". Ces variables pertinentes ne sont pas forcément visibles à l'œil du radiologue [47]. À partir de l'image, le modèle est capable de trouver des critères pronostiques, en partie connus de l'homme, comme l'agressivité tumorale. On sait, par exemple, que les cancers triples négatifs sont plutôt de forme ronde, alors que les lésions de type Luminal A et B (peu agressives) sont souvent spiculées; ou que la présence de calcifications est en faveur de lésions sur-exprimant HER2. L'imagerie serait alors la seule technique à prédire la biologie des cancers en visualisant la tumeur dans son ensemble, ce que l'histopathologie, à travers l'échantillonnage, ne peut réaliser. L'analyse radiomique de la tumeur deviendrait un élément clé de la prise en charge étant donné que l'hétérogénéité tumorale est un facteur de résistance au traitement.

Blatshe and al. a montré que la prise de contraste d'une tumeur mammaire en IRM pouvait prédire la fixation HER2. Les tumeurs HER2+ avaient un pourcentage plus élevé de pixel se rehaussant, corrélé au VEGF. De la même façon, le Ki67 (index de prolifération) pouvait être prédit via différents paramètres de perfusion en IRM (contraste, pente de la courbe, ...) [48]. C'est un élément pronostique essentiel qui témoigne de l'agressivité tumorale et indique une chimiothérapie s'il est élevé (> 10% à 20%).

Une autre étude s'est intéressée au score Oncotype DX®, score permettant de prédire le bénéfice d'une chimiothérapie pour les cancers invasifs de petite taille RH + HER2 - sans atteinte ganglionnaire, et donc d'orienter la décision thérapeutique. Il est déterminé à travers une analyse génomique d'un échantillon tumoral, nécessitant donc une biopsie ou chirurgie, avec un délai d'attente pouvant être de plusieurs semaines. *Saha et al.* a démontré que l'on pouvait prédire ce score (AUC à 0,77, p à 0,003) à travers des paramètres radiomiques tumoraux (wash-in, wash-out, pente de rehaussement, nombre de voxels réhaussés en MIP) [49]. Enfin, *Braman et al.* a quant à lui montré que la signature radiogénomique de la tumeur et de son environnement permettait de déterminer les cancers susceptibles de répondre au traitement anti-HER2 [50].

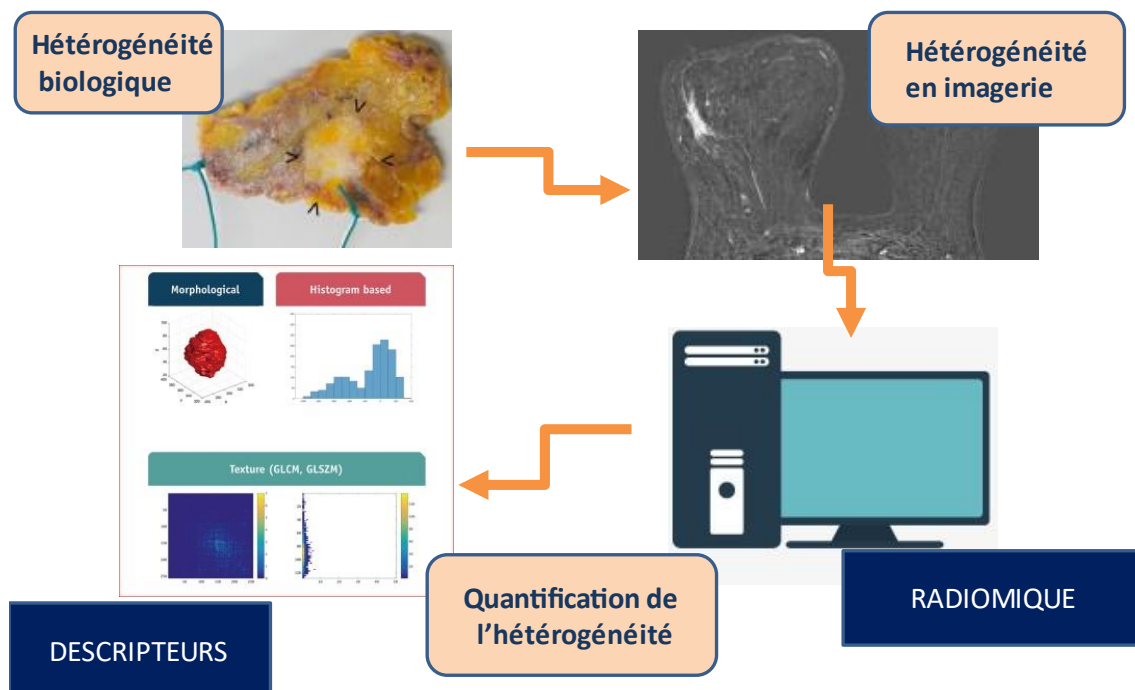


Figure 9. Résumé schématique de l'analyse radiomique. L'hétérogénéité tumorale, facteur pronostique essentiel, se traduit par une hétérogénéité en imagerie médicale. À travers l'analyse radiomique faisant intervenir des modèles d'intelligence artificielle, il y a une mesure quantitative des paramètres de l'image à partir de l'analyse des pixels et voxels. Il en résulte différents paramètres de textures (ou descripteurs). L'ensemble de ces descripteurs permet une analyse radiologique au-delà de la prédiction, comme le type histologique, la prédiction du Ki67 ou de la fixation HER2.

e. Dépistage personnalisé

Les recommandations françaises prévoient un dépistage particulier et annuel pour les patientes considérées à très haut risque de cancer (génétique ou considéré équivalent). Pour les patientes à risque intermédiaire et bas il n'y a pas de conduite spécifique à tenir en dehors du dépistage organisé. La question s'est posée de savoir si les patientes à bas risque pourraient espacer de plus de 2 ans leur dépistage alors que celles à risque intermédiaire - haut risque pourraient être suivies annuellement.

C'est l'objet de deux grandes études internationales en cours : WISDOW TRIAL au Etats-Unis et MyPEBS (My Personal Breast Screening) en Europe. Cette dernière est la première étude clinique européenne randomisée qui vise à évaluer les bénéfices d'un dépistage dont la fréquence sera adaptée au risque individuel de cancer du sein de chaque femme. Elle fait intervenir un modèle d'IA, MammoRisk®, qui prend en compte la densité mammaire, les antécédents au premier degré, l'âge de la patiente

et les antécédents de biopsies mammaires. Il s'en suit une stadification du risque à l'aide des données de l'algorithme associées à celles de l'étude du polymorphisme génétique (test salivaire) de chaque patiente incluse. Quatre niveaux de risques sont étudiés :

- Bas risque (densité A et absence d'antécédents) : mammographie tous les 4 ans.
- Risque modéré (population générale) : mammographies tous les deux ans.
- Risque élevé : mammographie annuelle.
- Très haut risque : mammographie et IRM annuelles.

3. IA en sénologie au CHUGA : Transpara®^[51]

Transpara®, de la société ScreenPoint Medical, est un logiciel conçu pour être utilisé par les médecins afin d'améliorer l'interprétation de la mammographie numérique et de la tomosynthèse chez des femmes asymptomatiques. Il est utilisé dans 20 pays à travers le monde.

Le système est destiné à être utilisé comme une aide à la lecture simultanée pour aider les lecteurs à détecter et à caractériser les anomalies potentielles et pour améliorer le flux de travail. Des algorithmes de DL sont appliqués pour la reconnaissance des calcifications et des lésions des tissus mous suspects (dont les masses, les distorsions architecturales et les asymétries).

Les algorithmes sont formés à l'aide d'une vaste base de données (plus d'un million de mammographies) contenant des exemples de cancer du sein, d'anomalies bénignes et de tissus normaux ayant fait l'objet de biopsies [52].

L'efficacité du logiciel n'a été établie que sur les incidences MLO (médio-latérale oblique), CC (cranio-caudale), ML (médio-latérale) et LM (latéro-médiale).

a. Marques fixes

Les régions considérées potentiellement suspectes de cancer par le logiciel sont marquées par un rectangle associé à une annotation, soit « calc » pour un groupe de microcalcifications soit « mass » pour une anomalie suspecte des parties molles (masse, distorsions architecturales ou asymétrie). Les clichés en tomosynthèse sont également analysés et le numéro de coupe est identifié.

b. Score régional (figure 10)

Pour chaque région détectée, l'algorithme donne un score régional de probabilité de cancer sur une échelle de 0 à 100 (plus le nombre est élevé plus le risque est important). Une seule marque est indiquée s'il coexiste dans la même localisation des microcalcifications et une anomalie des parties molles (« mass + calc »).

L'analyse des clichés MLO et CC est faite séparément ce qui explique la présence d'un score très souvent différent par cliché.

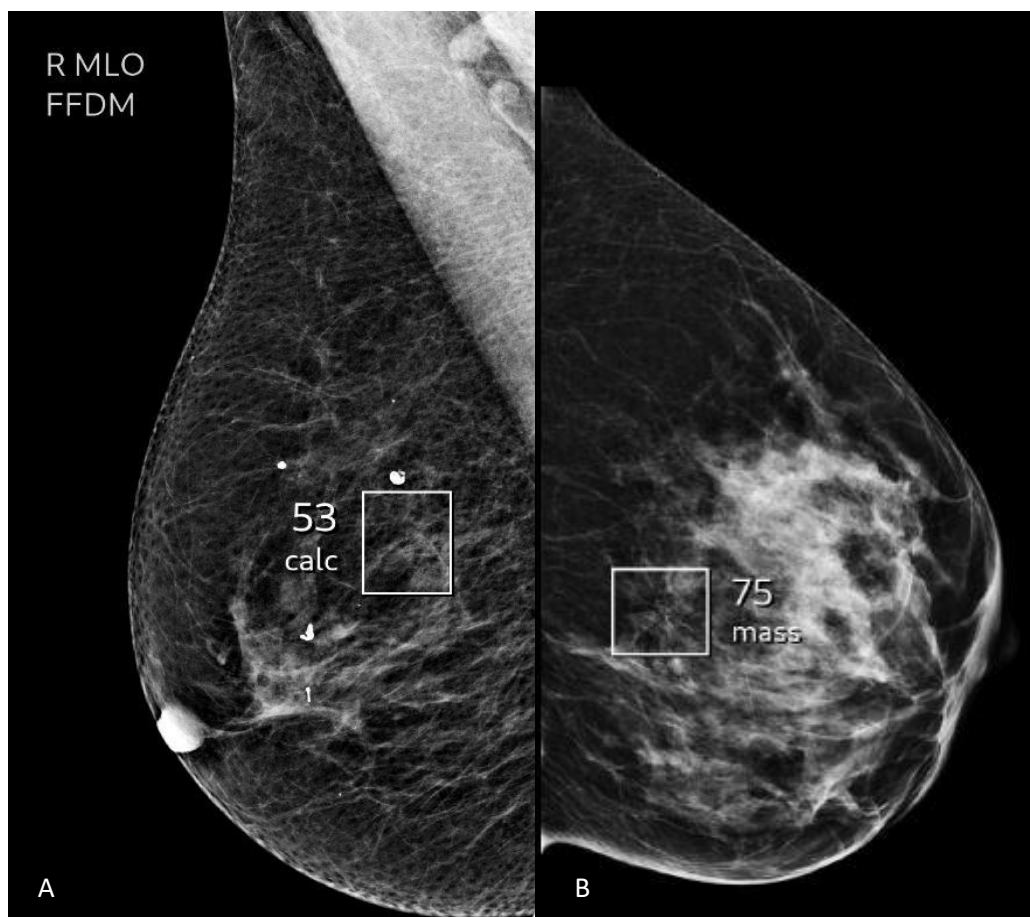


Figure 10. (A) Cliché oblique externe droit. Détection d'un foyer de microcalcification des quadrants supérieurs annoté « calc » avec un score régional de 53. **(B)** Cliché de profil gauche. Détection d'une distorsion architecturale de la région centrale annotée « mass » avec un score régional de 75.

c. Score d'examen (figures 11 et 12)

L'ensemble de l'examen mammographique, classiquement un cliché de face et oblique externe de chaque côté, est annoté d'un score plus général permettant de synthétiser la probabilité globale de cancer sur l'examen de dépistage. Il correspond au score

régional étant le plus proche de 100, s'il existe plusieurs anomalies détectées. Il est basé sur trois scores de risque différents :

- Faible : si les scores régionaux sont **< 43**. Ce seuil a été choisi comme étant le cut-off optimal en dessous duquel le risque de cancer est très faible parmi les 70% des examens les moins suspects (VPN à 99.98% [95% CI: 99.94, 99.99] en mammographie 2D et 99.99% [95% CI: 99.95, 99.99] en tomosynthèse) [40].
- Intermédiaire : si les scores se situent dans un intervalle **entre 43 et 75**. Dans cette catégorie la prévalence de cancer est équivalente à celle de la population générale du dépistage organisé, soit environ 6 à 7 cas pour 1000 patientes.
- Elevé : si au moins un score régional est **> 75**. Ce seuil a été défini par rapport au taux de détection des cancers du sein (entre 3 et 5%) des patientes dépistées. Ici la prévalence est estimée entre 1/5 à 1/10 avec une sensibilité comparable ou meilleure à la première lecture radiologique.

Ces différents seuils proviennent de la base de données de ScreenPoint et sont confirmés dans d'autres études sur la diminution du workflow ou l'analyse de la tomosynthèse par le logiciel seul [34,40]. L'absence de bijection entre le score BIRADS et ceux du logiciel est mentionné par le constructeur dans le manuel d'utilisation.



Figure 11. Score d'examen global à 53 (Intermédiaire) avec un seul foyer détecté dans le quadrant supéro-externe droit. Il est marqué sur l'oblique (53) et la face (40). Absence de foyer additionnel. Le foyer n'est pas détecté sur la tomosynthèse (flèche).

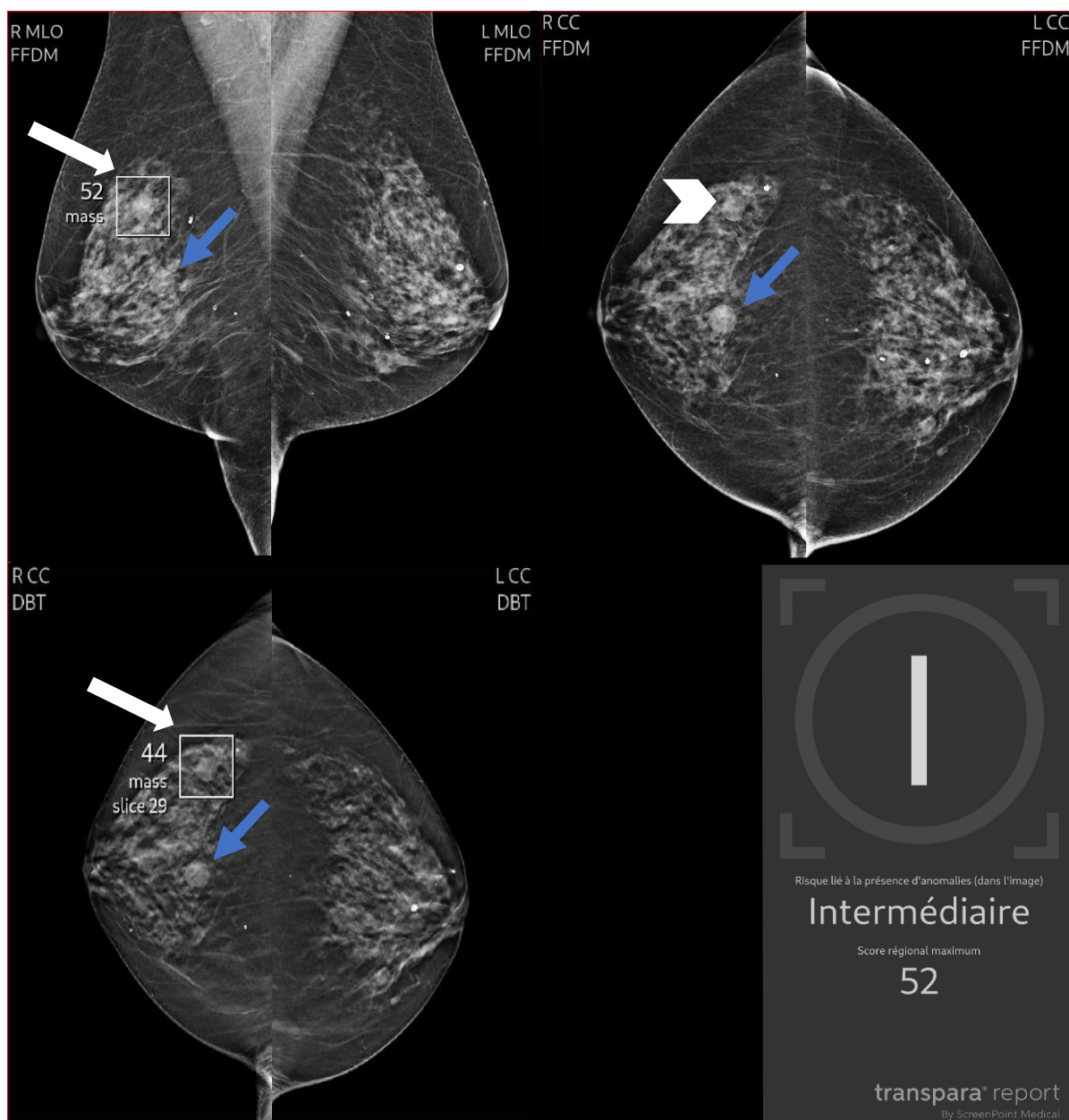


Figure 12. Masse du quadrant supéro-externe du sein droit en zone mammaire moyenne (flèche), marquée comme étant une « mass » par Transpara® avec un score régional de 52 sur l'oblique et 44 sur la tomosynthèse (coupe 29). Elle n'a pas été détectée sur l'incidence de face (tête de flèche). En échographie la lésion correspondait à un kyste simple. Le modèle n'a pas détecté l'autre masse rétromamelonnaire du sein droit (flèche bleue), correspondant à un autre kyste simple. Absence d'anomalie détectée sur le sein gauche par le modèle ou par le radiologue. Le score global de l'ensemble de l'examen est Intermédiaire (52) et correspond au score régional le plus élevé de l'ensemble des anomalies détectées par Transpara (masse droite sur le cliché oblique).

d. Marques itératives

C'est une aide à la décision interactive, c'est-à-dire que l'utilisateur pointe la lésion suspecte selon lui (non retenue par le logiciel initialement) et l'algorithme annote soit d'un cercle l'anomalie avec un score régional soit une croix s'il ne retient rien.

Les contestations sont corrélées, via l'algorithme, sur les clichés de face et oblique et le score est affiché automatiquement sur les deux clichés.

e. Aide à la navigation en tomosynthèse (DBT)

Il permet une analyse des coupes des données volumétriques de la DBT suivie de la localisation des constatations en 3D et envoie des liens vers les régions correspondantes en MLO et CC. À noter que la 2D synthétique (image 2D reconstruite à partir de la tomosynthèse) n'est pas traitée. Le logiciel ne permet pas de faire des comparaisons avec des examens antérieurs.

f. Résumé des études cliniques

À propos de Transpara®, il a été démontré par des études cliniques :

- Que son utilisation permettait une augmentation de la sensibilité des radiologues sans diminuer la spécificité ou allonger le temps de lecture [38].
- Que sa performance n'est pas inférieure à celle des radiologues [34], avec une sensibilité non inférieure ou supérieure (+ 6,3 % IC 95 % = (5,8 % - 6,8 %) ; $p = 0,016$), et une spécificité non inférieure ou supérieure (+ 0,9 % IC 95 % = (-1,1 %, + 2,8 %) ; $p = 0,38$).
- L'augmentation de la sensibilité de détection du cancer du sein sur les clichés tomosynthèse à l'aide de l'algorithme versus sans algorithme (AUC avec Transpara® = 0,863 vs AUC sans aide = 0,833 ; $p = 0,0025$).
- Une réduction du temps de lecture sur les examens normaux, ce qui implique une réduction moyenne du temps de lecture de -20 % (IC 95 % = 16 % - 25 %) dans le cadre d'un dépistage sans risque de cancer [39].

PARTIE 2 : Article

A) Introduction

Le cancer du sein est le cancer le plus fréquent chez les femmes et la première cause de décès par cancer chez ces dernières, ce qui en fait un réel enjeu de santé publique. Pour cette raison, les techniques radiologiques de dépistage se sont nettement améliorées ces dernières années. En conséquence, des lésions radiologiques de plus en plus subtiles sont découvertes et le nombre de biopsies mammaires réalisées augmente.

Les examens d'imageries sont classés en cinq catégories selon le lexique BIRADS de l'ACR : ACR 1 : examen normal, ACR 2 : anomalie bénigne, ACR 3 : anomalie probablement bénigne dans 98% des cas, nécessitant une surveillance, ACR 4 : anomalie suspecte nécessitant une biopsie et ACR 5 : lésion fortement suspecte de cancer, à biopsier également.

Les échantillons de biopsie sont analysés et classés en cinq catégories de B1 à B5 selon la classification du EWGBSP (European Working Group on Breast Screening Pathology [1]). Les différentes catégories correspondent à : B1, normal ; B2, bénin ; B3, potentiel malin incertain ; B4, suspect et B5, malin.

Les lésions classées B3, dites frontières, représentent environ 7 à 9% des biopsies [2] et sont un groupe hétérogène d'anomalies histologiques marqueurs d'un tissu mammaire à risque de cancer du sein. Les principales lésions B3 sont l'hyperplasie canalaire atypique (HCA), l'hyperplasie lobulaire atypique (HLA), la métaplasie cylindro-cubique atypique (MCA), le papillome simple (PS) ou atypique (PA), les cicatrices radiaires (CR), le carcinome lobulaire in situ de type classique (CLIS classique) et les tumeurs phyllodes (TP).

Leur prise en charge a longtemps reposé sur la chirurgie devant leur risque de sous-estimation (risque de trouver sur la pièce chirurgicale une lésion néoplasique non présente sur la pièce de biopsie) de 9,9% à 35,1% [3]. Cependant, depuis plusieurs années, la tendance est à la désescalade chirurgicale notamment avec le développement de la macrobiopsie d'exérèse comme alternative à la chirurgie. Si certains critères cliniques, radiologiques ou histologiques sont connus pour

majorer le risque de sous-estimation de ces lésions frontières, il n'existe pas à l'heure actuelle de recommandation officielle sur leur prise en charge qui reste pays et centre dépendante.

Depuis les quinze dernières années, le nombre de publications sur l'intelligence artificielle en imagerie médicale a considérablement augmenté pour atteindre environ 700 à 900 publications par an [4,5]. Le dépistage du cancer du sein est une des applications majeures de l'IA en radiologie pour plusieurs raisons, telles que la grande incidence du cancer du sein ou la grande variabilité d'interprétation des mammographies selon les radiologues [6,7].

De nombreux systèmes d'aide à la décision (computer-aided diagnosis, CAD) ont initialement été développés en imagerie médicale aux États-Unis durant les années 2000 avec des résultats encourageants [6]. Ils ont par la suite été largement abandonnés de par leurs performances médiocres, notamment d'une sensibilité excessive [7].

Depuis quelques années, de nouveaux modèles prédictifs plus matures ont vu le jour, avec des performances supérieures aux premières générations, dont certains font intervenir la notion d'apprentissage profond ou deep learning (DL), comme le logiciel Transpara® commercialisé par ScreenPoint, disponible au CHU Grenoble Alpes (CHUGA).

L'objectif principal de ce travail est d'étudier les performances de Transpara® pour la prédiction de malignité des pièces opératoires des anomalies mammographiques correspondant à des lésions frontières B3 sur macrobiopsies. L'hypothèse est que plus le score du modèle augmente plus le risque de cancer croît, comme pour le score radiologique BIRADS.

Après un rappel sur l'analyse descriptive de la population et des critères radiologiques associés à la malignité des pièces opératoires, étudiés par *S. Benhamiche* [8], les objectifs secondaires porteront sur la comparaison entre la distribution du score du modèle et la classification BIRADS, notamment dans l'étude des foyers de microcalcifications (distribution et morphologie) et des masses.

B) Matériels et méthodes

1. Design de l'étude

Il s'agit d'une étude diagnostique rétrospective monocentrique réalisée au CHUGA de mars 2016 à février 2022.

La validation du protocole par la Direction de la Recherche Clinique et de l'Innovation (DRCI) du CHUGA et l'autorisation à démarrer la recherche ont été obtenues en avril 2022.

Le service de radiologie a procédé à l'affichage d'une notice générique portant sur la recherche interne sur données. Cet affichage comporte toutes les mentions utiles sur la recherche au CHUGA et les modalités d'opposition.

Nous avons recensé au total 925 procédures de macrobiopsies sous mammographie (stéréotaxie ou tomosynthèse) réalisées dans cet intervalle. Le critère d'inclusion était un résultat histologique de biopsie classé B3 (165), avec des lésions de MCA, HCA, HLA, CLIS classique, CR, PS et de PA. Les lésions phyllodes dont la prise en charge est à part et systématiquement chirurgicale n'ont pas été incluses. Lorsque plusieurs types histologiques étaient associés, la lésion la plus péjorative a été déterminée selon les données de la littérature (papillome avec atypie > HCA > CLIS > HLA > MCA > CR > papillome sans atypie) [9,10].

Le critère d'exclusion principal était l'absence de chirurgie ou des résultats histologiques post-opératoires non disponibles. Une abstention chirurgicale a été décidée en RCP pour 60 patientes. L'information concernant la réalisation d'une chirurgie ou ses résultats était manquante pour 4 lésions (perdus de vue, refus de la chirurgie par la patiente, chirurgie toujours en attente au dernier jour d'inclusion).

Les autres critères d'exclusion étaient le sexe masculin, l'âge inférieur à 18 ans, le refus de la patiente à l'utilisation de ses données pour l'étude, la présence d'une lésion B5 (6 lésions) ou B4 (2 lésions) intriquée avec la lésion B3, l'absence des données de l'algorithme Transpara® (2 lésions) et un bilan mammographique diagnostique incomplet (2 lésions sans clichés de face). Les inclusions se sont terminées le 1^{er} avril 2022 et nous avons finalement inclus 89 lésions chez 86 patientes (*figure 1*).

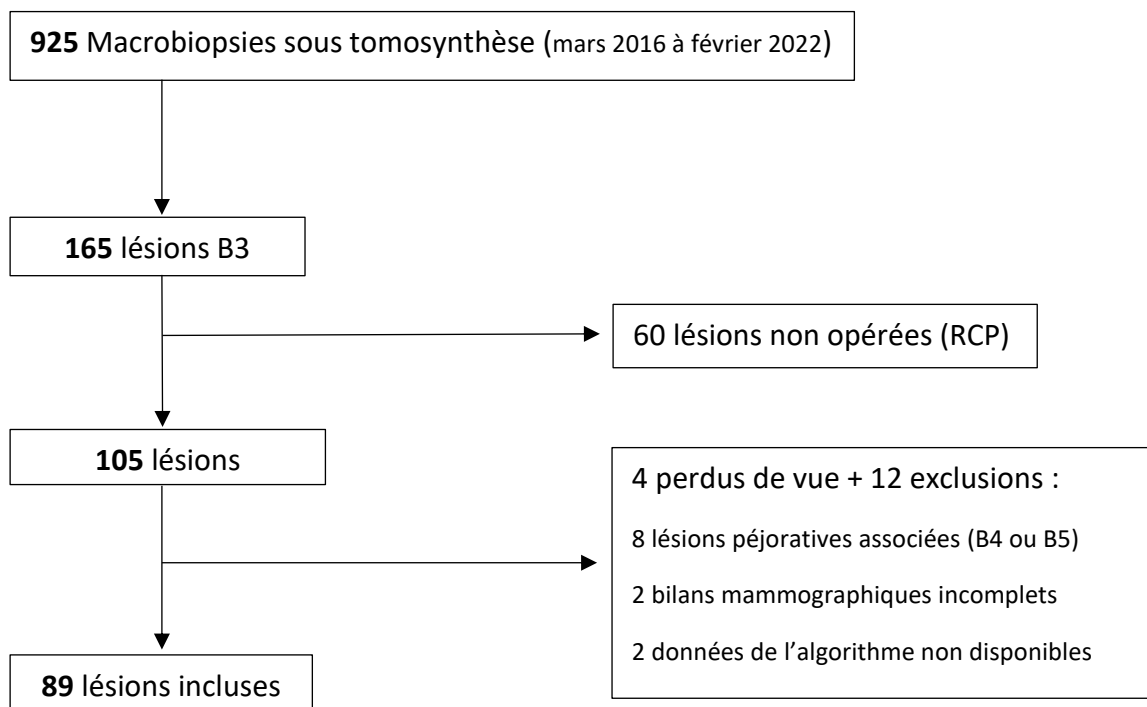


Figure 1. Flow Chart de l'étude

2. Algorithme d'intelligence artificielle Transpara®

Le système d'intelligence artificielle d'aide au diagnostic Transpara® (version 1.3.0 ; ScreenPoint Medical), qui a déjà été étudié dans d'autres publications [11-13], a été utilisé pour cette étude. Ce système utilise des algorithmes de deep learning pour analyser les mammographies 2D et les tomosynthèses ainsi que pour détecter les lésions suspectes de cancer du sein.

Les lésions les plus suspectes sont marquées et un score régional leur est attribué, de 1 à 100. Plus le score est élevé plus le risque de cancer du sein est significatif. Chaque cliché est analysé de façon indépendante et très souvent les scores diffèrent pour une même lésion.

Par la suite, une catégorie (score global) est attribuée à l'ensemble du bilan mammographique, quel que soit le nombre de lésions marquées :

- Faible : si l'ensemble des scores régionaux est **inférieur à 43**, traduisant un risque de cancer très faible (VPN à 99.98%).
- Intermédiaire : si les scores se situent dans un intervalle **entre 43 et 75**.

Dans cette catégorie la prévalence de cancer est équivalente à celle de la population générale du dépistage organisé, soit environ 6 à 7 cas pour 1000 patientes.

- Elevé : si les scores sont **supérieurs à 75**, avec une prévalence de cancer estimée entre 1/5 à 1/10.

Ces différents seuils proviennent de la base de données ScreenPoint et ont été validées par différentes études [11-13].

Le système a été installé en juin 2021 au CHUGA après validation par la DRCI.

Concernant la période de novembre 2016 à juin 2021, le CHUGA n'étant pas encore équipé du modèle, il a été appliqué à postériori sur les mammographies.

Dans cette étude, seul le score régional le plus élevé entre les deux clichés 2D a été pris en compte. Les images de tomosynthèse n'étant pas disponibles pour toutes les patientes (non réalisée sur les bilans les plus anciens), les scores attribués à ces images n'ont pas été pris en compte.

3. Analyse statistique

Pour la partie descriptive pure de notre étude, il a été réalisé des moyennes et des pourcentages. Pour la seconde partie, la population a été divisée en deux groupes selon leur histologie définitive sur pièce opératoire (B2 et B5), puis soumise à des analyses statistiques univariées à l'aide des logiciels en ligne pvalue.io et biostatgv, en utilisant le test paramétrique de Fisher. La découverte de lésions frontières B3 sur pièce n'a pas été étudiée compte tenu du fait que la prise en charge est non consensuelle, contrairement aux lésions B5 (indication de chirurgie) et B2 (absence de prise en charge thérapeutique). Pour la dernière partie, le test de normalité (Shapiro Wilk) de la variable « score IA » étant significatif ($p = 0,001$), le test non paramétrique de Wilcoxon-Mann-Whitney entre les groupes B2 et B5 a été utilisé. Les performances de l'algorithme ont été illustrées sous forme de courbe ROC (logiciel R (v 4.1.3) avec le package Epi (v 2.47)), et l'aire sous la courbe (AUC) a été estimée à l'aide de la règle trapézoïdale. Enfin, une analyse de la distribution des scores de Transpara® selon l'ACR, la distribution et la morphologie des 81 foyers de calcifications B3 a été réalisée à l'aide du test non paramétrique Kruskal-Wallis.

C) Résultats

1. Analyse descriptive de la population

a. Description de la population des lésions B3 sur biopsie [8]

Les caractéristiques histologiques (*figure 2*) et radiologiques (*tableau 1*) de la population étaient similaires à celles étudiées par S. Benhamiche, notamment près de la moitié des lésions (48,3%) était classée ACR 4B (suspicion de cancer entre 10% et 50%) et correspondait à des lésions d'hyperplasie canalaire atypique sur les biopsies dans la majorité des cas (51,7%). La seule différence notable était la nature des lésions étudiées. Contrairement à la précédente étude, 91% des lésions de notre étude étaient des foyers de microcalcifications contre 56%. À noter qu'une des cibles correspondait à l'association de microcalcifications et d'une distorsion architecturale.

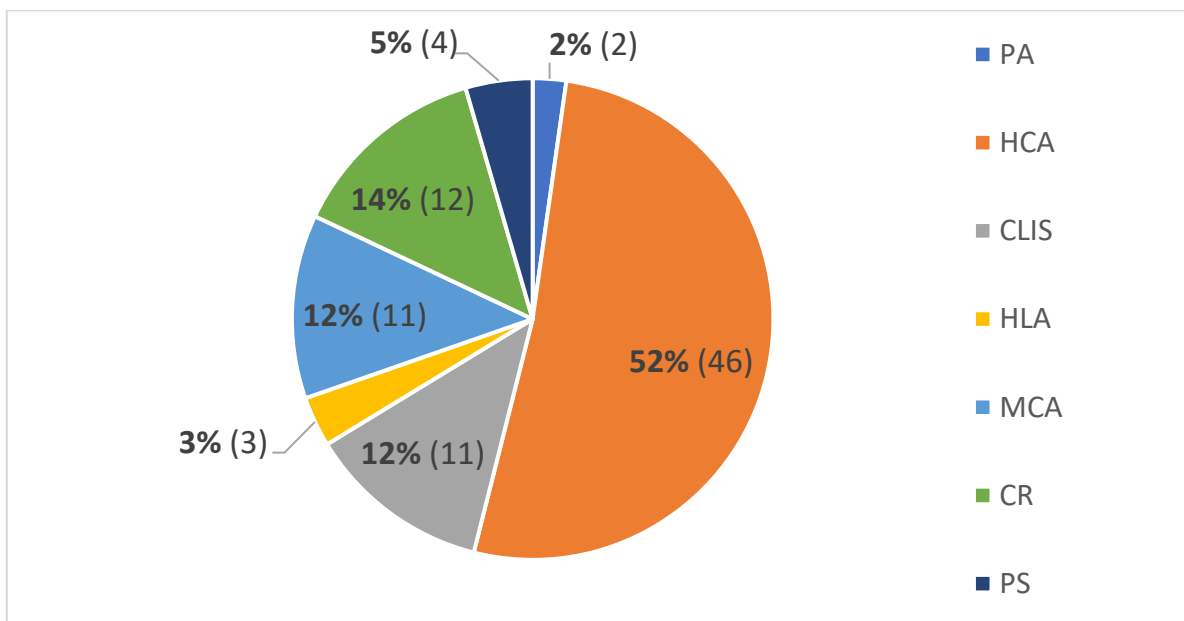


Figure 2. Répartition en pourcentage (N) des sous types histologiques B3 les plus péjoratifs sur les biopsies

		Toutes les lésions (89)	
		Nombre	%
Age moyen (années)		58	
Densité mammaire			
	A	1	1,1
	B	26	29,2
	C	58	65,1
	D	4	4,5
ACR			
	3	4	4,5
	4A	21	23,6
	4B	43	48,3
	4C	21	23,6
	5	0	0
Distorsion architecturale			
	Oui	9	10,1
	Non	80	89,9
Microcalcifications			
	Oui	81	91
	Non	8	9
Distribution			
	Régionale	13	16,1
	Amas	54	66,7
	Linéaire	10	12,3
	Segmentaire	4	4,9
Morphologie			
	Amorphes	21	25,9
	Grossières Hétérogènes	16	19,7
	Fines Polymorphes	43	53,1
	Linéaires ou linéaires branchées	1	1,2

Tableau 1. Caractéristiques radiologiques de la population B3 sur biopsie

b. Analyse descriptive des pièces opératoires

Dans cette partie, notre population a été divisée en trois groupes selon leur histologie définitive sur la pièce opératoire : B2 (15 lésions ; 16,85%), B3 (63 lésions ; 70,79%) et B5 (11 lésions ; 12,36 %). Nous souhaitons une description plus précise des patientes dont l'histologie était finalement bénigne (B2) (patientes qui auraient pu ne pas être opérées) et des patientes qui finalement avaient un cancer (B5) et pour lesquelles la chirurgie était vraiment bénéfique. Les *tableaux 2 et 3* récapitulent les caractéristiques histologiques définitives dans ces deux groupes.

		Toutes les lésions (89)	
		Nombre	%
Chirurgie			
	CHUGA	72	81
	Hors CHUGA	17	19
B2 définitif		15	16,8
B3 définitif		63	70,8
B5 définitif		11	12,4
Sous type	CCIS	9	81,8
B5 définitif	Carcinome infiltrant	2	18,2
	Type non spécifique	1	50
	Lobulaire	1	50

Tableau 2. Caractéristiques histologiques définitives de la population B3 opérée

		Histologie définitive		p value
		B2	B5	
Nombre de lésions		15	11	
Densité mammaire				0,837
	A	0	0	
	B	8 (53,3)	6 (54,5)	
	C	6 (40)	4 (36,3)	
	D	1 (6,7)	1 (9,1)	
ACR BI RADS				0,645
	3	0	0	
	4A	4 (26,7)	1 (9,1)	
	4B	8 (53,3)	8 (72,7)	
	4C	3 (20)	2 (18,2)	
Distorsion architecturale				1
	Oui	1 (6,7)	0	
	Non	14 (93,3)	11 (100)	
Microcalcifications				1
	Oui	14 (93,3)	11	
	Non	1 (6,7)	0	
Distribution				1
	Régionale	1 (7,1)	2 (18,2)	
	Amas	11 (78,6)	7 (63,6)	
	Linéaire	1 (7,1)	1 (9,1)	
	Segmentaire	1 (7,1)	1 (9,1)	
Morphologie				0,89
	Amorphes	4 (26,7)	5 (45,4)	
	Grossières hétérogènes	1 (6,7)	0	
	Fines et polymorphes	9 (60)	6 (54,6)	
	Fines linéaires ou branchées	0	0	

Tableau 3. Caractéristiques radiologiques des patientes selon le type histologique définitif (B2 et B5)

Description des groupes B2 et B5 définitifs

Comme pour la population incluse, les caractéristiques radiologiques et histologiques étaient équivalentes à celle de S. Benhamiche [8], en dehors du type d'anomalies mammographiques.

Pour le groupe B2, il s'agissait uniquement de microcalcifications (93,3% contre 56% précédemment) dont une lésion associait microcalcifications et distorsion architecturale (6,7%). La majorité des lames chirurgicales retrouvait seulement des remaniements cicatriciels des biopsies antérieures.

Toutes les lésions B5 correspondaient à des microcalcifications (contre 80%). L'histologie définitive correspondait à un carcinome canalaire in situ (CCIS) dans 9 cas sur 11 et un carcinome infiltrant dans 2 cas (un de type non spécifique (CITNS), l'autre de type lobulaire (CLI)) (*tableau 2*).

De façon similaire, il n'a pas été mis en évidence de critères radiologiques significativement différents entre les groupes B2 et B5 (*tableau 3*).

2. Analyse du score Transpara®

a. Analyse descriptive du score régional

Sur les 89 lésions détectées par le radiologue, 79 (88,8%) ont été signalées par le logiciel Transpara®. Le score régional moyen de l'anomalie étudiée était de 64, soit un risque considéré comme « Intermédiaire » par le constructeur. Seul le score le plus élevé entre les deux incidences a été pris en compte pour l'analyse statistique. La moyenne des scores régionaux sur l'incidence de face était de 52 (Intermédiaire), identique à celle sur les clichés obliques.

Dix lésions avaient un score nul car non détectées (11,2%) (*figure 3*), 4 « Faible » (4,5%), 37 « Intermédiaire » (41,6%) et 38 « Élevé » (42,7%). À noter que 5 lésions ont été retrouvées seulement sur les images de tomosynthèse, non disponibles pour l'ensemble de la population et donc non étudiée.

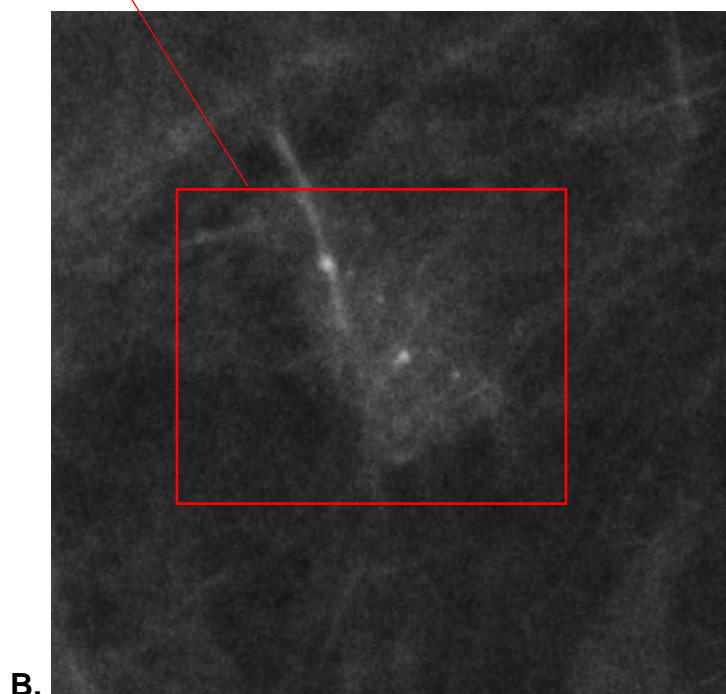
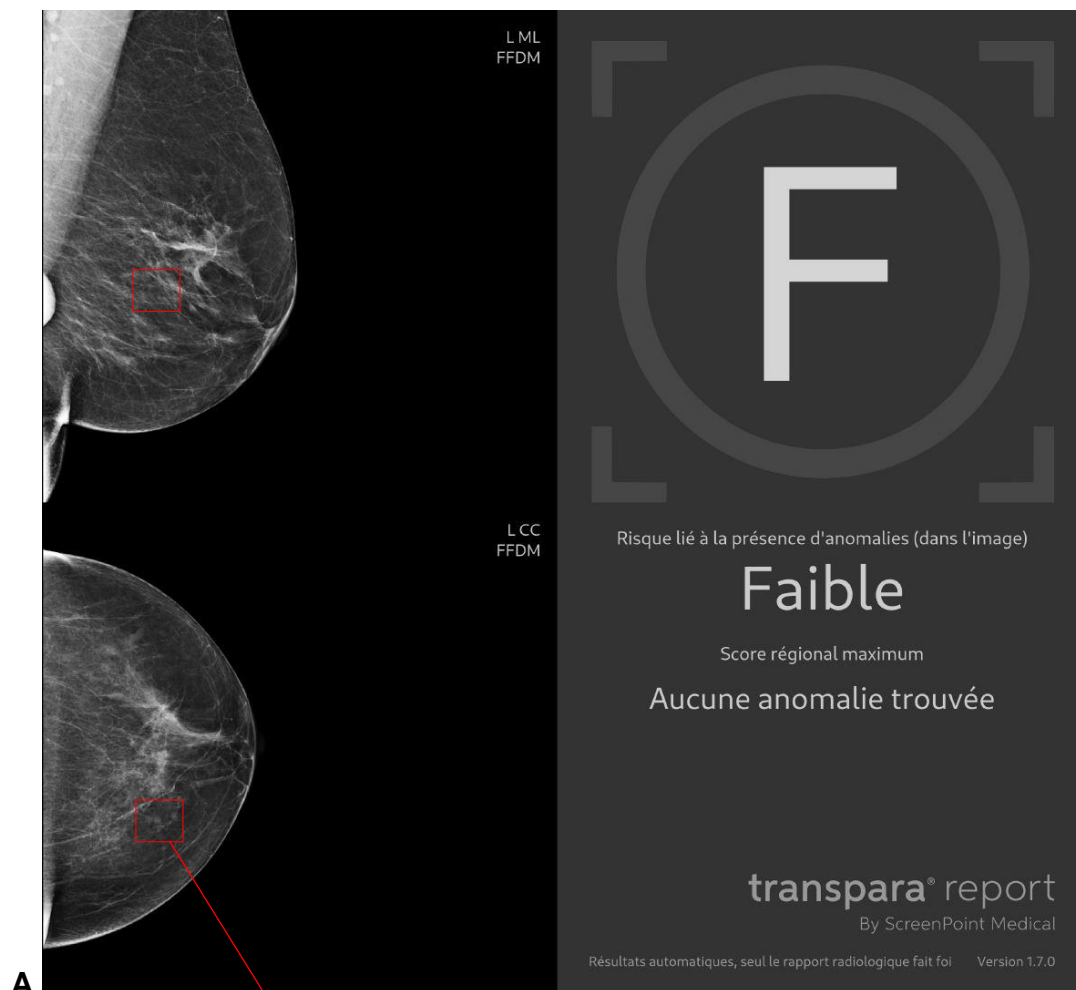


Figure 3. (A) Foyer de microcalcifications grossières hétérogènes de l'union des quadrants internes gauches (encadré rouge), non vu par Transpara (absence de score régional), correspondant à de l'hyperplasie canalaire atypique (B3) sur la biopsie et la pièce opératoire. **(B)** Agrandissement sur le foyer de microcalcifications ACR4A.

Les anomalies étaient pour plus de la moitié considérées par le modèle comme des calcifications (« calc ») (54,4%) et de façon moins fréquente par une double anomalie (« calc + mass ») (38%). Six anomalies étaient désignées comme des masses seules (« mass ») (7,6%) et correspondaient uniquement à des distorsions architecturales (figure 4).

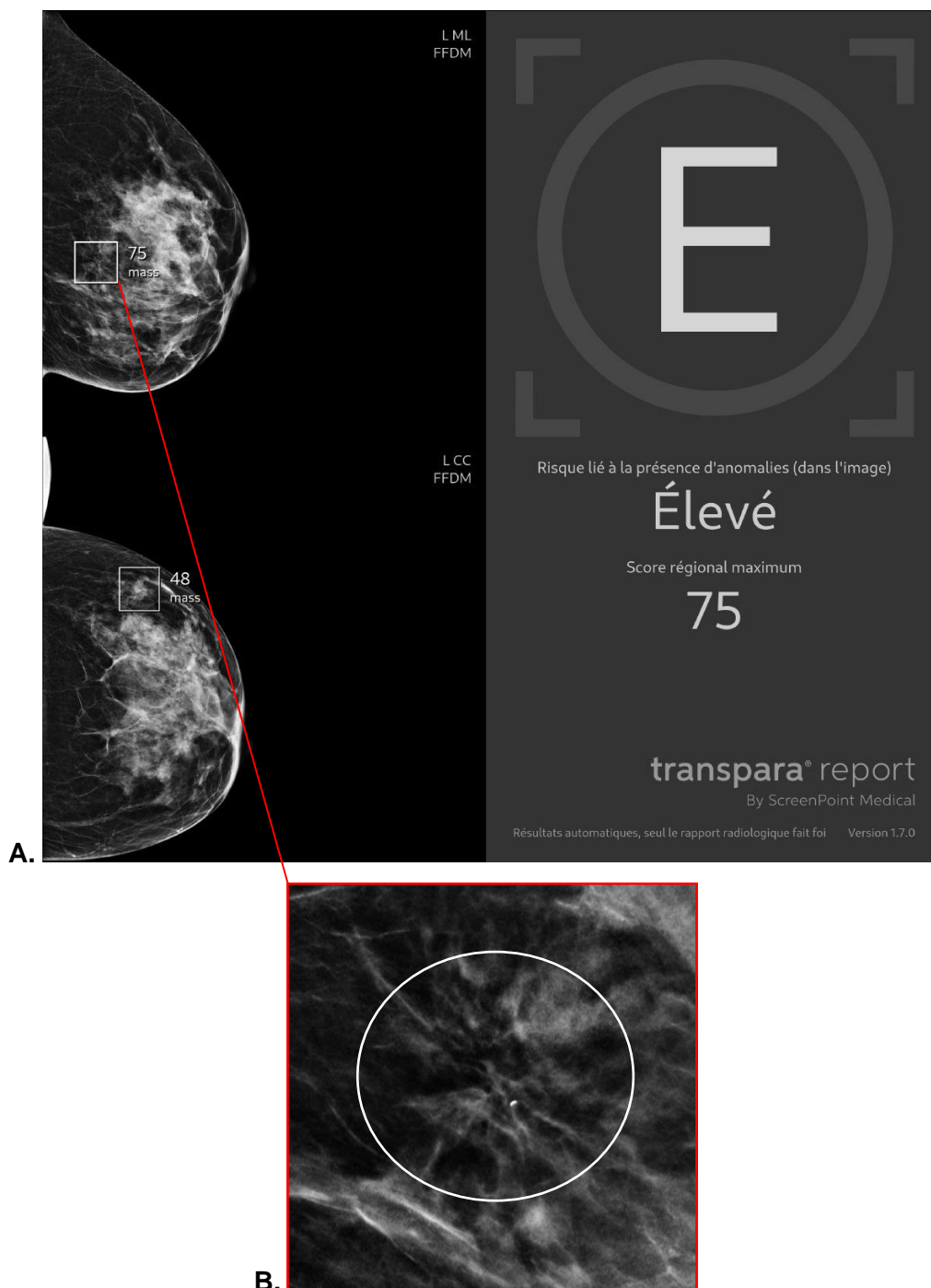


Figure 4. (A) Distorsion architecturale ACR4C de l’union des quadrants externes du sein gauche marquée par Transpara® avec l’annotation « mass » (encadré blanc). Score régional sur le cliché oblique externe à 75 et à 48 sur la face, avec un score global Elevé (correspondant au score régional le plus élevé). L’anatomopathologie définitive était une cicatrice radiaire (B3). **(B)** Agrandissement sur le cliché oblique centré sur la distorsion (cercle).

Les dix lésions non détectées étaient 9 foyers de microcalcifications (*figures 3 et 5*) et une distorsion architecturale. Elles correspondaient à 1 lésion B2 définitive (10%), 1 lésion B5 définitive (10%) (*figure 5*) et 8 lésions B3 définitives (80%). Comme précédemment nous souhaitions une analyse plus précise des groupes B2 et B5.

- **Descriptif du score Transpara® du groupe B2 (N=15) :**

Le score régional moyen était de 64,5. Les anomalies détectées correspondaient à 8 foyers de microcalcifications (57,1 %), 4 foyers de microcalcifications associés à une masse (28,6%) et 2 masses (14,3%) (toutes deux correspondant à une distorsion architecturale). Une lésion n'a pas été retrouvée par Transpara et correspondait à de l'hyperplasie canalaire simple. Ces anomalies étaient soit non retrouvées sur les tomosynthèses soit détectées mais avec un score régional plus bas.

- **Descriptif du score Transpara® du groupe B5 (N=11) :**

Le score régional moyen de 64,1 était statistiquement non différent du groupe B2 (p value 0,876, *tableau 5*). Les anomalies détectées correspondaient à 5 foyers de microcalcifications (50%) et 5 foyers de microcalcifications associés à une masse (50%). Une lésion (foyer de microcalcification ACR4B) signalée par le radiologue n'a pas été marquée par le modèle et correspondait à un carcinome invasif de type non spécifique (CITNS) (*figure 5*). Une des lésions avait un score un peu plus élevé avec la tomosynthèse (71 versus 61 sur le cliché 2D).

Les *tableaux 4 et 5* récapitulent les données du modèle sur notre population.

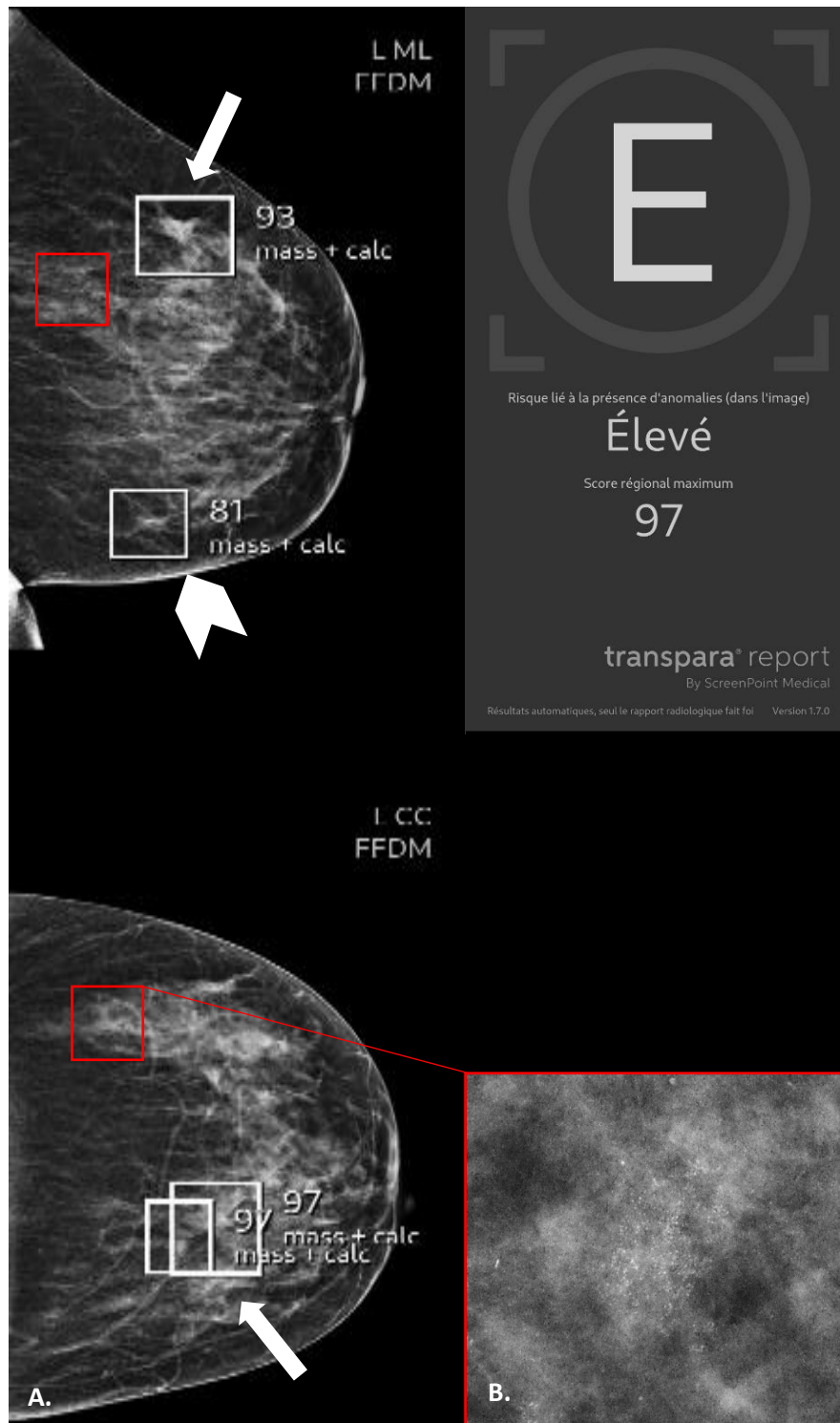


Figure 5. (A) Détection par Transpara® (et par le radiologue) d'une masse associée à des microcalcifications du quadrant supéro-interne gauche (flèche). Lésion doublement marquée sur la face (deux encadrés « mass+calc ») avec score régional à 97 associé à un score régional de 93 sur le cliché oblique. Le score global est Élevé (97). L'histologie définitive était un carcinome canalaire infiltrant (CITNS) avec une composante in situ (CCIS). Le foyer de microcalcification en amas ACR4B du quadrant supéro-externe (encadré rouge) n'a pas été marqué. Le score de 81 sur l'oblique (tête de flèche) était considéré comme de la superposition glandulaire par le radiologue. **(B)** Agrandissement sur les microcalcifications retenues seulement par le radiologue, correspondant à un CITNS (B5) sur la pièce opératoire.

	Total des lésions	
	Nombre	%
Score Régional Transpara®	89	
	Score Nul (non détecté)	10 11,2
	Score Faible	4 4,5
	Score Intermédiaire	37 41,6
	Score Élevé	38 42,7
Anomalie détectée (79 sur 89)		
	Calcifications	43 54,4
	Masse	6 7,6
	Masse + Calcifications	30 38
Histologie définitive des lésions non vues par Transpara®		
	B2	1 10
	B3	8 80
	B5	1 10

Tableau 4. Données de l'algorithme pour notre population B3 biopsiée

	Histologie définitive		P-value
	B2 (%)	B5 (%)	
Nombre total de lésions	15	11	
Lésions détectées par Transpara®	14 (93,3)	10 (90,9)	
Score moyen	64,5	64,1	0,876
Anomalies détectées			0,683
Microcalcifications	8 (57,1)	5 (50)	
Masse	2 (14,3)	0	
Masse + microcalcifications	4 (28,6)	5 (50)	
Lésions non détectées par Transpara®	1 (6,7)	1 (9)	

Tableau 5. Données de l'algorithme selon le type histologique définitif (B2 et B5).

b. Foyers additionnels (tableau 6)

L'algorithme a détecté 56 foyers (63% des patientes) en plus de celui étudié et à chaque fois un seul par patiente sauf pour une (deux foyers additionnels). Le score régional moyen était de 64 (« Intermédiaire »), équivalent au foyer principal étudié.

Seulement 48% (27/56) des lésions étaient retenues par le radiologue a priori.

Dix d'entre elles (37%) étaient des lésions B5 définitives (biopsiées dans le même temps ou à distance). Il s'agissait de 3 CITNS associés à des lésions de CCIS (figure 5), 2 CITNS et 5 CCIS seuls. Les 9 autres (34%) étaient en rapport avec la lésion principale étudiée doublement marquée (figure 6), systématiquement en lien avec des foyers de microcalcifications étendus (jusqu'à 50 mm). Par ailleurs, il y avait 6 lésions B2 (22,2%) dont un kyste, 1 lésion B3 (3,4%) et 1 foyer de calcifications ACR 3 (perdue de vue après la chirurgie de la lésion frontière étudiée).

Les 29 lésions (52%) non retenues étaient considérées par le radiologue comme de la superposition glandulaire (figure 7), des macro-calcifications, des microcalcifications éparses ou des calcifications vasculaires (ACR 2).

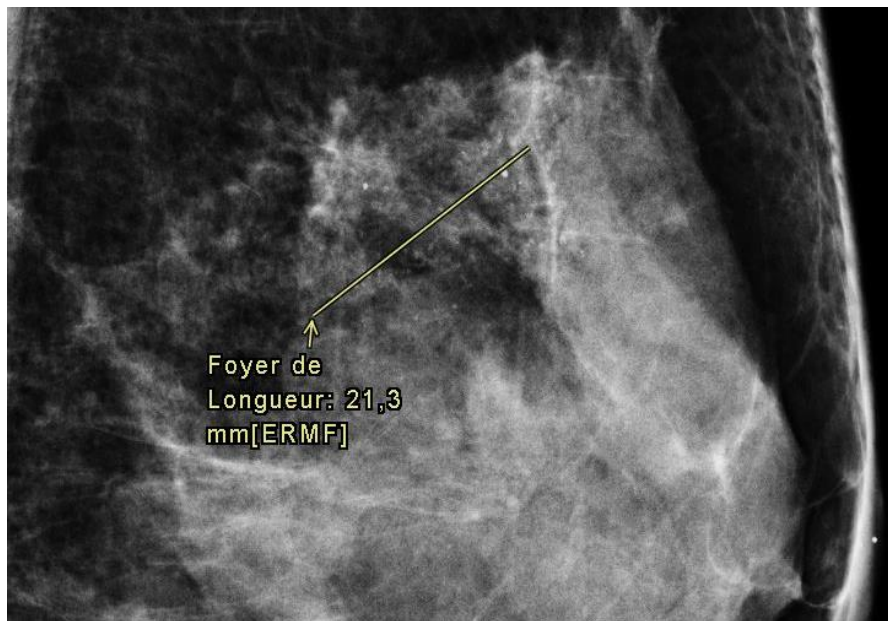
Deux foyers de microcalcifications ACR 4B (B2) détectés par le radiologue n'étaient pas vus par l'algorithme et une lésion vue par le modèle (score 81, élevé) et non vue initialement par le radiologue était un CITNS, vu en échographie dans un second temps (figure 8).

	Nombre	%
Foyers additionnels	56	
Retenus par le radiologue	27	48
Non retenus	29	52
Caractéristiques des foyers retenus		
Lésions B5 définitives	10	37
Lésion B3 définitive	1	3,4
Lésions B2 définitives	6	22,2
Microcalcifications ACR 3	1	3,4
Lésions principales doublements marquées (étendues)	9	34

Tableau 6. Foyers additionnels marqués par Transpara® sur le bilan mammographique avec les caractéristiques de ceux retenus par le radiologue.



A.



B.

Figure 6. (A) Double marquage (flèche) d'un foyer de microcalcifications sur le cliché oblique (66 et 73). Une seule marque (tête de flèche) sur le cliché de face (88). **(B)** Agrandissement sur le cliché oblique des microcalcifications régionales fines et polymorphes du quadrant supéro-externe gauche, ACR4B, étendues sur 21 mm, expliquant le double marquage.

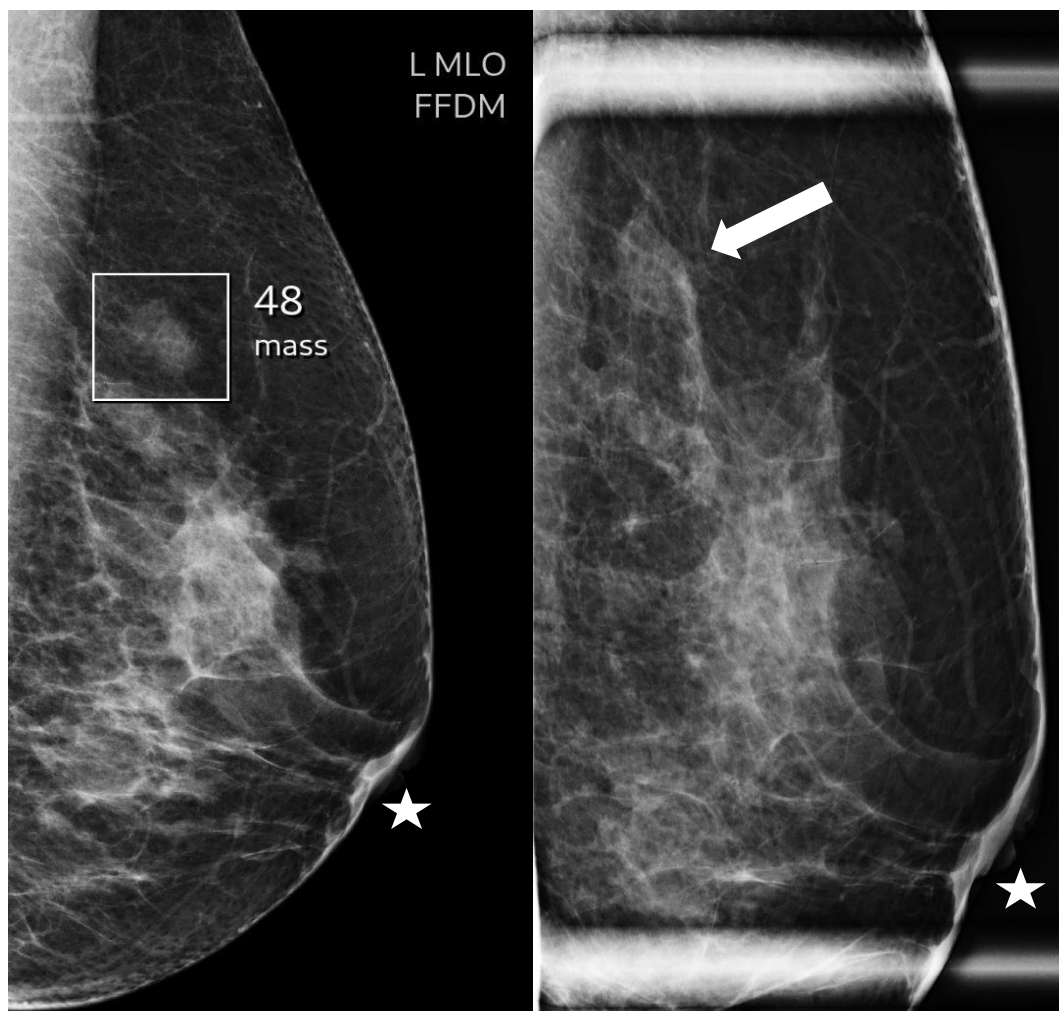


Figure 7. Détection d'une masse (score régional à 48) des quadrants supérieurs gauches sur le cliché oblique par Transpara®. Étalement de cette densité sur le cliché localisé comprimé correspondant à de la simple superposition glandulaire (flèche), non retenue par le radiologue. Mamelon (étoile).

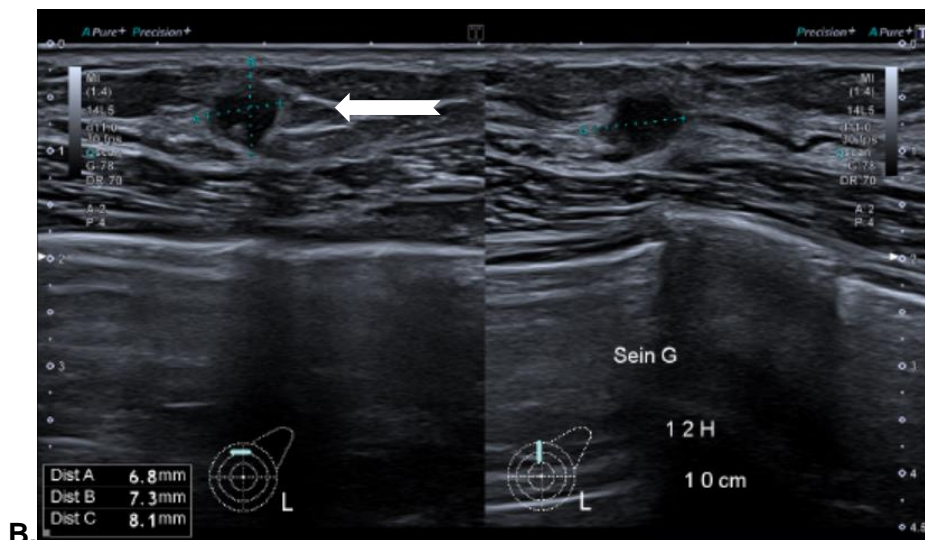
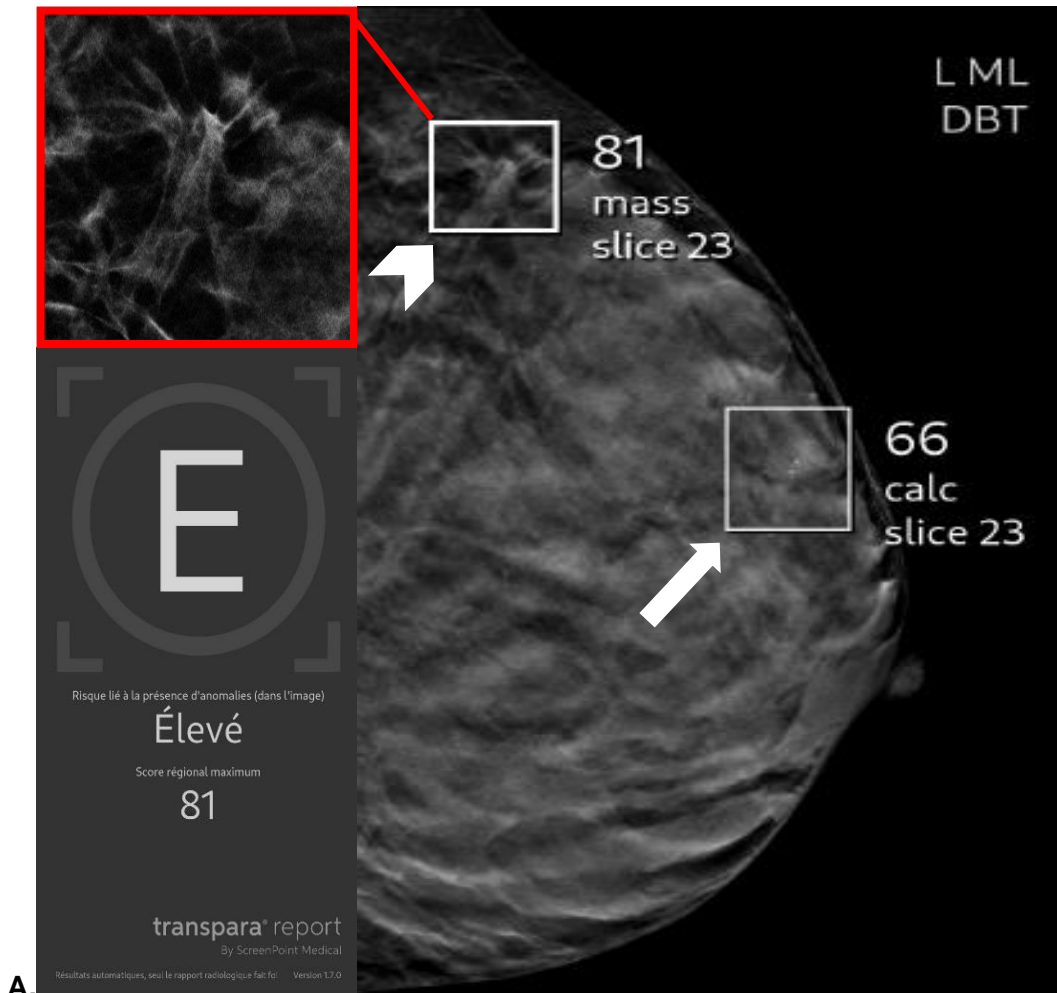


Figure 8. (A) Détection du foyer de microcalcifications ACR4B du quadrant supéro-interne (score à 66) (flèche) en lien avec la lésion principale étudiée (B3). Cible additionnelle (tête de flèche) de l'union des quadrants supérieurs, marquée comme étant une masse avec un score régional de 81 sur la coupe 23 (cliché de tomosynthèse). Agrandi sur cette masse spiculée (encadré rouge), détectée par Transpara® sur les clichés de tomosynthèse et non vue par le radiologue. **(B)** En échographie, masse infracentimétrique des quadrants supérieurs gauches correspondant à la masse initialement marquée par Transpara®. La pièce opératoire était un CITNS (B5).

c. Performance de Transpara® sur la population B2 et B5

La courbe ROC du modèle est montrée dans la *figure 9*. L'aire sous la courbe (AUC, area under the curve) était de 0,52, soit à peine plus que le hasard. Les meilleures performances du test étaient atteintes pour un score supérieur à 82 avec une sensibilité de 36% et une spécificité de 87%. La valeur prédictive négative (VPN) était de 0,65 et la valeur prédictive positive (VPP) de 0,667. Pour comparaison, la sensibilité moyenne d'un radiologue en mammographie est de 86,9% avec une spécificité de 88,9% (Lehman 2017).

Sur ces données descriptives, la VPP du modèle semblait dans l'intervalle de celle du score ACR 4C (VPP de cancer : 50-95%). Le modèle positivait 4/11 lésions B5 alors que le radiologue classait seulement 2 lésions B5 avec un score ACR4C. Les autres étaient classées en ACR 4A (VPP de cancer : 2%-10%) ou ACR 4B (VPP de cancer : 10-50%) et donc considérées comme moins suspectes par le radiologue. De la même façon, le modèle positivait 2/15 lésions B2 alors que le radiologue avait classé 3/15 lésions ACR 4C.

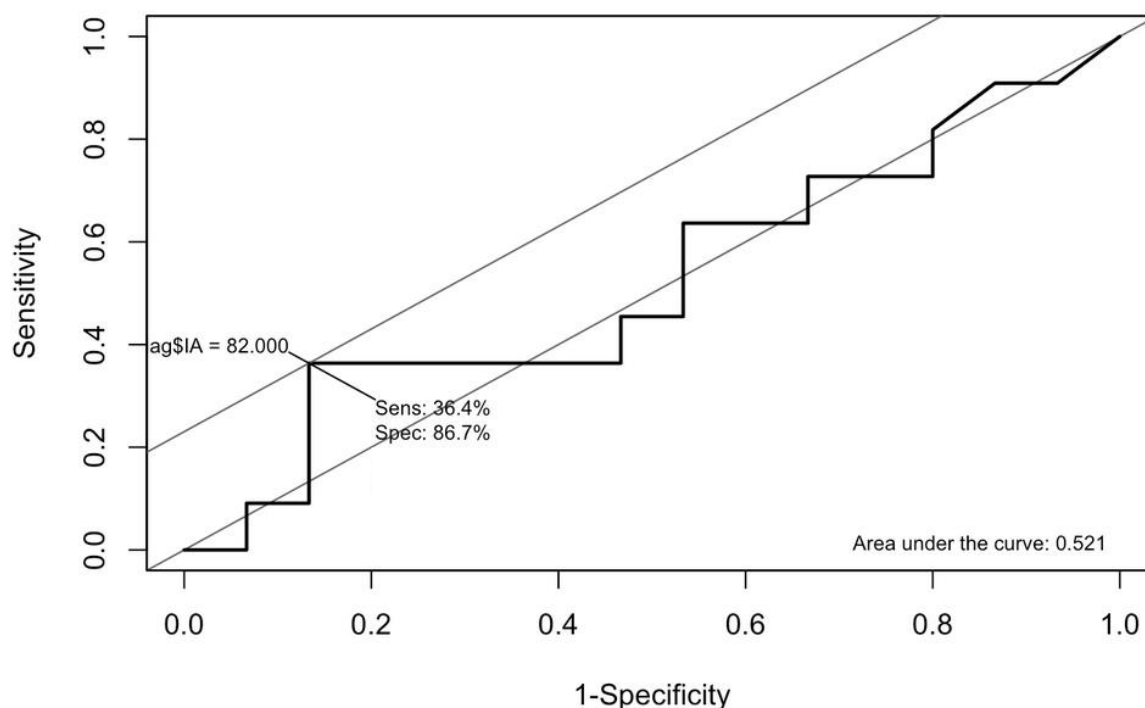


Figure 9. Performance et Aire sous la courbe du modèle Transpara® dans le groupe B2 et B5 de notre population.

d. Analyse descriptive de la distribution des scores du modèle sur les 81 foyers de microcalcifications de notre population B3 biopsiée (tableau 7)

- *Selon l'ACR :*

Les scores de l'algorithme étaient statistiquement différents selon le score ACR et, au vu des données descriptives, semblaient croissants selon la gravité du score BIRADS : ACR 3 < ACR 4A < ACR4B < ACR4C (p=0,004).

- *Selon les caractéristiques des microcalcifications*

Les scores du modèle n'étaient pas significativement différents selon le type de distribution des 81 foyers de calcifications (p = 0,052).

Concernant la morphologie, il existait une différence statistiquement significative (p = 0,006). Aux vues des données descriptives, les scores de l'IA étaient croissants dans le sens Amorphes < Fines et polymorphes < Grossières Hétérogènes, et donc différents du score prédictif BI RADS actuellement validé (Amorphes < Grossières Hétérogènes < Fines et Polymorphes < Fines Linéaires et/ou Branchées). La catégorie de microcalcifications de morphologie fine linéaire n'a pas été pris en compte dans l'analyse statistique étant donné qu'il y avait qu'une seule valeur.

	Lésions B3	Score IA ¹	p value ²
ACR BI RADS			0,004
	3	26	
	4A	57	
	4B	68	
	4C	77	
Distribution			0,052
	Régionale	69	
	Amas	62	
	Linéaire	64	
	Segmentaire	88	
Morphologie			0,006
	Amorphes	54	
	Grossières hétérogènes	77	
	Fines et polymorphes	65	

Tableau 7. Données de l'algorithme selon la classification BIRADS de l'ACR, la morphologie et la distribution des microcalcifications de notre population B3 biopsiée (81).

¹ Moyenne, ² Kruskal-Wallis rank sum test

e. Analyse descriptive du score Transpara® pour les masses du groupe B5

Sur 11 lésions B5, 5 étaient considérées par le modèle comme des masses associées à des calcifications (« mass + calc »), toutes en lien avec des lésions de CCIS. Les deux lésions infiltrantes étaient soit non vues pour une (*figure 5*), soit considérées comme étant des microcalcifications seules pour l'autre avec un score intermédiaire (*figure 10*). L'ensemble des lésions B5 était considéré comme des microcalcifications seules par le radiologue.

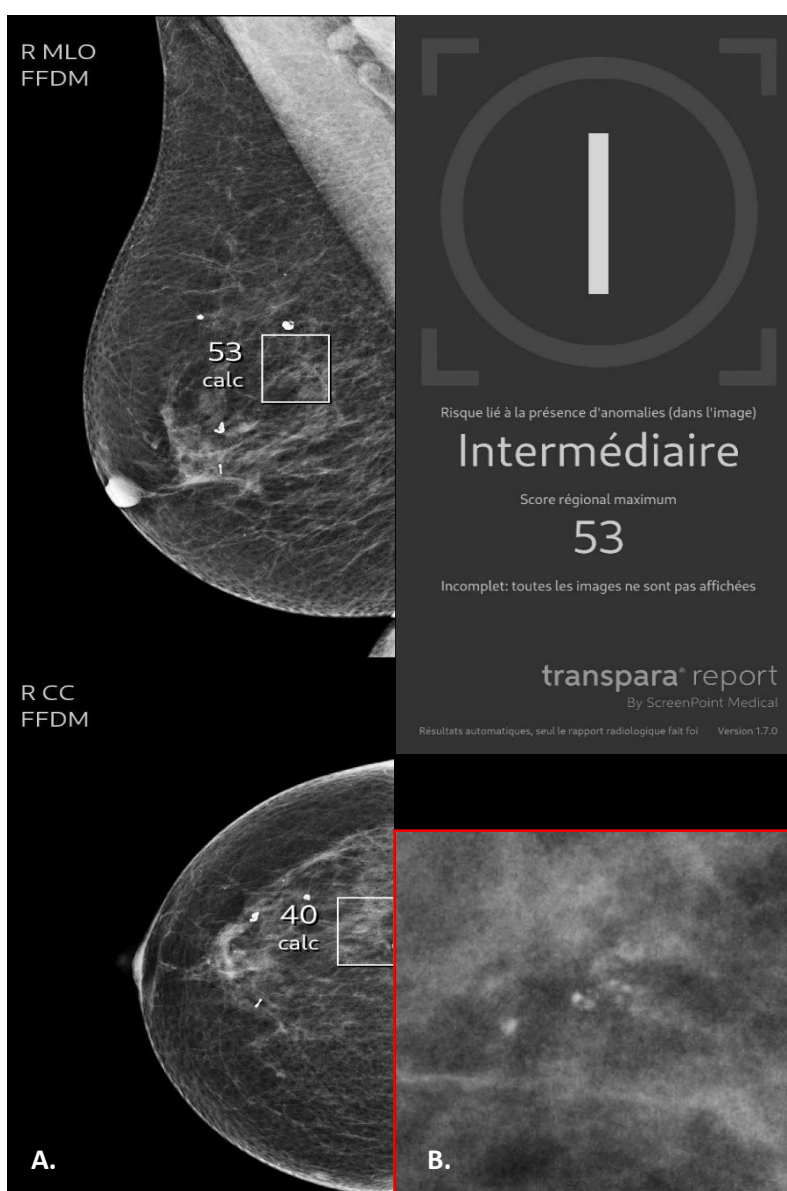


Figure 10. (A) Détection de calcifications (« calc ») en amas et amorphes ACR4B dans le quadrant supéro-externe droit avec score global intermédiaire (score régional de 40 sur la face et de 53 sur l'oblique). **(B)** Agrandissement sur le cliché de face (encadré rouge). L'anomalie correspondait à un carcinome lobulaire infiltrant (B5b) sur la pièce opératoire.

D) Discussion

1. Description de la population B3

Les lésions frontières ont été retrouvées sur 17,84 % (165/925) de l'ensemble des macrobiopsies mammaires sous mammographie réalisées en 6 ans au CHUGA. Ce pourcentage est concordant avec les données de la littérature, qui les situent entre 4,5% et 22,9% [14]. Quatre-vingt-onze pourcent des anomalies macrobiopsiées sous mammographie étaient des microcalcifications, résultats également similaires à la littérature (36,3% à 98,7% avec une médiane de 85,7%) [14].

De 2016 à 2021 au CHUGA, 63,6% des lésions B3 étaient opérées après discussion en RCP, également en accord avec les données disponibles dans la littérature malgré leur hétérogénéité (de 33,5% pour *Rageth* à 86,3% pour *Cohen et al.*) [3,9] ce qui montre l'absence consensuelle de prise en charge thérapeutique à ce jour.

2. Analyse du score Transpara®

Le modèle a été appliqué à posteriori sur les mammographies réalisées entre novembre 2016 et juin 2021 pour 79 lésions frontières. Le score Transpara n'était pas disponible initialement et le classement ACR de chaque lésion a été déterminé par le radiologue de façon indépendante, limitant le biais d'incorporation. Pour les 10 autres lésions détectées sur des mammographies réalisées après juin 2021, le modèle émettait un score instantané sur la console d'interprétation, ce qui pouvait influencer le radiologue dans sa classification BIRADS.

L'absence de performance satisfaisante du modèle dans notre population B3 tend vers les résultats de deux grandes études. *Fenton et al.* a montré une diminution de la performance dans l'interprétation des mammographies par des logiciels d'IA par rapport aux médecins, une augmentation du taux de biopsies et l'absence d'association dans l'amélioration de la détection des cancers invasifs [7]. *Lehman et al.* a quant à lui estimé que les performances des CAD n'étaient pas meilleures que le radiologue et qu'elles diminuaient la sensibilité du radiologue lorsqu'il était aidé par ces modèles [15]. Cependant, ces études analysaient en grandes parties des CAD d'anciennes générations dérivées du machine learning à la différence du logiciel Transpara® utilisé pour notre étude.

Dès lors, notre analyse est en grande partie discordante avec les données récentes de la littérature faisant intervenir des modèles de DL.

Dans l'ensemble de ces études, *Rodriguez-Ruiz et al* a démontré la non-infériorité ou la supériorité du modèle Transpara® seul pour la détection des cancers (AUC du modèle à 0,840 vs 0,814 pour les médecins, avec un AUC du modèle supérieur à 61,4% des 101 radiologues) ou en association avec le radiologue (AUC 0,89 avec l'aide d'IA vs 0,87 sans aide) ainsi qu'une diminution du flux de travail de 20% sans perte de chance pour les patientes [11-13].

Kim et al. a prouvé également l'importante amélioration de la détection des lésions infiltrantes par les algorithmes par rapport aux radiologues (sensibilité de 90,24% contre 75,55%) [16]. Dans notre étude, les deux lésions infiltrantes étaient soit non vue pour l'une, soit marquée comme des microcalcifications seules pour l'autre avec un score intermédiaire (53).

Une des principales explications est que notre population diffère de façon majeure à celles incluses dans les études actuelles. Comme l'a montré *Freeman et al.* dans sa revue systématique [17], ces études sont pour la plupart rétrospectives et sans analyse en routine clinique (analyse en laboratoire avec biais de sélection et d'incorporation). Elles se basent sur des données internes sans qu'il y ait de validation externe satisfaisante, notamment à travers la validation croisée, ce qui surestime les performances (overfitting) et entraîne un problème de généralisation de ces tests [18]. Les inclusions sont par ailleurs toujours réalisées dans le cadre du dépistage avec des patientes asymptomatiques cliniquement et sans antécédents personnels de cancer du sein ou de lésions frontalières. Les populations sont soit enrichies en cancer (jusqu'à 39,7% avec le modèle Transpara par *Rodriguez-Ruiz* [11]), soit avec une prévalence proche de la population générale (0,7%).

Notre population au CHUGA est très différente. Sur 89 lésions, 19 patientes (21,34%) n'étaient pas dans le cadre du dépistage. Il y avait 3 antécédents de lésions B3 opérées (1 HLA, 1 MCA et 1 CR + CLIS), 2 antécédents de cancer, 6 lésions ACR 6 homolatérales (3 CCIS + CITNS, 2 CITNS et 1 CCIS) et 8 lésions ACR 6 controlatérales (2 CLI, 3 CITNS, 2 CCIS, 1 CCI + CCIS). Dans les groupes B2 et B5, où les performances du modèle ont été analysées, il y avait 6/26 patientes (23%) en dehors du cadre du dépistage.

Une des autres principales explications est le manque de puissance.

Avec 15 lésions B2, 11 lésions B5 et une puissance consentie de 80%, on pouvait mettre en évidence une AUC à 0,79. Pour mettre en évidence une AUC de 0,8 avec une puissance de 90%, il aurait été nécessaire d'avoir 17 lésions dans chaque groupe.

Cinq lésions ont été détectées seulement sur les clichés de tomosynthèse (score de 39 à 72, non pris en compte) dont 3 correspondaient à une distorsion architecturale. Cela est concordant avec les données de la littérature où la performance des modèles seuls ou du radiologue aidé par les modèles semble meilleure sur les tomosynthèses par rapport aux clichés 2D, notamment pour les anomalies non calciques (distorsion, masse et asymétrie) [19]. Cependant, ces 5 anomalies étaient des lésions B3 définitives, ce qui n'avait pas d'impact dans notre étude.

Concernant la distribution du score de l'IA selon l'ACR dans les foyers de microcalcifications ; il semblerait que compte tenu des données descriptives, il soit croissant selon la gravité du BIRADS. Pour s'en assurer, il faudrait comparer nos moyennes deux à deux en prenant en compte le risque d'inflation du risque alpha, ce qui n'a pas été fait.

3. Limites de l'étude

Les lésions n'étaient pas toutes indépendantes car nous avons 89 lésions chez 86 patientes.

Le manque de puissance, précédemment décrit, était une des principales limites de notre étude.

Bien que les performances de Transpara® dans notre population étaient médiocres, il n'y a pas eu de comparaison avec celles des radiologues du CHUGA. L'interprétation médicale est actuellement basée sur le score BIRADS de l'ACR, une variable non continue et une échelle non ordinale, difficilement analysable pour la comparaison de courbe ROC [20]. En conséquence, il n'existe pas à ce jour d'études comparant les performances des modèles avec l'intégralité du score BIRADS [17]. Pour une partie des analyses comparant les performances des modèles à celles des radiologues, notamment Rodriguez-Ruiz avec Transpara®, l'ensemble des lésions classées ACR 3 (VPP de cancer à 2%) à ACR 5, voir ACR 2 (VPP de cancer nul), étaient considérées comme positives (cancéreuses). Seules les lésions ACR 1 plus ou moins ACR 2 étaient considérées comme indemnes de cancer. L'autre principale alternative est l'utilisation de probabilité de malignité (probability of malignancy , POM) émise par les médecins [21], sans lien avec le lexique BIRADS validé. Comme le souligne Freeman, l'hétérogénéité méthodologique de ces études comporte un important risque de biais et pose problème quant à l'applicabilité de ces modèles en routine clinique.

Comme pour la majorité des études actuelles, notre population sur laquelle les performances de Transpara® ont été étudiées (B2 et B5) n'était pas représentative de la population en routine clinique. Les lésions frontières définitives (B3) ont été exclues de l'analyse alors qu'elles représentaient la majorité de l'ensemble des anomalies (70,8%). Cependant, l'analyse statistique de ces lésions est hasardeuse compte tenu de leur hétérogénéité et du risque néoplasique non binaire à la différence des lésions B5 (cancer) et B2 (bénigne).

Huit lésions ont été exclues de l'étude car non opérées en raison d'une exérèse complète de la lésion sur la macrobiopsie diagnostique. Nous aurions pu inclure ces patientes et considérer les résultats comme des résultats post opératoires et étudier le score Transpara associé. Cela aurait pu augmenter notre nombre de patientes B3 de 89 à 97. Le reste des lésions (52/60) était considéré comme ayant un faible risque néoplasique en RCP, et donc non opérées. Cela peut constituer un risque de biais de sélection, car le risque de sous-estimation persiste même s'il est moins important que pour les lésions dont l'indication opératoire est posée en RCP. Un suivi de 10 ans aurait été nécessaire pour analyser d'éventuels cancers apparus ou de lésions opérées dans cet intervalle, puis intégrer les scores Transpara® à notre population.

Dix-sept patientes ont été opérées hors CHUGA. Les lames post-opératoires ont donc été lues par des anatomopathologistes différents, ce qui est à la fois une limite (reproductibilité inter-observateur parfois moyenne, par exemple de 0,27 à 0,69 [19] pour l'HCA) et une force car de cette manière, nous nous sommes affranchis du biais de lecteur.

3. Conclusion

Le modèle Transpara® semble non adapté comme aide supplémentaire à la décision thérapeutique des lésions frontières. Ses performances s'avèrent médiocres pour la prédiction de cancer en mammographie dans une population différente (lésions frontières) de celle entraînée (dépistage), malgré l'absence de comparaison avec les radiologues. Les preuves actuelles sur l'utilisation des systèmes d'IA dans le dépistage du cancer du sein sont loin d'avoir la qualité requise pour leur mise en œuvre dans la routine clinique et encore moins dans des populations différentes de celle du dépistage.

Les lésions frontières restent une catégorie très hétérogène d'anomalies mammaires et difficile à appréhender, nécessitant une expertise médicale spécifique. En l'attente de nouvelles recommandations claires et d'éventuelles études sur l'IA en dehors du dépistage, les RCP radio-clinico-histologiques restent cruciales dans la décision thérapeutique des lésions frontières.

THÈSE SOUTENUE PAR : Kévin GUTTIN

TITRE :

PERFORMANCE DIAGNOSTIQUE D'UN MODÈLE DE DEEP LEARNING POUR LA PREDICTION DE CANCER DU SEIN EN MAMMOGRAPHIE : EXPERIENCE DU SYSTÈME TRANSPARA® SUR 89 LÉSIONS FRONTIÈRES (B3) AU CHU GRENOBLE ALPES

CONCLUSION :

Introduction : L'amélioration des techniques d'imageries pour le dépistage du cancer du sein est responsable depuis plusieurs années d'une augmentation du nombre de biopsies mammaires et de la découverte de lésions frontières (B3). Elles représentent environ 7 à 9% des biopsies et sont un groupe hétérogène d'anomalies histologiques marqueurs d'un tissu mammaire à risque de cancer du sein. Leur prise en charge a longtemps reposé sur la chirurgie devant leur risque de sous-estimation néoplasique de 9.9% à 35.1%. Si certains critères cliniques, radiologiques ou histologiques sont connus pour majorer le risque de sous-estimation de ces lésions frontières, il n'existe pas à l'heure actuelle de recommandation officielle sur leur prise en charge qui reste pays et centre dépendant. Notre objectif est d'étudier les performances d'un modèle d'intelligence artificielle (Transpara®) pour la prédiction de malignité des pièces opératoires d'anomalies mammographiques macrobiopsiées et correspondant à des lésions frontières B3.

Méthodes : Cette étude rétrospective monocentrique a inclus 89 lésions frontières macrobiopsiées correspondant à des anomalies mammographies chez 86 patientes, de mars 2016 à février 2022 au CHUGA. L'ensemble des patientes a été opéré avec des résultats histologiques définitifs (11 cancers, 15 lésions bénignes et 63 lésions frontières). Le modèle Transpara® a été appliqué en grande partie à posteriori (79 lésions) et à priori (10 lésions) sur les mammographies 2D. Pour chaque anomalie mammographique, le modèle attribuait un score de risque de cancer de 0 à 100 (plus le score est élevé, plus le risque est important). Les performances de l'algorithme ont été analysées sur les lésions B2 définitives (absence de prise en charge) et des lésions B5 définitives (chirurgie complémentaire). Elles ont été illustrées sous la forme d'une courbe ROC et d'une aire sous la courbe.

Résultats : Les performances du modèle avaient une aire sous la courbe à 0,521 avec une sensibilité à 0,36 et une spécificité à 0,87. Les scores de l'IA étaient croissants dans le sens du BIRADS ($p=0,004$) alors qu'il n'existait pas de différences statistiques selon la morphologie ou la distribution des microcalcifications ($p=0,052$). Les cancers infiltrants étaient soit non vu par le modèle (1 carcinome infiltrant de type non spécifique), soit vu avec un score intermédiaire (1 carcinome lobulaire infiltrant).

Conclusion : Le modèle Transpara® semble non performant dans notre population constituée uniquement de lésions frontières dont certaines patientes sont en dehors du cadre du dépistage. Ces résultats sont en marge des études actuelles qui se basent uniquement sur des mammographies de dépistage.

VU ET PERMIS D'IMPRIMER

Grenoble, le : 05/09/22

LE DOYEN DE LA FACULTÉ DE MÉDECINE


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

LE PRÉSIDENT / LA PRÉSIDENTE DU JURY


Pr. Gilbert FERRETTI

RÉFÉRENCES Partie 1

- [1] « Cancer du sein ». <https://www.santepubliquefrance.fr/maladies-et-traumatismes/cancers/cancer-du-sein> (consulté le 27 décembre 2021).
- [2] « Le programme de dépistage organisé - Dépistage du cancer du sein ». <https://www.e-cancer.fr/Professionnels-de-sante/Depistage-et-detection-precoce/Depistage-du-cancer-du-sein/Le-programme-de-depistage-organise> (consulté le 27 décembre 2021).
- [3] N. Anderson *et al.*, « Guidelines for non-operative diagnostic procedures and reporting in breast cancer screening (February 2017), Royal College of Pathologists Publication G150. », févr. 2017.
- [4] C. Cullinane, J. Byrne, L. Kelly, M. O. Sullivan, M. A. Corrigan, et H. P. Redmond, « The positive predictive value of vacuum assisted biopsy (VAB) in predicting final histological diagnosis for breast lesions of uncertain malignancy (B3 lesions): A systematic review & meta-analysis », *Eur. J. Surg. Oncol.*, vol. 0, n° 0, avr. 2022, doi: 10.1016/j.ejso.2022.04.005.
- [5] P. Orsaria *et al.*, « Heterogeneous risk profiles among B3 breast lesions of uncertain malignant potential », *Tumori J.*, vol. 106, n° 2, p. 115-125, avr. 2020, doi: 10.1177/0300891619868301.
- [6] C. J. Rageth *et al.*, « Second International Consensus Conference on lesions of uncertain malignant potential in the breast (B3 lesions) », *Breast Cancer Res. Treat.*, vol. 174, n° 2, p. 279-296, avr. 2019, doi: 10.1007/s10549-018-05071-1.
- [7] S. E. Pinder *et al.*, « NHS Breast Screening multidisciplinary working group guidelines for the diagnosis and management of breast lesions of uncertain malignant potential on core biopsy (B3 lesions) », *Clin. Radiol.*, vol. 73, n° 8, p. 682-692, août 2018, doi: 10.1016/j.crad.2018.04.004.
- [8] « Macrobiopsie mammaire par aspiration note technique préparée pour l'AETMIS / par Wilber Deck | BANQ numérique ». <https://numerique.banq.qc.ca/patrimoine/details/52327/56620> (consulté le 22 juin 2022).
- [9] S. Benhamiche, « Lésions frontières B3 sur biopsies mammaires : étude descriptive monocentrique d'une population de 146 patientes opérées entre 2015 et 2021 avec création d'un ECRF ».
- [10] M. Cohen *et al.*, « When to do surgical resection for atypical breast lesions: Results of a prospective cohort of 518 lesions », *Surg. Oncol.*, vol. 34, p. 80-85, sept. 2020, doi: 10.1016/j.suronc.2020.04.002.
- [11] É. Larousse, « intelligence artificielle - LAROUSSE ». https://www.larousse.fr/encyclopedie/divers/intelligence_artificielle/187257 (consulté le 28 décembre 2021).
- [12] A. M. Turing, « I.—COMPUTING MACHINERY AND INTELLIGENCE », *Mind*, vol. LIX, n° 236, p. 433-460, oct. 1950, doi: 10.1093/mind/LIX.236.433.
- [13] « Intelligence artificielle, machine learning, deep learning : kézako ? », *L:EDigitalab*, 2 octobre 2017. <https://www.ledigitalab.com/2017/10/02/intelligence-artificielle-machine-learning-deep-learning-kezako/> (consulté le 28 décembre 2021).
- [14] « Les réseaux de neurones pour l'apprentissage profond : comment cela fonctionne ? », *Société Française de Radiologie*. <http://lamediatheque.radiologie.fr/mediatheque/media.aspx?mediaId=73650&playlistId=73654&channel=3277> (consulté le 28 décembre 2021).
- [15] T. W. Freer et M. J. Ulissey, « Screening Mammography with Computer-aided Detection: Prospective Study of 12,860 Patients in a Community Breast Center », *Radiology*, vol. 220, n° 3, p. 781-786, sept. 2001, doi: 10.1148/radiol.2203001282.
- [16] « ImageNet Large Scale Visual Recognition Competition 2012 (ILSVRC2012) ». <https://image-net.org/challenges/LSVRC/2012/results.html> (consulté le 28 décembre 2021).
- [17] B. A. Jonsson *et al.*, « Brain age prediction using deep learning uncovers associated sequence variants », *Nat. Commun.*, vol. 10, n° 1, Art. n° 1, nov. 2019, doi: 10.1038/s41467-019-13163-9.
- [18] A. Guermazi *et al.*, « Improving Radiographic Fracture Recognition Performance and Efficiency Using Artificial Intelligence », *Radiology*, vol. 302, n° 3, p. 627-636, mars 2022, doi: 10.1148/radiol.210937.

- [19] G. Chartrand *et al.*, « Deep Learning: A Primer for Radiologists », *RadioGraphics*, vol. 37, n° 7, p. 2113-2131, nov. 2017, doi: 10.1148/rg.2017170077.
- [20] R. A. Castellino, « Computer aided detection (CAD): an overview », *Cancer Imaging*, vol. 5, n° 1, p. 17-19, 2005, doi: 10.1102/1470-7330.2005.0018.
- [21] I. Thomassin-Naggara *et al.*, « Artificial intelligence and breast screening: French Radiology Community position paper », *Diagn. Interv. Imaging*, vol. 100, n° 10, p. 553-566, oct. 2019, doi: 10.1016/j.diii.2019.08.005.
- [22] « European Commission Initiative on Breast and Colorectal Cancer ». <https://healthcare-quality.jrc.ec.europa.eu> (consulté le 20 juillet 2022).
- [23] « Digital Mammography versus Digital Mammography Plus Tomosynthesis in Breast Cancer Screening: The Oslo Tomosynthesis Screening Trial | Radiology ». <https://pubs.rsna.org/doi/10.1148/radiol.2019182394> (consulté le 26 juin 2022).
- [24] « Breast cancer screening using synthesized two-dimensional mammography: A systematic review and meta-analysis | Elsevier Enhanced Reader ».
- [25] « Identifying Women with Dense Breasts at High Risk of Interval Cancers - PMC ». <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4443857/> (consulté le 25 juin 2022).
- [26] B. L. Sprague *et al.*, « Variation in mammographic breast density assessments among radiologists in clinical practice: Findings from a multicenter observational study », *Ann. Intern. Med.*, vol. 165, n° 7, p. 457-464, oct. 2016, doi: 10.7326/M15-2934.
- [27] A. Ciritsis *et al.*, « Determination of mammographic breast density using a deep convolutional neural network », *Br. J. Radiol.*, vol. 92, n° 1093, p. 20180691, janv. 2019, doi: 10.1259/bjr.20180691.
- [28] K. Kerlikowske, « Performance of Screening Mammography among Women with and without a First-Degree Relative with Breast Cancer », *Ann. Intern. Med.*, vol. 133, n° 11, p. 855, déc. 2000, doi: 10.7326/0003-4819-133-11-200012050-00009.
- [29] C. D. Lehman, R. D. Wellman, D. S. M. Buist, K. Kerlikowske, A. N. A. Tosteson, et D. L. Miglioretti, « Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection », *JAMA Intern. Med.*, vol. 175, n° 11, p. 1828, nov. 2015, doi: 10.1001/jamainternmed.2015.5231.
- [30] F. J. Gilbert *et al.*, « Single Reading with Computer-Aided Detection for Screening Mammography », *N. Engl. J. Med.*, vol. 359, n° 16, p. 1675-1684, oct. 2008, doi: 10.1056/NEJMoa0803545.
- [31] D. M. Ikeda, R. L. Birdwell, K. F. O'Shaughnessy, E. A. Sickles, et R. J. Brenner, « Computer-aided Detection Output on 172 Subtle Findings on Normal Mammograms Previously Obtained in Women with Breast Cancer Detected at Follow-Up Screening Mammography », *Radiology*, vol. 230, n° 3, p. 811-819, mars 2004, doi: 10.1148/radiol.2303030254.
- [32] R. K. Samala, H.-P. Chan, L. M. Hadjiiski, et M. A. Helvie, « Analysis of computer-aided detection techniques and signal characteristics for clustered microcalcifications on digital mammography and digital breast tomosynthesis », *Phys. Med. Biol.*, vol. 61, n° 19, p. 7092-7112, sept. 2016, doi: 10.1088/0031-9155/61/19/7092.
- [33] J. J. Fenton *et al.*, « Effectiveness of Computer-Aided Detection in Community Mammography Practice », *JNCI J. Natl. Cancer Inst.*, vol. 103, n° 15, p. 1152-1161, août 2011, doi: 10.1093/jnci/djr206.
- [34] A. Rodriguez-Ruiz *et al.*, « Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists », *JNCI J. Natl. Cancer Inst.*, vol. 111, n° 9, p. 916-922, sept. 2019, doi: 10.1093/jnci/djy222.
- [35] K. Freeman *et al.*, « Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy », *The BMJ*, vol. 374, p. n1872, sept. 2021, doi: 10.1136/bmj.n1872.
- [36] A. Pfof *et al.*, « The importance of multi-modal imaging and clinical information for humans and AI-based algorithms to classify breast masses (INSPIRED 003): an international, multicenter analysis », *Eur. Radiol.*, vol. 32, n° 6, p. 4101-4115, 2022, doi: 10.1007/s00330-021-08519-z.
- [37] T. Schaffter *et al.*, « Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms », *JAMA Netw. Open*, vol. 3, n° 3, p. e200265, mars 2020, doi: 10.1001/jamanetworkopen.2020.0265.

- [38] A. Rodríguez-Ruiz *et al.*, « Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System », *Radiology*, vol. 290, n° 2, p. 305-314, févr. 2019, doi: 10.1148/radiol.2018181371.
- [39] A. Rodríguez-Ruiz *et al.*, « Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study », *Eur. Radiol.*, vol. 29, n° 9, p. 4825-4832, sept. 2019, doi: 10.1007/s00330-019-06186-9.
- [40] J. L. Raya-Povedano, S. Romero-Martín, E. Elías-Cabot, A. Gubern-Mérida, A. Rodríguez-Ruiz, et M. Álvarez-Benito, « AI-based Strategies to Reduce Workload in Breast Cancer Screening with Mammography and Tomosynthesis: A Retrospective Evaluation », *Radiology*, vol. 300, n° 1, p. 57-65, juill. 2021, doi: 10.1148/radiol.2021203555.
- [41] A. Graewingholt et S. Duffy, « Retrospective comparison between single reading plus an artificial intelligence algorithm and two-view digital tomosynthesis with double reading in breast screening », *J. Med. Screen.*, vol. 28, n° 3, p. 365-368, sept. 2021, doi: 10.1177/0969141320984198.
- [42] K. Lång, S. Hofvind, A. Rodríguez-Ruiz, et I. Andersson, « Can artificial intelligence reduce the interval cancer rate in mammography screening? », *Eur. Radiol.*, vol. 31, n° 8, p. 5940-5947, août 2021, doi: 10.1007/s00330-021-07686-3.
- [43] A. J. T. Wanders *et al.*, « Interval Cancer Detection Using a Neural Network and Breast Density in Women with Negative Screening Mammograms », *Radiology*, vol. 303, n° 2, p. 269-275, mai 2022, doi: 10.1148/radiol.210832.
- [44] M. Zanotel *et al.*, « Automated breast ultrasound: basic principles and emerging clinical applications », *Radiol. Med. (Torino)*, vol. 123, n° 1, p. 1-12, janv. 2018, doi: 10.1007/s11547-017-0805-z.
- [45] M. Sasaki *et al.*, « Artificial intelligence for breast cancer detection in mammography: experience of use of the ScreenPoint Medical Transpara system in 310 Japanese women », *Breast Cancer*, vol. 27, n° 4, p. 642-651, juill. 2020, doi: 10.1007/s12282-020-01061-8.
- [46] H.-E. Kim *et al.*, « Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study », *Lancet Digit. Health*, vol. 2, n° 3, p. e138-e148, mars 2020, doi: 10.1016/S2589-7500(20)30003-0.
- [47] « Radiomics : une révolution ? », *Cleverdoc*, 20 avril 2021. <https://cleverdoc.io/radiomics-une-revolution/> (consulté le 30 décembre 2021).
- [48] M.-W. Juan, J. Yu, G.-X. Peng, L.-J. Jun, S.-P. Feng, et L.-P. Fang, « Correlation between DCE-MRI radiomics features and Ki-67 expression in invasive breast cancer », *Oncol. Lett.*, vol. 16, n° 4, p. 5084-5090, oct. 2018, doi: 10.3892/ol.2018.9271.
- [49] A. Saha, M. R. Harowicz, W. Wang, et M. A. Mazurowski, « A study of association of Oncotype DX recurrence score with DCE-MRI characteristics using multivariate machine learning models », *J. Cancer Res. Clin. Oncol.*, vol. 144, n° 5, p. 799-807, mai 2018, doi: 10.1007/s00432-018-2595-7.
- [50] N. Braman *et al.*, « Association of Peritumoral Radiomics With Tumor Biology and Pathologic Response to Preoperative Targeted Therapy for HER2 (ERBB2)-Positive Breast Cancer », *JAMA Netw. Open*, vol. 2, n° 4, p. e192561, avr. 2019, doi: 10.1001/jamanetworkopen.2019.2561.
- [51] « Transpara® User Manual ».
- [52] « Transpara Clinical Summaries ».

RÉFÉRENCES Partie 2

- [1] N. Anderson *et al.*, « Guidelines for non-operative diagnostic procedures and reporting in breast cancer screening (February 2017), Royal College of Pathologists Publication G150. », févr. 2017.
- [2] S. E. Pinder *et al.*, « NHS Breast Screening multidisciplinary working group guidelines for the diagnosis and management of breast lesions of uncertain malignant potential on core biopsy (B3 lesions) », *Clin. Radiol.*, vol. 73, n° 8, p. 682-692, août 2018, doi: 10.1016/j.crad.2018.04.004.
- [3] C. J. Rargeth *et al.*, « Second International Consensus Conference on lesions of uncertain malignant potential in the breast (B3 lesions) », *Breast Cancer Res. Treat.*, vol. 174, n° 2, p. 279-296, avr. 2019, doi: 10.1007/s10549-018-05071-1.
- [4] J.-P. Beregi *et al.*, « Radiology and artificial intelligence: An opportunity for our specialty », *Diagn. Interv. Imaging*, vol. 99, n° 11, p. 677-678, nov. 2018, doi: 10.1016/j.diii.2018.11.002.
- [5] « Artificial intelligence and medical imaging 2018: French Radiology Community white paper », *Diagn. Interv. Imaging*, vol. 99, n° 11, p. 727-742, nov. 2018, doi: 10.1016/j.diii.2018.10.003.
- [6] R. A. Castellino, « Computer aided detection (CAD): an overview », *Cancer Imaging*, vol. 5, n° 1, p. 17-19, 2005, doi: 10.1102/1470-7330.2005.0018.
- [7] J. J. Fenton *et al.*, « Effectiveness of Computer-Aided Detection in Community Mammography Practice », *JNCI J. Natl. Cancer Inst.*, vol. 103, n° 15, p. 1152-1161, août 2011, doi: 10.1093/jnci/djr206.
- [8] S. Benhamiche, « Lésions frontières B3 sur biopsies mammaires : étude descriptive monocentrique d'une population de 146 patientes opérées entre 2015 et 2021 avec création d'un ECRF ».
- [9] M. Cohen *et al.*, « When to do surgical resection for atypical breast lesions: Results of a prospective cohort of 518 lesions », *Surg. Oncol.*, vol. 34, p. 80-85, sept. 2020, doi: 10.1016/j.suronc.2020.04.002.
- [10] F. Catanzariti *et al.*, « High-risk lesions of the breast: concurrent diagnostic tools and management recommendations », *Insights Imaging*, vol. 12, p. 63, mai 2021, doi: 10.1186/s13244-021-01005-6.
- [11] A. Rodriguez-Ruiz *et al.*, « Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists », *JNCI J. Natl. Cancer Inst.*, vol. 111, n° 9, p. 916-922, sept. 2019, doi: 10.1093/jnci/djy222.
- [12] J. L. Raya-Povedano, S. Romero-Martín, E. Elías-Cabot, A. Gubern-Mérida, A. Rodríguez-Ruiz, et M. Álvarez-Benito, « AI-based Strategies to Reduce Workload in Breast Cancer Screening with Mammography and Tomosynthesis: A Retrospective Evaluation », *Radiology*, vol. 300, n° 1, p. 57-65, juill. 2021, doi: 10.1148/radiol.2021203555.
- [13] A. Rodriguez-Ruiz *et al.*, « Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study », *Eur. Radiol.*, vol. 29, n° 9, p. 4825-4832, sept. 2019, doi: 10.1007/s00330-019-06186-9.
- [14] S. Bianchi *et al.*, « Positive predictive value for malignancy on surgical excision of breast lesions of uncertain malignant potential (B3) diagnosed by stereotactic vacuum-assisted needle core biopsy (VANCB): A large multi-institutional study in Italy », *The Breast*, vol. 20, n° 3, p. 264-270, juin 2011, doi: 10.1016/j.breast.2010.12.003.
- [15] C. D. Lehman, R. D. Wellman, D. S. M. Buist, K. Kerlikowske, A. N. A. Tosteson, et D. L. Miglioretti, « Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection », *JAMA Intern. Med.*, vol. 175, n° 11, p. 1828, nov. 2015, doi: 10.1001/jamainternmed.2015.5231.
- [16] H.-E. Kim *et al.*, « Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study », *Lancet Digit. Health*, vol. 2, n° 3, p. e138-e148, mars 2020, doi: 10.1016/S2589-7500(20)30003-0.
- [17] K. Freeman *et al.*, « Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy », *The BMJ*, vol. 374, p. n1872, sept. 2021, doi: 10.1136/bmj.n1872.

- [18] S. H. Park et K. Han, « Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction », *Radiology*, vol. 286, n° 3, p. 800-809, mars 2018, doi: 10.1148/radiol.2017171920.
- [19] K. J. Geras, R. M. Mann, et L. Moy, « Artificial Intelligence for Mammography and Digital Breast Tomosynthesis: Current Concepts and Future Perspectives », *Radiology*, vol. 293, n° 2, p. 246-259, nov. 2019, doi: 10.1148/radiol.2019182627.
- [20] Y. Jiang et C. E. Metz, « BI-RADS Data Should Not Be Used to Estimate ROC Curves1 », *Radiology*, vol. 256, n° 1, p. 29-31, juill. 2010, doi: 10.1148/radiol.10091394.
- [21] A. Rodríguez-Ruiz *et al.*, « Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System », *Radiology*, vol. 290, n° 2, p. 305-314, févr. 2019, doi: 10.1148/radiol.2018181371.

Le 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.

”