

UNIVERSITÉ TOULOUSE III – PAUL SABATIER
FACULTÉS DE MÉDECINE

ANNÉE 2022

2022 TOU3 1652

THÈSE

POUR LE DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE
MÉDECINE SPÉCIALISÉE CLINIQUE

Présentée et soutenue publiquement

par

Vincent JEHANNO

le 17 octobre 2022

**Apport de la segmentation préalable des objets biologiques dans
des images microscopiques de cancer du sein, pour la prédiction
de leur sous-type moléculaire**

Directeur de thèse : Dr Camille FRANCHET

JURY

Madame le Professeur Janick SELVES	Président
Madame le Professeur Emmanuelle URO-COSTE	Assesseur
Monsieur le Professeur Pierre BROUSSET	Assesseur
Madame le Professeur Florence DALENC	Assesseur
Monsieur le Docteur Camille FRANCHET	Suppléant
Monsieur le Docteur Guillaume BATAILLON	Invité
Monsieur le Professeur Sylvain CUSSAT-BLANC	Invité

Tableau du personnel hospitalo-universitaire



FACULTE DE SANTE
Département Médecine Maieutique et Paramédicaux
Tableau des personnels HU de médecine
Mars 2022

Professeurs Honoraires

Doyen Honoraire	M. CHAP Hugues	Professeur Honoraire	M. GHISOLFI Jacques
Doyen Honoraire	M. GUIRAUD-CHAUMEIL Bernard	Professeur Honoraire	M. GLOCK Yves
Doyen Honoraire	M. LAZORTHES Yves	Professeur Honoraire	M. GOUZI Jean-Louis
Doyen Honoraire	M. PUEL Pierre	Professeur Honoraire	M. GRAND Alain
Doyen Honoraire	M. ROUGE Daniel	Professeur Honoraire	M. GUIRAUD CHAUMEIL Bernard
Doyen Honoraire	M. VINEL Jean-Pierre	Professeur Honoraire	M. HOFF Jean
Professeur Honoraire	M. ABBAL Michel	Professeur Honoraire	M. JOFFRE Francis
Professeur Honoraire	M. ADER Jean-Louis	Professeur Honoraire	M. LAGARRIGUE Jacques
Professeur Honoraire	M. ADOUE Daniel	Professeur Honoraire	M. LANG Thierry
Professeur Honoraire	M. ARBUS Louis	Professeur Honoraire	Mme LARENG Marie-Blanche
Professeur Honoraire	M. ARLET Philippe	Professeur Honoraire	M. LAURENT Guy
Professeur Honoraire	M. ARLET-GUAU Elisabeth	Professeur Honoraire	M. LAZORTHES Franck
Professeur Honoraire	M. ARNE Jean-Louis	Professeur Honoraire	M. LAZORTHES Yves
Professeur Honoraire	M. BARRET André	Professeur Honoraire	M. LEOPHONTE Paul
Professeur Honoraire	M. BARTHE Philippe	Professeur Honoraire	M. MAGNAVAL Jean-François
Professeur Honoraire	M. BAYARD Francis	Professeur Honoraire	M. MALECAZE François
Professeur Honoraire	M. BLANCHER Antoine	Professeur Honoraire	M. MANELFE Claude
Professeur Honoraire	M. BOCCALON Henri	Professeur Honoraire	M. MANSAT Michel
Professeur Honoraire	M. BONAFE Jean-Louis	Professeur Honoraire	M. MARCHOU Bruno
Professeur Honoraire	M. BONEU Bernard	Professeur Honoraire	M. MASSIP Patrice
Professeur Honoraire	M. BONNEVILLE Paul	Professeur Honoraire	Mme MARTY Nicole
Professeur Honoraire	M. BOUNHOURS Jean-Paul	Professeur Honoraire	M. MAZIERES Bernard
Professeur Honoraire	M. BOUTAULT Franck	Professeur Honoraire	M. MONROZIES Xavier
Professeur Honoraire Associé	M. BROS Bernard	Professeur Honoraire	M. MOSCOVICI Jacques
Professeur Honoraire	M. BUGAT Roland	Professeur Honoraire	M. MURAT
Professeur Honoraire	M. CAHUZAC Jean-Philippe	Professeur Honoraire associé	M. NICODEME Robert
Professeur Honoraire	M. CARATERO Claude	Professeur Honoraire	M. OLIVES Jean-Pierre
Professeur Honoraire	M. CARLES Pierre	Professeur Honoraire	M. PARINAUD Jean
Professeur Honoraire	M. CARON Philippe	Professeur Honoraire	M. PASCAL Jean-Pierre
Professeur Honoraire	M. CARRIERE Jean-Paul	Professeur Honoraire	M. PERRET Bertrand
Professeur Honoraire	M. CARTON Michel	Professeur Honoraire	M. PESSEY Jean-Jacques
Professeur Honoraire	M. CATHALA Bernard	Professeur Honoraire	M. PLANTE Pierre
Professeur Honoraire	M. CHABANON Gérard	Professeur Honoraire	M. PONTONNIER Georges
Professeur Honoraire	M. CHAMONTIN Bernard	Professeur Honoraire	M. POURRAT Jacques
Professeur Honoraire	M. CHAP Hugues	Professeur Honoraire	M. PRADERE Bernard
Professeur Honoraire	M. CHAVOIN Jean-Pierre	Professeur Honoraire	M. PRIS Jacques
Professeur Honoraire	M. CLANET Michel	Professeur Honoraire	Mme PUEL Jacqueline
Professeur Honoraire	M. CONTE Jean	Professeur Honoraire	M. PUEL Pierre
Professeur Honoraire	M. COSTAGLIOLA Michel	Professeur Honoraire	M. PUJOL Michel
Professeur Honoraire	M. COTONAT Jean	Professeur Honoraire	M. QUERLEU Denis
Professeur Honoraire	M. DABERNAT Henri	Professeur Honoraire	M. RAILHAC Jean-Jacques
Professeur Honoraire	M. DAHAN Marcel	Professeur Honoraire	M. REGIS Henri
Professeur Honoraire	M. DALOUS Antoine	Professeur Honoraire	M. REGNIER Claude
Professeur Honoraire	M. DALY-SCHWEITZER Nicolas	Professeur Honoraire	M. REME Jean-Michel
Professeur Honoraire	M. DAVID Jean-Frédéric	Professeur Honoraire	M. RISCHMANN Pascal
Professeur Honoraire	M. DELSOL Georges	Professeur Honoraire	M. RIVIERE Daniel
Professeur Honoraire	Mme DELISLE Marie-Bernadette	Professeur Honoraire	M. ROCHE Henri
Professeur Honoraire	Mme DIDIER Jacqueline	Professeur Honoraire	M. ROCHICCIOLI Pierre
Professeur Honoraire	M. DUCOS Jean	Professeur Honoraire	M. ROLLAND Michel
Professeur Honoraire	M. DUFFAUT Michel	Professeur Honoraire	M. ROQUES-LATRILLE Christian
Professeur Honoraire	M. DUPRE M.	Professeur Honoraire	M. RUMEAU Jean-Louis
Professeur Honoraire	M. DURAND Dominique	Professeur Honoraire	M. SALVADOR Michel
Professeur Honoraire associé	M. DUTAU Guy	Professeur Honoraire	M. SALVAYRE Robert
Professeur Honoraire	M. ESCHAPASSE Henri	Professeur Honoraire	M. SARRAMON Jean-Pierre
Professeur Honoraire	M. ESCOURROU Jean	Professeur Honoraire	M. SERRE Guy
Professeur Honoraire	M. ESQUERRE J.P.	Professeur Honoraire	M. SIMON Jacques
Professeur Honoraire	M. FABIE Michel	Professeur Honoraire	M. SUC Jean-Michel
Professeur Honoraire	M. FABRE Jean	Professeur Honoraire	M. THOUVENOT Jean-Paul
Professeur Honoraire	M. FOURNIAL Gérard	Professeur Honoraire	M. TREMOULET Michel
Professeur Honoraire	M. FOURNIE Bernard	Professeur Honoraire	M. VALDIGUIE Pierre
Professeur Honoraire	M. FORTANIER Gilles	Professeur Honoraire	M. VAYSSE Philippe
Professeur Honoraire	M. FRAYSSE Bernard	Professeur Honoraire	M. VINEL Jean-Pierre
Professeur Honoraire	M. FREXINOS Jacques	Professeur Honoraire	M. VIRENQUE Christian
Professeur Honoraire	Mme GENESTAL Michèle	Professeur Honoraire	M. VOIGT Jean-Jacques
Professeur Honoraire	M. GERAUD Gilles		

Professeurs Emérites

Professeur ARLET Philippe
Professeur BOUTAULT Franck
Professeur CARON Philippe
Professeur CHAMONTIN Bernard
Professeur CHAP Hugues
Professeur GRAND Alain
Professeur LAGARRIGUE Jacques
Professeur LAURENT Guy
Professeur LAZORTHES Yves
Professeur MAGNAVAL Jean-François
Professeur MARCHOU Bruno
Professeur PERRET Bertrand
Professeur RISCHMANN Pascal
Professeur RIVIERE Daniel
Professeur ROUGE Daniel

FACULTE DE SANTE
Département Médecine Maieutique et Paramédicaux

P.U. - P.H.
Classe Exceptionnelle et 1ère classe

M. ACAR Philippe	Pédiatrie	Mme LAMANT Laurence (C.E)	Anatomie Pathologique
M. ACCADBLE Franck (C.E)	Chirurgie Infantile	M. LANGIN Dominique (C.E)	Nutrition
M. ALRIC Laurent (C.E)	Médecine Interne	Mme LAPRIE Anne	Radiothérapie
M. AMAR Jacques	Thérapeutique	M. LARRUE Vincent	Neurologie
Mme ANDRIEU Sandrine	Epidémiologie, Santé publique	M. LAUQUE Dominique (C.E)	Médecine d'Urgence
M. ARBUS Christophe	Psychiatrie	M. LAUWERS Frédéric	Chirurgie maxillo-faciale
M. ARNAL Jean-François (C.E)	Physiologie	M. LEOBON Bertrand	Chirurgie Thoracique et Cardio-vasculaire
M. ATTAL Michel (C.E)	Hématologie	M. LEVADE Thierry (C.E)	Biochimie
M. AVET-LOISEAU Hervé	Hématologie, transfusion	M. LIBLAU Roland (C.E)	Immunologie
M. BERRY Antoine	Parasitologie	M. MALAUAUD Bernard	Urologie
Mme BERRY Isabelle (C.E)	Biophysique	M. MANSAT Pierre	Chirurgie Orthopédique
M. BIRMES Philippe	Psychiatrie	M. MARQUE Philippe (C.E)	Médecine Physique et Réadaptation
M. BONNEVILLE Fabrice	Radiologie	M. MAS Emmanuel	Pédiatrie
M. BOSSAVY Jean-Pierre (C.E)	Chirurgie Vasculaire	M. MAURY Jean-Philippe (C.E)	Cardiologie
M. BRASSAT David	Neurologie	Mme MAZEREEUW Juliette	Dermatologie
M. BROUCHET Laurent	Chirurgie thoracique et cardio-vascul	M. MAZIERES Julien (C.E)	Pneumologie
M. BROUSSET Pierre (C.E)	Anatomie pathologique	M. MINVILLE Vincent	Anesthésiologie Réanimation
M. BUJAN Louis (C.E)	Urologie-Andrologie	M. MOLINIER Laurent (C.E)	Epidémiologie, Santé Publique
Mme BURA-RIVIERE Alessandra (C.E)	Médecine Vasculaire	M. MONTASTRUC Jean-Louis (C.E)	Pharmacologie
M. BUREAU Christophe	Hépatogastro-entérologie	Mme MOYAL Elisabeth (C.E)	Cancérologie
M. BUSCAIL Louis (C.E)	Hépatogastro-entérologie	M. MUSCARI Fabrice	Chirurgie Digestive
M. CALVAS Patrick (C.E)	Génétique	Mme NOURHASHEMI Fatemeh (C.E)	Gériatrie
M. CANTAGREL Alain (C.E)	Rhumatologie	M. OLIVOT Jean-Marc	Neurologie
M. CARRERE Nicolas	Chirurgie Générale	M. OSWALD Eric (C.E)	Bactériologie-Virologie
M. CARRIE Didier (C.E)	Cardiologie	M. PARIENTE Jérémie	Neurologie
M. CHAIX Yves	Pédiatrie	M. PAUL Carle (C.E)	Dermatologie
Mme CHARPENTIER Sandrine	Médecine d'urgence	M. PAYOUX Pierre (C.E)	Biophysique
M. CHAUFOUR Xavier	Chirurgie Vasculaire	M. PAYRASTRE Bernard (C.E)	Hématologie
M. CHAUVEAU Dominique	Néphrologie	M. PERON Jean-Marie (C.E)	Hépatogastro-entérologie
M. CHAYNES Patrick	Anatomie	M. RASCOL Olivier (C.E)	Pharmacologie
M. CHIRON Philippe (C.E)	Chir. Orthopédique et Traumatologie	Mme RAUZY Odile	Médecine Interne
M. CHOLLET François (C.E)	Neurologie	M. RAYNAUD Jean-Philippe (C.E)	Psychiatrie Infantile
M. CONSTANTIN Amaud	Rhumatologie	M. RECHER Christian(C.E)	Hématologie
M. COURBON Frédéric	Biophysique	M. RITZ Patrick (C.E)	Nutrition
Mme COURTADE SAIDI Monique (C.E)	Histologie Embryologie	M. ROLLAND Yves (C.E)	Gériatrie
M. DAMBRIN Camille	Chir. Thoracique et Cardiovasculaire	M. RONCALLI Jérôme	Cardiologie
M. DE BOISSEZON Xavier	Médecine Physique et Réadapt Fonct.	M. ROUGE Daniel (C.E)	Médecine Légale
M. DEGUINE Olivier (C.E)	Oto-rhino-laryngologie	M. ROUSSEAU Hervé (C.E)	Radiologie
M. DELABESSE Eric	Hématologie	M. ROUX Franck-Emmanuel	Neurochirurgie
M. DELOBEL Pierre	Maladies Infectieuses	M. SAILLER Laurent (C.E)	Médecine Interne
M. DELORD Jean-Pierre (C.E)	Cancérologie	M. SALES DE GAUZY Jérôme (C.E)	Chirurgie Infantile
M. DIDIER Alain (C.E)	Pneumologie	M. SALLES Jean-Pierre (C.E)	Pédiatrie
M. DUCOMMUN Bernard	Cancérologie	M. SANS Nicolas	Radiologie
Mme DULY-BOUHANICK Beatrice (C.E)	Thérapeutique	M. SCHMITT Laurent (C.E)	Psychiatrie
M. ELBAZ Meyer	Cardiologie	Mme SELVES Janick (C.E)	Anatomie et cytologie pathologiques
M. FERRIERES Jean (C.E)	Epidémiologie, Santé Publique	M. SENARD Jean-Michel (C.E)	Pharmacologie
M. FOURCADE Olivier	Anesthésiologie	M. SERRANO Elle (C.E)	Oto-rhino-laryngologie
M. FOURNIE Pierre	Ophthalmologie	M. SIZUN Jacques (C.E)	Pédiatrie
M. GALINIER Michel (C.E)	Cardiologie	M. SOL Jean-Christophe	Neurochirurgie
M. GAME Xavier	Urologie	Mme SOTO-MARTIN Maria-Eugenia	Gériatrie et biologie du vieillissement
Mme GARDETTE Virginie	Epidémiologie, Santé publique	M. SOULAT Jean-Marc	Médecine du Travail
M. GEERAERTS Thomas	Anesthésiologie et réanimation	M. SOULIE Michel (C.E)	Urologie
Mme GOMEZ-BROUCHET Anne-Muriel	Anatomie Pathologique	M. SUC Bertrand	Chirurgie Digestive
M. GOURDY Pierre (C.E)	Endocrinologie	Mme TAUBER Marie-Thérèse (C.E)	Pédiatrie
M. GROLLEAU RAOUX Jean-Louis (C.E)	Chirurgie plastique	M. TELMON Norbert (C.E)	Médecine Légale
Mme GUIMBAUD Rosine	Cancérologie	Mme TREMOLLIERES Florence	Biologie du développement
Mme HANAIRE Hélène (C.E)	Endocrinologie	Mme URO-COSTE Emmanuelle (C.E)	Anatomie Pathologique
M. HUYGHE Eric	Urologie	M. VAYSSIERE Christophe (C.E)	Gynécologie Obstétrique
M. IZOPET Jacques (C.E)	Bactériologie-Virologie	M. VELLAS Bruno (C.E)	Gériatrie
M. KAMAR Nassim (C.E)	Néphrologie	M. VERGEZ Sébastien	Oto-rhino-laryngologie
P.U. Médecine générale			
M. OUSTRIC Stéphane (C.E)			

FACULTE DE SANTE
Département Médecine Maieutique et Paramédicaux

P.U. - P.H. 2ème classe	Professeurs Associés
M. ABBO Olivier	Chirurgie Infantile
M. AUSSEIL Jérôme	Biochimie et biologie moléculaire
Mme BONGARD Vanina	Epidémiologie, Santé publique
M. BONNEVILLE Nicolas	Chirurgie orthopédique et traumatologique
M. BOUNES Vincent	Médecine d'urgence
Mme BOURNET Barbara	Gastro-entérologie
Mme CASPER Charlotte	Pédiatrie
M. CAVAGNAC Etienne	Chirurgie orthopédique et traumatologie
M. CHAPUT Benoit	Chirurgie plastique
M. COGNARD Christophe	Radiologie
Mme CORRE Jill	Hématologie
Mme DALENC Florence	Cancérologie
M. DE BONNECAZE Guillaume	Anatomie
M. DECRAMER Stéphane	Pédiatrie
M. EDOUARD Thomas	Pédiatrie
M. FAGUER Stanislas	Néphrologie
Mme FARUCH BILFELD Marie	Radiologie et imagerie médicale
M. FRANCHITTO Nicolas	Addictologie
M. GARRIDO-STOWHAS Ignacio	Chirurgie Plastique
M. GUIBERT Nicolas	Pneumologie
M. GUILLEMINAULT Laurent	Pneumologie
M. HERIN Fabrice	Médecine et santé au travail
M. LAIREZ Olivier	Biophysique et médecine nucléaire
M. LAROCHE Michel	Rhumatologie
Mme LAURENT Camille	Anatomie Pathologique
M. LE CAIGNEC Cédric	Génétique
M. LEANDRI Roger	Biologie du dével. et de la reproduction
M. LOPEZ Raphael	Anatomie
M. MARCHEIX Bertrand	Chirurgie thoracique et cardiovasculaire
M. MARTIN-BONDEL Guillaume	Maladies Infectieuses, maladies tropicales
Mme MARTINEZ Alejandra	Gynécologie
M. MARX Mathieu	Oto-rhino-laryngologie
M. MEYER Nicolas	Dermatologie
M. PAGES Jean-Christophe	Biologie cellulaire
Mme PASQUET Marlène	Pédiatrie
M. PORTIER Guillaume	Chirurgie Digestive
M. PUGNET Grégory	Médecine Interne
M. REINA Nicolas	Chirurgie orthopédique et traumatologique
M. RENAUDINEAU Yves	Immunologie
Mme RUYSSSEN-WITRAND Adeline	Rhumatologie
Mme SAVAGNER Frédérique	Biochimie et biologie moléculaire
M. SAVALL Frédéric	Médecine légale
M. SILVA SIFONTES Stein	Réanimation
M. SOLER Vincent	Ophthalmologie
Mme SOMMET Agnès	Pharmacologie
M. TACK Ivan	Physiologie
Mme VAYSSE Charlotte	Cancérologie
Mme VEZZOSI Delphine	Endocrinologie
M. YRONDI Antoine	Psychiatrie
M. YSEBAERT Loïc	Hématologie
P.U. Médecine générale	
M. MESTHE Pierre	
Mme ROUGE-BUGAT Marie-Eve	
Professeur Associé de Médecine Générale	
M. ABITTEBOUL Yves	
Mme BOURGEOIS Odile	
M. BOYER Pierre	
M. CHICOULAA Bruno	
Mme IRI-DELAHAYE Motoko	
M. PIPONNIER David	
M. POUTRAIN Jean-Christophe	
M. STILLMUNKES André	
Professeur Associé de Bactériologie-Hygiène	
Mme MALAUAUD Sandra	

FACULTE DE SANTE
Département Médecine Maieutique et Paramédicaux

MCU - PH

Mme ABRAVANEL Florence	Bactériologie Virologie Hygiène	Mme GENNERO Isabelle	Biochimie
M. APOIL Poi Andre	Immunologie	Mme GENOUX Annelise	Biochimie et biologie moléculaire
Mme ARNAUD Catherine	Epidémiologie	Mme GRARE Marion	Bactériologie Virologie Hygiène
Mme AUSSEIL-TRUDEL Stéphanie	Biochimie	M. GUERBY Paul	Gynécologie-Obstétrique
Mme BASSET Céline	Cytologie et histologie	Mme GUILBEAU-FRUGIER Céline	Anatomie Pathologique
Mme BELLIERES-FABRE Julie	Néphrologie	Mme GUYONNET Sophie	Nutrition
Mme BERTOLI Sarah	Hématologie, transfusion	M. HAMDJ Safouane	Biochimie
M. BIETH Eric	Génétique	Mme HITZEL Anne	Biophysique
Mme BREHIN Camille	Pneumologie	Mme INGUENEAU Cécile	Biochimie
M. BUSCAIL Etienne	Chirurgie viscérale et digestive	M. IRIART Xavier	Parasitologie et mycologie
Mme CAMARE Caroline	Biochimie et biologie moléculaire	Mme JONCA Nathalie	Biologie cellulaire
M. CAMBUS Jean-Pierre	Hématologie	M. KIRZIN Sylvain	Chirurgie générale
Mme CANTERO Anne-Vaérie	Biochimie	Mme LAPEYRE-MESTRE Maryse	Pharmacologie
Mme CARFAGNA Luana	Pédiatrie	M. LEPAGE Benoit	Biostatistiques et Informatique médicale
Mme CASPAR BAUGUIL Sylvie	Nutrition	M. LHERMUSIER Thibaut	Cardiologie
Mme CASSAGNE Myriam	Ophtalmologie	M. LHOMME Sébastien	Bactériologie-virologie
Mme CASSAING Sophie	Parasitologie	Mme MASSIP Clémence	Bactériologie-virologie
Mme CASSOL Emmanuelle	Biophysique	Mme MAUPAS SCHWALM Françoise	Biochimie
Mme CHANTALAT Elodie	Anatomie	Mme MONTASTIER Emilie	Nutrition
M. CHASSAING Nicolas	Génétique	M. MONTASTRUC François	Pharmacologie
M. CLAVEL Cyril	Biologie Cellulaire	Mme MOREAU Jessika	Biologie du dev. Et de la reproduction
Mme COLOMBAT Magali	Anatomie et cytologie pathologiques	Mme MOREAU Marion	Physiologie
M. CONGY Nicolas	Immunologie	M. MOULIS Guillaume	Médecine Interne
Mme COURBON Christine	Pharmacologie	Mme NASR Nathalie	Neurologie
M. CUROT Jonathan	Neurologie	Mme NOGUEIRA M.L.	Biologie Cellulaire
Mme DAMASE Christine	Pharmacologie	Mme PERROT Aurore	Hématologie
Mme DE GLISEZENSKY Isabelle	Physiologie	M. PILLARD Fabien	Physiologie
M. DEDUIT Fabrice	Médecine Légale	Mme PLAISANCIE Julie	Génétique
M. DEGBOE Yannick	Rhumatologie	Mme PUISSANT Bénédicte	Immunologie
M. DELMAS Clément	Cardiologie	Mme QUELVEN Isabelle	Biophysique et médecine nucléaire
M. DELPLA Pierre-André	Médecine Légale	Mme RAYMOND Stéphanie	Bactériologie Virologie Hygiène
M. DESPAS Fabien	Pharmacologie	M. REVET Alexis	Pédo-psychiatrie
M. DUBOIS Damien	Bactériologie Virologie Hygiène	M. RIMAILHO Jacques	Anatomie et Chirurgie Générale
Mme ESQUIROL Yolande	Médecine du travail	Mme SABOURDY Frédérique	Biochimie
Mme EVRARD Solène	Histologie, embryologie et cytologie	Mme SAUNE Karine	Bactériologie Virologie
Mme FILLAUX Judith	Parasitologie	Mme SIEGFRIED Aurore	Anatomie et cytologie pathologiques
Mme FLOCH Pauline	Bactériologie-Virologie	M. TAFANI Jean-André	Biophysique
Mme GALINIER Anne	Nutrition	M. TREINER Emmanuel	Immunologie
Mme GALLINI Adeline	Epidémiologie	Mme VALLET Marion	Physiologie
M. GANTET Pierre	Biophysique	M. VERGEZ François	Hématologie
M. GASQ David	Physiologie	Mme VIJA Lavinia	Biophysique et médecine nucléaire
M. GATIMEL Nicolas	Médecine de la reproduction		

M.C.U. Médecine générale

M. BISMUTH Michel
M. BRILLAC Thierry
Mme DUPOUY Julie
M. ESCOURROU Emile

Maitres de Conférence Associés

M.C.A. Médecine Générale

M. BIREBENT Jordan
Mme BOUSSIER Nathalie
Mme FREYENS Anne
Mme LATROUS Lella
Mme PUECH Marielle

Remerciements

A ma présidente du Jury,

Madame la Professeure Janick Selves

Professeur des Universités - Praticien hospitalier
Anatomie et cytologie pathologiques

Vous nous faites l'honneur et le plaisir de présider ce jury de thèse. Votre enthousiasme à enseigner, communiquer et apprendre est admirable et inspirant. Nous avons la chance de bénéficier de votre enseignement et de votre encadrement bienveillant tout au long de notre internat en tant que coordonnatrice de DES. Vous prenez ce rôle très à cœur. Vous veillez à la qualité de notre formation et nous encouragez à toujours donner le meilleur de nous-même. Soyez assurée de ma plus sincère reconnaissance.

A mon directeur de thèse,

Monsieur le Docteur Camille Franchet

Praticien hospitalier
Anatomie et cytologie pathologiques

Tu m'as fait l'honneur d'accepter et de diriger ce travail de thèse et je t'en remercie sincèrement, car ce projet comptait beaucoup pour moi. De mes premières lames en pathologie mammaire jusqu'à ce travail d'intelligence artificielle, tu as toujours été un excellent pédagogue, enseignant dans la bonne humeur et la bienveillance. Ta rigueur, ta curiosité et ta passion pour ce métier sont exemplaires à mes yeux. Je te suis très reconnaissant et espère être digne de la confiance que tu m'accordes.

Aux membres du Jury,

Monsieur le Professeur Pierre Brousset

Professeur des Universités - Praticien hospitalier
Anatomie et cytologie pathologiques

Vous nous faites l'honneur de siéger à ce jury et nous vous sommes très reconnaissant de juger notre travail. Nous avons la chance de bénéficier de votre engouement pour la recherche, dans ce service que vous dirigez. Ce projet d'intelligence artificielle comme d'autres, a pu voir le jour grâce à votre intérêt pour ce domaine. Je vous en suis très reconnaissant et espère que ce travail vous donnera satisfaction. Je vous prie d'agréer l'expression de notre profond respect et de notre sincère reconnaissance.

Madame la Professeure Emmanuelle Uro-Coste

Professeure des Universités - Praticien hospitalier
Anatomie et cytologie pathologiques

Vous nous faites l'honneur de siéger à ce jury et nous vous sommes très reconnaissant de juger notre travail.

Votre bonne humeur et votre humour viennent égayer chaque évènement du service. Et nous avons la chance de bénéficier de votre enthousiasme à travers tous les cours que vous réalisez. Votre pédagogie captive beaucoup d'internes partout en France. Veuillez trouver ici l'expression de ma sincère reconnaissance.

Madame la Professeure Florence Dalenc

Professeure des Universités - Praticien hospitalier
Oncologie sénologique

Vous nous faites l'honneur de siéger à ce jury et nous vous sommes très reconnaissant de juger notre travail.

Nous connaissons et admirons vos qualités professionnelles appréciées de vos confrères et de vos patientes. Soyez assurée de notre profond respect et de notre sincère reconnaissance.

Monsieur le Docteur Guillaume Bataillon

Praticien hospitalier
Anatomie et cytologie pathologiques

Tu nous fais l'honneur de siéger à ce jury et nous te sommes très reconnaissant de juger notre travail.

Je n'ai eu que de rares occasions de travailler avec toi pour l'instant. Ton dynamisme, ta gentillesse et ta rigueur te précèdent et te bâtissent une belle réputation. Il me tarde d'apprendre la gynécopathologie à tes côtés.

Monsieur le Professeur Sylvain Cussat-Blanc

Professeur assistant à l'Université Toulouse 1 Capitole

Tu nous fais l'honneur de siéger à ce jury et nous te sommes très reconnaissant de juger notre travail.

Tu as supervisé le travail de Mehdi tout au long de ce projet et ton implication a été très enrichissante. Tu nous as énormément appris sur l'intelligence artificielle, toujours avec facilité et bonne humeur. Nous te sommes sincèrement reconnaissant de ton implication dans ce projet et dans tous les autres.

A ceux et celles qui ont participé à ce travail,
Mehdi, Robin, Sandrine, Stéphanie, Samira et Nadia,

Au service d'anatomopathologie de l'Oncopole,
A l'ensemble des séniors,
A l'ensemble des techniciens, secrétaires et cadres du labo,
Aux assistants et anciens co-internes,
A mes actuels co-internes,

Au service d'anatomopathologie de Valence,

A mes amis de Toulouse,

A mes amis de Rennes,

Et enfin à ma famille,
A mes parents,
A mes frères.

Table des matières

1 – Introduction	14
1.1 - Epidémiologie et prise en charge thérapeutique du cancer du sein en France	14
1.2 - Les critères clinico-pathologiques traditionnels (CCPT)	15
1.2.1 - Caractéristiques cliniques	15
1.2.2 - Caractéristiques histologiques	15
1.2.3 - Caractéristiques immunohistochimiques	22
1.2.4 - Les limites des CCPT	23
1.3 - Caractéristiques moléculaires : les signatures moléculaires	27
1.4 - Prédiction avec l'aide de l'intelligence artificielle	29
1.4.1 - Définition de l'intelligence artificielle	29
1.4.2 - Les systèmes experts	30
1.4.3 - Le machine learning (ML)	30
1.4.4 - Le deep learning (DL)	32
1.4.5 - Prédiction des altérations moléculaires	35
1.5 - Objectifs de l'étude	37
2 - Matériel et méthodes	38
2.1 - Populations de l'étude	38
2.1.1 - Cohorte d'entraînement	38
2.1.2 - Cohorte de test	39
2.1.3 - Description des cohortes	39
2.2 - Méthodes	42
2.2.1 - Préparation des whole slide images	42
2.2.2 - Prédiction à partir des IHC et d'autres CCPT	43
2.2.3 - Prédiction par analyse d'image en DL	44
2.2.4 - Prédiction par analyse d'image en ML à partir des features nucléaires	48
2.2.5 - Combinaisons des différentes méthodes	52
2.2.6 - Evaluation des performances et de l'interprétabilité des modèles	53
2.2.6.1 - Evaluation des performances	53
2.2.6.2 - Evaluation de l'interprétabilité	53
3 - Résultats et discussion	56
3.1 - Performances de la prédiction à partir des IHC et d'autres CCPT dans un dataset monocentrique (Oncopole)	56
3.1.1 - Selon St Gallen et Maisonneuve	56
3.1.2 - Prédiction et interprétabilité par ML avec XG boost	56
3.1.3 - Conclusion préliminaire sur la prédiction par les IHC et les CCPT	59
3.2 – Performances de la prédiction par analyse d'image en DL	60
3.2.1 - Performances sur patch	60
3.2.2 - Performances sur lame entière	61
3.2.3 - Apport du DL à la prédiction par les IHC et les CCPT	63
3.2.4 - Conclusions préliminaires sur la prédiction par DL	64
3.3 – Performances de la prédiction par analyse d'image en ML à partir de features extraites au préalable	65
3.3.1 - Performances sur les features seules	65
3.3.2 - Apport des features sur la prédiction par les IHC et les CCPT	69
3.3.3 - Conclusion préliminaire sur la prédiction par les features nucléaires	73

4 - Discussion des limites, forces et perspectives de l'étude	76
4.1 - Limites de l'étude.....	76
4.2 - Forces de l'étude	77
4.3 - Perspectives.....	78
5 - Conclusion	81

Liste des tableaux et figures

Tableau 1 : classification pTNM pour les cancers du sein selon l'UICC (Union for International Cancer Control) 2017 (8ème édition)

Tableau 2 : les différents types histologiques de carcinomes mammaires invasifs

Tableau 3 : détail du grade histopronostique de Nottingham

Tableau 4 : caractéristiques clinico-pathologiques et moléculaire des populations de l'étude

Tableau 5 : hyperparamètres utilisés lors de l'entraînement de nos modèles

Tableau 6 : liste des 103 features extraites

Tableau 7 : performances de prédiction du sous-type intrinsèque sur les IHC proposés par St Gallen et Maisonneuve

Tableau 8 : performances de prédiction du sous-type intrinsèque par ML (XGboost) sur les IHC et les CCPT

Tableau 9 : répartition selon les sous-types intrinsèques des valeurs de la taille tumorale, du nombre de mitoses / mm², du nombre de ganglions envahis, et du pourcentage de positivité aux RE et RP dans la cohorte d'entraînement

Tableau 10 : répartition selon les sous-types intrinsèques des valeurs de la taille tumorale, du nombre de mitoses / mm², du nombre de ganglions envahis, et du pourcentage de positivité aux RE et RP dans la cohorte de test

Tableau 11 : performances des prédictions du sous-type intrinsèque par DL seul sur patch

Tableau 12 : performances des prédictions du sous-type intrinsèque par DL seul sur lame entière en effectuant une moyenne des prédictions

Tableau 13 : performances des prédictions du sous-type intrinsèque par DL seul sur lame entière en effectuant un vote à la majorité

Tableau 14 : performances de prédiction du sous-type intrinsèque par ML (XGboost) sur la combinaison des IHC, des CCPT et du DL

Tableau 15 : performances des prédictions du sous-type intrinsèque par ML (random forest) sur les features seules

Tableau 16 : performances des prédictions du sous-type intrinsèque par ML (random forest) sur la combinaison des features aux IHC et aux CCPT

Tableau 17 : performances des prédictions du sous-type intrinsèque par ML (random forest) sur la combinaison des features aux IHC et aux CCPT, après élimination des features corrélées

Figure 1 : exemple de WSI utilisée pour nos modèles

Figure 2 : exemple de patch 512 x 512 pixels

Figure 3 : exemple de zone d'annotation (en jaune)

Figure 4 : exemples d'augmentation de données

Figure 5 : exemple de représentation des prédictions de l'ensemble des patchs de la lames (bleu = luminal A ; jaune = luminal B ; violet = autre)

Figure 6 : exemple de segmentation par HoVer-Net

Figure 7 : exemple de bonne segmentation (à gauche) et de mauvaise segmentation (à droite)

Figure 8 : distinction des cellules tumorales (haut) et des lymphocytes (bas) sur les features

Figure 9 : exemple de matrice de confusion

Figure 10 : exemple de Grad-CAM pour un patch (à gauche) et de Heatmap à l'échelle d'une WSI (à droite)

Figure 11 : exemple de SHAP Value, avec les données d'entrée triées du haut vers le bas, par ordre d'importance pour l'algorithme, les plus importantes étant en haut

Figure 12 : SHAP value de la prédiction du sous-type intrinsèque par ML (XGboost) sur la combinaison des IHC et des CCPT

Figure 13 : exemple de Gradcam pour un patch, avec les zones d'intérêt en rouge pour l'algorithme
(a = prédiction luminal A ; b = prédiction luminal B ; c = prédiction autre)

Figure 14 : exemple de carte de prédiction et de heatmaps à l'échelle d'une lame entière. Ici, il s'agit d'une tumeur classée luminaire B par le modèle. (a = WSI brute ; b = cartographie des prédictions par patch avec les patchs prédits luminal A en bleu, les patchs prédits luminal B en jaune et les patchs prédits autre en violet ; c = heatmaps pour la prédiction luminal B ; d = heatmaps pour la prédiction luminal A)

Figure 15 : SHAP value de la prédiction du sous-type intrinsèque par ML (XGboost) sur la combinaison des IHC, des CCPT et du DL

Figure 16 : SHAP value de la prédiction du sous-type intrinsèque par ML (random forest) sur les features seules

Figure 17 : SHAP value de la prédiction du sous-type intrinsèque par ML (random forest) sur la combinaison des features aux IHC

Figure 18 : SHAP value de la prédiction du sous-type intrinsèque par ML (random forest) sur la combinaison des features aux IHC et aux CCPT

Figure 19 : SHAP value de la prédiction du sous-type intrinsèque par ML (random forest) sur la combinaison des features aux IHC et aux CCPT après élimination des features corrélées

Figure 20 : analyse de patchs présentant des valeurs d'`eccentricity_std` extrêmes (en haut : sous-type luminal A avec une `eccentricity_std` haute (0,17) ; en bas : sous-type luminal B avec une `eccentricity_std` basse (0,14))

Liste des abréviations

AUC : area under the curve
ARN : acide ribonucléique
cDNA : acide désoxyribonucléique complémentaire
CNN : convolutional neural networks
DL : deep learning
HAS : haute autorité de santé
HE : hémalun-éosine
HIFs : human interpretable features
IA : intelligence artificielle
INCa : Institut national du cancer
IHC : Immunohistochimie
IUCT : Institut Universitaire du Cancer de Toulouse
ML : machine learning
NN : neural networks
NPI : Nottingham prognostic index
NST : non special type
OMS : organisation mondiale de la santé
RCP : réunion de concertation pluri-disciplinaire
RE : récepteur aux oestrogènes
RGB : Red Green Blue
RH : récepteur hormonal
RP : récepteur à la progestérone
ROR (Risk Of Recurrence)
RT-PCR : Reverse transcription polymerase chain reaction
SHAP : shapley additive explanations
TILS : tumor infiltrating lymphocytes
TMA : tissue microarray
TNBC : triple negative breast cancer
UICC : Union for International Cancer Control
WSI : whole slide images
XG boost : extrem gradient boost

1 – Introduction

1.1 - Epidémiologie et prise en charge thérapeutique du cancer du sein en France

Le cancer du sein est le cancer le plus fréquent chez la femme, avec près de 2,3 millions de nouveaux cas diagnostiqués dans le monde en 2020 (1). En France métropolitaine, l'incidence annuelle chez la femme a presque doublé entre 1990 et 2018, passant de 29 970 à 58 459 cas annuels en 2018, soit une augmentation de 1,1% par an en moyenne (2). En revanche, on observe une tendance inverse du taux de mortalité durant cette même période, avec une diminution de 1,3% par an en moyenne, passant de 20,2 pour 100 000 en 1990, à 14,0 en 2018.

La prise en charge thérapeutique des cancers du sein infiltrants localisés (stades I, II et III) comprend un traitement locorégional (chirurgie +/- radiothérapie) associé ou non à un traitement systémique (pouvant combiner chimiothérapie, hormonothérapie et thérapie ciblée anti-HER2).

Pour un même stade de la maladie, la survie des patientes atteintes d'un cancer du sein présente une hétérogénéité (3). De multiples études (4, 5) ont démontré le bénéfice de ces thérapies adjuvantes en termes de pronostic, mais il existe néanmoins une variabilité dans la réponse à ces traitements selon les patientes, certaines ne bénéficiant que très peu de l'amélioration du pronostic (6, 7). Pour affiner l'évaluation de l'agressivité de ces tumeurs, prédire la réponse aux traitements et ainsi, déterminer les modalités de la séquence thérapeutique à adopter, des critères cliniques et pathologiques sont utilisés comme facteurs pronostiques de la maladie et/ou comme facteurs prédictifs de réponse au traitement.

1.2 - Les critères clinico-pathologiques traditionnels (CCPT)

Outre le stade de la maladie, les facteurs pronostiques et prédictifs utilisés pour prédire la survie sont des caractéristiques cliniques et anatomopathologiques regroupées dans le terme “Critères Clinico-Pathologiques Traditionnels”.

1.2.1 - Caractéristiques cliniques

De nombreuses études ont montré que les cancers du sein de la femme jeune (moins de 45 ans) ont un moins bon pronostic (8 - 11). Ceci s’explique en partie par le fait qu’ils présentent souvent des caractéristiques clinico-pathologiques plus agressives, avec une taille tumorale plus grande, un envahissement vasculaire et ganglionnaire plus fréquent et un grade histo-pronostique plus élevé (12, 13). Ils comportent également une répartition différente des phénotypes moléculaires avec une plus grande proportion de tumeurs triple négatives ainsi que de tumeurs lumineales B (14). Mais outre ces caractéristiques péjoratives, l’âge jeune est également un facteur pronostique indépendant défavorable sur la survie (12, 13). Le seuil actuellement utilisé en France pour définir un âge jeune est fixé à 35 ans. Pour ces tumeurs, l’attitude thérapeutique sera plus agressive.

L’autre caractéristique clinique utilisée dans la décision thérapeutique est l’état général de la patiente, qui conditionne sa capacité à recevoir un traitement systémique ; les comorbidités étant corrélées au pronostic individuel (15). Dans ces cas, l’attitude thérapeutique pourra être adaptée.

1.2.2 - Caractéristiques histologiques

Parmi les caractéristiques déterminées par l’analyse anatomopathologique des tumeurs du sein, celles utilisées comme facteurs pronostiques ou prédictifs sont l’extension tumorale (pT), l’extension ganglionnaire régionale (pN), le type histologique, le grade histopronostique de Nottingham, l’invasion vasculaire et les lymphocytes infiltrant la tumeur (TILS).

pT : extension tumorale locale et pN : extension ganglionnaire régionale

L’extension tumorale locale notée pT, et l’extension tumorale aux ganglions régionaux notée pN, sont 2 des 3 paramètres évalués dans le stade TNM (tableau 1), qui est une stadification internationale des tumeurs comprenant l’extension tumorale locale (T),

l'extension ganglionnaire régionale (N) et la présence de métastases à distance (M). Ce stade est utilisé très largement en cancérologie et son intérêt pronostique est clairement établi (16).

pT	
pTx	Détermination de taille de la tumeur primitive impossible
pT0	Absence de signe de tumeur primitive
pTis	Carcinome in situ
	pTis (CCIS) Carcinome canalaire in situ
	pTis (CLIS) Carcinome lobulaire in situ
	pTis (Paget) Maladie de Paget du mamelon sans tumeur décelable
pT1	Tumeur ≤ 2 cm
	pT1mi tumeur ≤ 0,1 cm
	pT1a 0,1 cm < tumeur ≤ 0,5 cm
	pT1b 0,5 cm < tumeur ≤ 1 cm
	pT1c 1 cm < tumeur ≤ 2 cm
pT2	2 cm < tumeur ≤ 5 cm
pT3	tumeur > 5 cm
pT4	extension locale à la paroi thoracique et/ou la peau
	pT4a extension à la paroi thoracique
	pT4b extension à la peau
	pT4c extension à la paroi thoracique et à la peau
	pT4d carcinome inflammatoire
pN	
pNx	Atteinte ganglionnaire impossible à préciser
pN0	Absence d'envahissement ganglionnaire
	pN0 (i+) Présence de cellules tumorales isolées ≤ 0.2mm
	pN0 (mol+) Absence de signe histologique d'envahissement mais test moléculaire positif (RT-PCR)
pN1	
	pN1mi métastase unique de plus de 200 cellules 0,2 mm < taille ≤ 2 mm

pN1a	métastases ganglionnaires dans 1-3 ganglions axillaires, avec au moins une de plus de 2 mm
pN1b	métastases ganglionnaires dans des ganglions sentinelles de la chaîne mammaire interne
pN1c	métastases ganglionnaires dans 1-3 ganglions axillaires et dans des ganglions sentinelles de la chaîne mammaire interne
pN2	
pN2a	métastases ganglionnaires axillaires ipsilatérales dans 4 à 9 ganglions avec au moins une > 2 mm
pN2b	métastases ganglionnaires cliniques dans la chaîne mammaire interne seule (hors ganglion sentinelle)
pN3	
pN3a	métastases ganglionnaires infraclaviculaires ou dans au moins 10 ganglions axillaires
pN3b	métastases ganglionnaires axillaires ipsilatérales et dans la chaîne mammaire interne clinique ou dans 3 ganglions axillaires et dans un ganglion sentinelle de la chaîne mammaire interne non vu cliniquement
pN3c	métastases ganglionnaires supraclaviculaires
pM0	Absence de métastase à distance
pM1	Présence de métastase à distance

Tableau 1 : classification pTNM pour les cancers du sein selon l'UICC (Union for International Cancer Control) 2017 (8ème édition)

La présence de métastase(s) à distance fait chuter la survie à 5 ans en dessous de 30 % (17). Cet évènement permet de séparer les cancers du sein en 2 groupes : localisés et métastatiques. Dans notre étude, nous nous intéresserons exclusivement aux cancers du sein localisés.

Le pT et le pN sont des facteurs pronostiques indépendants, corrélés au risque de récurrence locale, locorégionale et à distance (18).

Le pT comprend l'évaluation de la taille tumorale dans sa plus grande dimension et l'extension aux autres tissus adjacents à la tumeur (paroi thoracique, peau). A statut ganglionnaire pN0 équivalent, la survie globale à 5 ans passe de 96,3 % pour les tumeurs pT1, à 89,4 % pour les tumeurs pT2, et 82,2 % pour les tumeurs pT3 (19).

Le pN comprend l'évaluation du statut métastatique ou non des ganglions axillaires homolatéraux, de la chaîne mammaire interne, et des régions infra et supra-claviculaires. A l'heure actuelle, c'est le facteur pronostique le plus important dans la prédiction de la survie liée au cancer et de la rechute métastatique (16, 20). La survie

sans récurrence à 5 ans pour des tumeurs de tailles équivalentes est de 82,8 % pour les patientes pN0, de 73 % pour les patientes pN1a, de 45,7 % pour les patientes ayant entre 4 et 12 ganglions envahis, et 28,4 % pour les patientes ayant au moins 13 ganglions envahis (21). Le nombre de ganglions envahis est également corrélé au risque de rechute métastatique à distance (19).

L'évaluation de ces critères est objective et possède donc une bonne reproductibilité inter et intra-observateur.

Type histologique

Actuellement, la dernière classification de l'Organisation Mondiale de la Santé (OMS 2019, 5ème édition) définit un type histologique par la présence d'un pattern histologique dans plus de 90% de la tumeur. Elle reconnaît 9 types de carcinome mammaire invasif, dont les différences sont résumées en tableau 2 (22). Les tumeurs ne présentant pas de pattern histologique particulier sont appelées NST (no special type). C'est le type histologique le plus fréquent.

Fréquence	Type histologique	Pronostic
60 - 75 %	Carcinome NST	Variable
5 - 15 %	Carcinome lobulaire infiltrant	Variable
2 %	Carcinome mucineux	Très bon
1,6 %	Carcinome tubulaire	Très bon
0,9 - 2 %	Carcinome micropapillaire	Mauvais
1 %	Carcinome à différenciation apocrine	Mauvais
0,2 - 1 %	Carcinome métaplasique	Mauvais (mais variable en fonction de la différenciation)
0,4 %	Carcinome cribriforme	Très bon
très rare (30 cas)	Cystadénocarcinome mucineux	Bon (mais peu de données)

Tableau 2 : les différents types histologiques de carcinomes mammaires invasifs

Grade histopronostique de Nottingham

Le grade histopronostique de Nottingham est un système de grading des tumeurs du sein développé initialement par Bloom et Richardson en 1957, puis modifié en 1991 par Elston et Ellis (23, 24). Le grade est calculé par addition de 3 scores (tableau 3), notés de 1 à 3 et attribués à 3 caractéristiques histologiques de la tumeur :

- *la différenciation tumorale* : évaluée par le pourcentage de structures glandulaires que forme la tumeur
- *les atypies cyto-nucléaires* : comprenant la taille du noyau, le pléomorphisme, l'anisocaryose, l'aspect du noyau
- *l'index mitotique* : c'est le nombre de mitoses par mm², évalué sur 10 champs au fort grossissement x 400

Différenciation architecturale	proportion de formation glandulaire dans la tumeur
	1 > 75% de glandes
	2 10 - 75 % de glandes
	3 < 10 % de glandes
Pléomorphisme nucléaire	quantification des atypies cyto-nucléaires
	1 cellules petites, régulières et uniformes
	2 augmentation modérée de la taille et du pléomorphisme nucléaire
	3 pléomorphisme marqué
Index mitotique	nombre de mitoses par mm ² , évalué sur 10 champs au fort grossissement x 400 (ici : pour 10 champs de diamètre de champs 0,54 mm)
	1 < 9 mitoses
	2 9 à 16 mitoses
	3 > 16 mitoses

Tableau 3 : détail du grade histopronostique de Nottingham

Le grade histopronostique est donc un score global compris entre 3 et 9, et séparé en 3 catégories : grade 1 (scores 3, 4, 5), grade 2 (scores 6, 7) et grade 3 (scores 8, 9) ; les grades élevés étant de moins bon pronostic. Dans l'étude de Rakha et al. publiée en 2008 (25), malgré son association à certains critères pronostics comme la taille tumorale, le statut ganglionnaire, l'âge de la patiente, l'invasion vasculaire et

l'expression des récepteurs aux oestrogènes, le grade histopronostique de Nottingham possède également une valeur pronostique indépendante. Dans cette étude, la survie à 10 ans liée au cancer était supérieure à 90 % pour les tumeurs de grade 1, de 90 % pour les tumeurs de grade 2 et de 70% pour les tumeurs de grade 3. Un fait intéressant dans cette étude était la différence de pronostic entre les tumeurs de grade 2 de stade ganglionnaire pN1 (1 à 3 ganglions axillaires ipsilatéraux métastatiques) qui ont un meilleur pronostic que les tumeurs de grade 3 de stade ganglionnaire pNo.

L'équipe de Nottingham a également proposé un index (le Nottingham Prognostic Index ou NPI) qui est une combinaison des 3 facteurs pronostiques majeurs (pT, pN et grade histologique) permettant d'améliorer la décision thérapeutique (26). Il se calcule selon la formule suivante :

$$\text{NPI} = \text{Taille (cm)} \times 0,2 + \text{nb de ganglions métastatiques (1 à 3)} + \text{Grade histopronostique (1 à 3)}$$

Cet index pronostique a été validé sur plusieurs séries indépendantes (27, 28) et est utilisé en pratique courante au Royaume Uni.

Concernant la reproductibilité inter-observateur, elle a été évaluée dans plusieurs études entre 50 et 89% de concordance (29 - 33) avec une amélioration quand les pathologistes bénéficient d'une aide au grading (31). Dans ces études, il est important de noter que les corrélations étaient plutôt bonnes pour les grades 1 et 3, et moins bonnes pour les grades intermédiaires. Également, parmi les critères évalués dans le grade, l'évaluation de la différenciation glandulaire semble avoir une meilleure reproductibilité que l'évaluation du grade nucléaire et de l'index mitotique.

Invasion vasculaire

L'invasion vasculaire, lymphatique et sanguine, est l'objectivation en analyse histologique d'embolus tumoraux dans la paroi de vaisseaux, en périphérie de la tumeur. L'invasion vasculaire est corrélée à la taille, au statut ganglionnaire, au grade histopronostique, au type histologique et à l'âge du patient (34, 35). Elle est également un facteur pronostique en termes de récurrence (locale, locorégionale et à distance), de survie spécifique et de survie globale (36, 37). En ce qui concerne la reproductibilité inter-observateurs, il semble qu'il est possible d'atteindre de bonnes performances avec des critères diagnostiques définis préalablement (38).

L'infiltration lymphocytaire de la tumeur (TILS)

Une prolifération néoplasique altère l'équilibre ou la structure du tissu, et entraîne une réponse immunitaire pour la contrer. La majorité des tumeurs prises en charge, le sont à la phase d'échappement de la tumeur face au système immunitaire. Même à des stades avancés d'une maladie tumorale, les caractéristiques de la réponse immunologique telles que le type, la densité et la localisation de l'infiltrat inflammatoire semblent influencer le pronostic (39).

Dans les tumeurs du sein, le pourcentage de TILS semble être un facteur prédictif dans les tumeurs triple négatives (TNBC) et dans les tumeurs surexprimant HER2, mais pas dans les tumeurs dites lumineales, exprimant les récepteurs aux hormones. En effet, pour ces tumeurs (TNBC et HER2+) un pourcentage de TILS élevé est corrélé à un meilleur pronostic et à une meilleure réponse à un traitement néoadjuvant par chimiothérapie ou immunothérapie (40 - 43).

L'utilité des TILS a entraîné un groupe d'experts à établir des recommandations pour standardiser leur évaluation en 2014 (44). L'évaluation s'effectue comme un pourcentage d'occupation du stroma (dénominateur) par l'infiltrat lymphoïde (numérateur). Ils préconisent d'évaluer les TILS sur les zones périphériques de la tumeur et de ne pas se focaliser sur un hotspot mais de faire une évaluation globale. Il convient de prendre en compte toute cellule inflammatoire mononucléée sans les polynucléaires et de ne pas retenir les cellules en dehors de la zone tumorale et dans les zones de carcinome in situ. En ce qui concerne la précision, ils recommandent d'être le plus précis possible, étant donné l'absence de valeur seuil pour l'impact clinique. Dans cette même publication, il semble qu'aucune étude n'ait encore montré de différence intra ou inter-personnelle dans l'évaluation des TILS.

1.2.3 - Caractéristiques immunohistochimiques

Parmi les caractéristiques immunohistochimiques utilisées pour prédire le pronostic ou la réponse à un traitement, figurent l'index de prolifération Ki67, la présence des récepteurs hormonaux aux oestrogènes et à la progestérone (RE, RP), et l'expression de la protéine HER2.

Index de prolifération Ki67

Le Ki67 est une protéine nucléaire impliquée dans les phases précoces de synthèse des ARN ribosomiaux. Il est exprimé lors des phases du cycle cellulaire G₁, S, G₂ et M. Son expression est mesurée par immunohistochimie (IHC) et le résultat est exprimé en pourcentage de cellules tumorales marquées par rapport aux cellules tumorales examinées. Ce marqueur est utilisé depuis plus de 20 ans et de nombreuses études ont validé son rôle dans la prédiction du pronostic (45 - 47) et la prédiction d'un meilleur bénéfice d'une chimiothérapie adjuvante dans les valeurs de Ki67 élevées (48).

Expression des récepteurs hormonaux

La détermination de l'expression des récepteurs hormonaux (RH) aux oestrogènes (RE) et à la progestérone (RP) s'effectue par immunohistochimie, avec une technique largement répandue et standardisée. Elle permet d'évaluer le pourcentage de cellules marquées ainsi que l'intensité du marquage (faible, modérée, forte) afin d'élaborer un score de positivité. Le seuil de positivité varie en fonction des pays ; il est de 10% en Europe (quelle que soit l'intensité du marquage) et de 1% aux Etats-Unis (utilisant un système de scoring sensiblement différent). L'intérêt de ce marqueur est double.

L'expression des RH est tout d'abord un facteur pronostique indépendant (49). La coexpression des 2 récepteurs est un facteur de bon pronostic, et ce d'autant plus que le niveau d'expression est élevé ; tandis qu'une expression dissociée des récepteurs (RE+ RP-) est de moins bon pronostic (50, 51). Au contraire, une absence d'expression de ces deux récepteurs est corrélée à un pronostic défavorable (52).

Outre son rôle pronostique, l'expression des RH est également un facteur prédictif de réponse à l'hormonothérapie. Environ 60% des tumeurs exprimant le RE répondent initialement aux traitements hormonaux et en particulier au tamoxifène, contre moins de 10% de réponse en l'absence d'expression des RH (53). Le niveau d'expression des récepteurs est également corrélé à la qualité de la réponse, avec des variations allant

de 42 à 63 % de réponse objective, entre les différents niveaux d'intensité d'expression (54). La coexpression des 2 récepteurs est également prédictive d'une meilleure réponse en situation métastatique (55, 56).

Surexpression de HER2

HER2 est un oncogène découvert en 1985, et dont le produit de sa traduction est une glycoprotéine transmembranaire de la famille des récepteurs de type I des facteurs de croissance. Une amplification ou une activation du signal provenant de HER2 entraîne une stimulation de l'activité mitotique (57). Environ 13 à 20% des cancers du sein présentent une amplification du gène HER2 accompagnée d'une surexpression de la protéine (58). Ces tumeurs présentent un pronostic défavorable de par un grade histopronostique plus élevé, des rechutes précoces, une évolution métastatique rapide avec une atteinte ganglionnaire plus fréquente et une sensibilité moindre à l'hormonothérapie (59 - 63). Cependant, l'avènement des thérapies ciblées anti-HER2 est venue contrebalancer ces pronostics avec des résultats spectaculaires de l'ordre de 50% d'amélioration de la survie sans récurrence (64, 65). Cette thérapie est désormais prescrite aux patientes présentant un cancer du sein HER2+ en phase adjuvante, en association à la chimiothérapie, et le statut HER2 est donc devenu un critère prédictif de réponse au traitement.

1.2.4 - Les limites des CCPT

Une chimiothérapie peut être administrée dans le but de réduire le risque de récurrence du cancer du sein. Cette décision prend en compte le degré d'extension et le profil biologique de la tumeur ainsi que l'état général et les préférences de la patiente.

Le protocole thérapeutique est bien défini dans certaines situations :

- Pour les tumeurs métastatiques d'emblée, les tumeurs surexprimant HER2, les tumeurs triple négatives et les tumeurs RH+/HER2- à risque élevé de récurrence, une chimiothérapie adjuvante et/ou néoadjuvante sera le plus souvent recommandée
- A l'inverse, pour les tumeurs RH+/HER2- à risque faible de récurrence (pT1a-b), une chimiothérapie ne sera pas recommandée

Il existe donc une catégorie de tumeurs dites de risque intermédiaire de récurrence (les tumeurs RH+/HER2-, pT1c - pT2, pN0 ou pN1mi - pN1a) pour laquelle le choix d'une chimiothérapie n'est pas bien défini, discuté au cas par cas en réunion de concertation pluri-disciplinaire (RCP) et laissé à l'appréciation des cliniciens et de la patiente. Ce qui semble créer une hétérogénéité des pratiques entre les centres (66). Pour ces tumeurs, 5 critères de mauvais pronostic, recommandés par la haute autorité de santé (HAS) sont utilisés (67) :

- L'envahissement ganglionnaire (pN1) ;
- L'âge jeune de la patiente (< 35/40 ans) ;
- La taille tumorale pT2 (T >2-3 cm) ;
- La présence d'emboles lymphovasculaires ;
- Le grade 3 histopronostique

L'existence d'au moins un de ces CCPT de mauvais pronostic permet habituellement de retenir l'indication d'une chimiothérapie adjuvante. Il en découle un risque de surtraitement de patientes qui n'auraient pas eu besoin de chimiothérapie, et qui les expose alors à une morbidité liée aux effets indésirables de celle-ci.

Cette difficulté à définir des recommandations précises pour ces tumeurs à risque intermédiaire prend son origine dans le manque de précision de la valeur pronostique des CCPT. Par exemple, 30% des patientes avec envahissement ganglionnaire (pN1-pN2) ne récidivent pas après un traitement locorégional adéquat sans traitement systémique (68). Dans cette étude, malgré l'association de certains CCPT à la survie des patientes, aucun ne pouvait clairement identifier celles qui ne récidivent pas. Par ailleurs, la majorité des tumeurs du sein (30 à 60 %) sont de grade histopronostique intermédiaire. Or, nous avons vu précédemment que cette catégorie était la moins reproductible des 3. De plus, dans l'étude de Sotiriou et al. sur le genomic grade index qui est un profil moléculaire corrélé au grade (69), il semblerait que l'ensemble des tumeurs de grade 2 soit en fait constitué de tumeurs de profil moléculaire grade 1, de tumeurs de profil moléculaire grade 3, et de tumeurs de profil moléculaire mixte exprimant des gènes du grade 1 et des gènes du grade 3.

Dans cette situation ambiguë, l'utilisation de modèles multivariés de prédiction combinant plusieurs CCPT indépendants a été envisagée (70, 71). Deux modèles sont prédominants dans la littérature : Adjuvant ! Online (abandonné aujourd'hui) et PREDICT.

- Adjuvant ! Online a été développé au début des années 2000, puis amélioré par la suite, dans le but de prédire le risque de récurrence et de mortalité à 10 ans, avec ou sans traitement systémique adjuvant. Il prend en compte plusieurs marqueurs cliniques, biologiques et anatomo-pathologiques : l'âge, les comorbidités, la taille tumorale, le grade histopronostique, l'expression du RE, le nombre de ganglions axillaires atteints. Mais malgré sa facilité d'utilisation, cet outil est imprécis et semble surestimer la survie (72, 73), possiblement à cause de biais dans la méthodologie des études, et l'absence de prise en compte d'autres facteurs pronostiques importants comme l'expression du RP et de HER2, la présence d'embolies vasculaires, l'index de prolifération Ki67 et l'absence de prise en compte du degré d'expression des récepteurs hormonaux.
- PREDICT est disponible depuis 2010 avec quelques révisions effectuées. Il fournit une estimation de la survie globale à 5 et 10 ans avec et sans traitement systémique adjuvant. Il combine plusieurs variables : l'âge au diagnostic, les modalités de diagnostic et du traitement, la taille tumorale, le statut ganglionnaire, le grade histopronostique, l'expression du RE, de HER2, et la valeur du Ki67.

D'autres scores ont également été développés, qui ne sont pas utilisés en pratique courante à l'heure actuelle :

- NPI+ est un algorithme développé plus récemment en 2014. Il se base sur 10 marqueurs immunohistochimiques (RE, RP, CK5/6, CK7/8, HER1, HER2, HER3, HER4, P53 et Mucin1) permettant de classer les tumeurs en 7 sous-types biologiques : 3 luminaux (A, B et N), 2 basaux (P53 normal ou muté) et 2 HER2 (RE+/HER2+, RE-/HER2+). Ces sous-types sont ensuite stratifiés en différents groupes pronostiques par un algorithme combinant la taille tumorale, le statut ganglionnaire, le grade histopronostique et les embolies vasculaires (74). Outre sa meilleure stratification du pronostic par rapport au NPI classique dans cette

étude, sa reproductibilité et sa performance ont par la suite été prouvés sur des séries indépendantes (75, 76).

- IHC4 est un outil développé en 2011 qui utilise de manière semi-quantitative les marqueurs immunohistochimiques RE, RP, HER2 et Ki67, sous la forme d'un score calculé de cette manière :
$$\text{IHC4} = 94.7 \times \{-0.100 \text{ ER10} - 0.079 \text{ PgR10} + 0.586 \text{ HER2} + 0.240 \ln(1 + 10 \times \text{Ki67})\}$$

En association à un score clinique comprenant la taille tumorale, le statut ganglionnaire, le grade histopronostique et le type d'hormonothérapie administrée (inhibiteur d'aromatase ou tamoxifène), il fournit un score pronostique similaire à celui donné par la signature multigénique Oncotype DX® (77).

- Le score CTS (Clinical Treatment Score) est calculé à partir de données cliniques et anatomopathologiques telles que l'âge, le traitement hormonal administré, la taille tumorale et le statut ganglionnaire. Il a prouvé son potentiel pronostique pour l'estimation des récurrences à distance après 5 ans d'hormonothérapie (78, 79). Il a également été utilisé en combinaison avec le score IHC4 pour prédire le risque de récurrence locorégionale de patientes n'ayant pas bénéficié de radiothérapie externe adjuvante (80).

L'utilisation de ces outils est néanmoins peu fréquente bien qu'elle ait été récemment encouragée par l'American Joint Committee on Cancer (81). En France, l'INCa ne recommande pas leur utilisation en routine (82). Des études sont en cours pour le développement et la validation de ces modèles en vue d'une utilisation en pratique courante (83 - 85).

Les CCPT montrent donc des limites dans la prédiction de survie, de récurrence et de bénéfice d'un traitement adjuvant, dans cette classe de tumeurs dites de risque intermédiaire. C'est pour affiner cette prédiction et mieux individualiser la prise en charge thérapeutique de ces patientes que les signatures moléculaires se sont développées.

1.3 - Caractéristiques moléculaires : les signatures moléculaires

En 2000, Perou et Sorlie (86 - 88) ont élaboré une classification moléculaire des tumeurs du sein basée sur le niveau d'expression des gènes, analysés en technique cDNA array. Cette classification moléculaire a défini cinq grands groupes :

- Luminal A
- Luminal B
- HER2-enrichi
- Basal-like
- Normal breast-like

Ces cinq groupes peuvent être séparés en 2 catégories : les tumeurs lumineales (A et B) exprimant le gène *ESR1* codant pour le RE ; et les tumeurs n'exprimant pas ce gène (HER2-enrichi, basal-like et normal breast-like). Ces catégories représentent respectivement 80 % et 20 % des tumeurs infiltrantes du sein. Des études ont par la suite démontré l'intérêt pronostique de cette classification (87 - 89), les tumeurs lumineales A ayant un meilleur pronostic que les autres. Dans ces tumeurs lumineales, les lumineales A expriment fortement les gènes de la voie du RE et faiblement les gènes impliqués dans la prolifération cellulaire, contrairement aux tumeurs lumineales B.

Suite à ces travaux, des signatures moléculaires ont été développées pour calculer des scores pronostiques, basés sur l'analyse du niveau d'expression d'un panel de gènes plus restreint, par la quantification de leurs ARN messagers (90, 91). Elles ont ensuite été affinées pour aider les cliniciens dans la décision d'une chimiothérapie adjuvante, particulièrement dans ces situations d'ambiguïté décisionnelle, en identifiant les patientes pour lesquelles une chimiothérapie serait bénéfique (92). Quatre signatures génomiques sont inscrites au référentiel des actes innovants hors nomenclatures (RIHN) en France, et sont actuellement financées par un acte global (N137) dans le cancer du sein : Oncotype Dx®, Mammaprint®, Prosigna®, Endopredict®. Les trois signatures d'expression génique les plus utilisées en France sont Oncotype DX®, Prosigna® PAM50 et Endopredict®. Cependant, seule la signature Prosigna® PAM50 (93) permet de donner la classe moléculaire (Luminal A ou B, HER2-enrichi ou Basal-like). Les deux signatures actuellement utilisées à l'Institut Universitaire du Cancer de

Toulouse sur la base des données scientifiques les plus récentes sont Oncotype DX® et Prosigna® PAM50.

La signature Prosigna® PAM50 se base sur le niveau d'expression de 50 gènes, impliqués dans l'expression du RE, de HER2 et dans la prolifération, analysés par RT-PCR quantitative à partir d'ARN extraits de tumeurs fixées en formol et inclus en paraffine (93). Le résultat fourni comprend 4 données (annexe 1) :

- Le sous-type intrinsèque : luminal A, luminal B, HER2 enrichi et basal-like
- Le score ROR (Risk Of Recurrence) : un risque de récurrence calculé à partir du profil moléculaire, couplé à 2 variables cliniques que sont le pT et le pN
- La probabilité individuelle de récurrence à 10 ans après 5 ans d'hormonothérapie
- La catégorie de risque de récurrence à 10 ans : faible, moyen et haut risque

Ces signatures sont utilisées dans les cas complexes de tumeurs lumineuses de risque intermédiaire décrites précédemment, afin d'orienter la décision d'une chimiothérapie adjuvante. Dans son rapport (94) l'HAS recommande d'utiliser ces signatures pour des cas équivoques d'adénocarcinomes infiltrants (préférentiellement NST), exprimant les RE, ayant un degré de prolifération intermédiaire, classé pT1c-2, pNo-N1mi, de grade histopronostique 2 et sans critère de mauvais pronostic (âge jeune et embolies lymphovasculaires). Elle permet de séparer ces tumeurs en 2 catégories : lumineuses A et B. Les tumeurs lumineuses B étant de moins bon pronostic, elles seront plus susceptibles de recevoir une chimiothérapie.

Il est important de noter qu'une partie importante des gènes étudiés dans les tumeurs lumineuses sont des gènes impliqués dans la prolifération. Également, le niveau d'expression des gènes mesuré a une répartition continue, gaussienne, et la distinction de ces deux sous-types est donc un seuil fixé arbitrairement (95).

Dans beaucoup de laboratoires, ces tests ne sont pas accessibles pour des raisons financières ou logistiques. C'est pourquoi, la prédiction de ces sous-types par d'autres méthodes plus répandues, notamment l'immunohistochimie, a été étudiée. Pour cela, les marqueurs d'intérêt qui étaient étudiés étaient l'expression des récepteurs de la progestérone ainsi que le Ki67 (marqueur de choix compte tenu de la forte implication des gènes de la prolifération dans les signatures moléculaires). La conférence de

consensus de St Gallen de 2013 (96) proposait un seuil de 20 % pour ces 2 marqueurs pour distinguer les luminal A-like et les luminal B-like ; une tumeur luminal B-like étant une tumeur présentant un $Ki67 \geq 20\%$ ou un $RP \leq 20\%$. Mais si des valeurs élevées ou faibles de Ki67 sont relativement fiables pour distinguer ces deux sous-types, la définition d'un seuil dans les valeurs intermédiaires semble difficile (97). Ce qui a entraîné une variation dans les recommandations successives faites dans les conférences de consensus de St Gallen, qui proposaient un seuil à 14% en 2011, puis à 20% en 2013, avant de cesser de proposer un seuil en 2015 (98 - 100). En 2021, le International Ki67 in Breast Cancer Working Group (IKWG) proposa d'utiliser le Ki67 comme facteur prédictif de bénéfice d'une chimiothérapie que pour des valeurs en dessous de 5 % ou au-dessus de 30 % (101). Ce manque de précision trouve son origine d'une part dans le manque de reproductibilité et de standardisation de l'évaluation du Ki67 par les pathologistes, et d'autre part dans la variabilité des scores des tumeurs lumineuses dûe à l'hétérogénéité d'expression des gènes de la prolifération (102). Les seuils proposés dans la conférence de St Gallen ont été retravaillés par l'équipe de Maisonneuve et al. qui proposait en 2014, de définir les tumeurs lumineuses B par un $Ki67 \geq 20\%$ ou un $Ki67$ compris entre 14 et 20% associé à un $RP < 20\%$ (103). Par rapport à la définition de St Gallen, les tumeurs RE+ et RP- (<20%) qui auraient été classées comme lumineuses B, sont ici reclassées en lumineuses A si elles présentent un $Ki67 < 14\%$.

1.4 - Prédiction avec l'aide de l'intelligence artificielle

1.4.1 - Définition de l'intelligence artificielle

Le domaine de l'intelligence artificielle (IA) est vaste. Globalement, il regroupe tout processus artificiel capable d'effectuer des actions complexes nécessitant généralement l'intelligence humaine. Le terme d'IA a été popularisé à partir de 1956 par McCarthy suite à la conférence "Dartmouth Summer Research Project on Artificial Intelligence", après le développement du premier modèle mathématique d'un réseau de neurones en 1943 par McCulloch et Pitts, et la création du premier ordinateur à réseau de neurones en 1950 par Minsky et Edmonds.

Au fil des années, l'utilisation de cet outil informatique s'est largement répandue dans de nombreux domaines, notamment dans l'analyse d'image et par conséquent en anatomopathologie. Un des premiers articles traitant d'analyse d'image de cellules est

celui de Mendelsohn et coll. et date de 1965 (104). Mais ce n'est que depuis les années 90 que l'analyse d'image à l'aide de l'IA connaît un regain d'intérêt grâce au développement technologique (augmentation des puissances de calcul avec le développement des graphics processing unit ou GPU), et qu'elle se popularisera en anatomopathologie à partir des années 2000, du fait du développement croissant de la numérisation par scanner de lames (105, 106).

Parmi les méthodes utilisées en IA, on distingue 3 grandes catégories de systèmes qui ont successivement été développées : les systèmes experts, le machine learning et le deep learning.

1.4.2 - Les systèmes experts

Les systèmes experts ont été les premiers à être développés. Ce sont des outils capables de reproduire le comportement d'un expert humain dans un domaine ou une tâche donnée. C'est un système qui traite des faits connus, avec un système de règles définies et implémentées au préalable. Un exemple en médecine pourrait être un algorithme d'aide au diagnostic utilisant des règles à la façon d'un arbre diagnostique. Les performances de ces modèles sont bonnes dans un champ d'utilisation restreint, mais diminuent avec l'augmentation du volume de données et l'augmentation des paramètres. D'autant que le paramétrage d'un tel système nécessite beaucoup de travail.

1.4.3 - Le machine learning (ML)

Les ambitions suivantes ont donc été de créer une machine capable d'apprendre par elle-même ; c'est ce qu'on appelle l'apprentissage autonome ou Machine Learning. C'est un système qui construit un modèle prédictif, à partir de données d'apprentissage. Il se distingue du système expert par le fait qu'il ne possède pas de règle définie au préalable, et que ce sont les données fournies en apprentissage qui serviront à créer ces règles. Le ML est en quelque sorte le nouveau nom de la statistique. En anatomopathologie, de nombreuses études ont utilisé des techniques de ML pour comparer des caractéristiques morphologiques extraites au préalable, comme des statistiques de formes et textures des noyaux, du stroma ou des différents tissus présents sur la lame.

En pathologie mammaire et pour des tâches de routine déjà réalisées par le pathologiste, ces méthodes ont montré de bonnes performances dans la segmentation de glandes (107), de noyaux (108), et la détection de mitoses (109). Dans des tâches plus complexes, J. Diamond et al. ont également créé un algorithme pour différencier sur des lames HE, les zones d'adénocarcinome prostatique en comparant les textures des zones normales et pathologiques (110)

Les méthodes de ML ont également été utilisées pour des prédictions de survie. Kong et al. ont analysé les caractéristiques morphométriques des noyaux tumoraux de glioblastomes par le biais de 74 caractéristiques morphologiques (couramment appelées features) sélectionnées au préalable. Ils les ont ensuite comparées aux données de survie et aux analyses moléculaires, et ont ainsi pu mettre en évidence différents clusters cellulaires morphométriques en corrélation avec les profils moléculaires et les différences de survie (111).

Dans les tumeurs du poumon, Yu et al. ont extrait et analysé une multitude de caractéristiques morphométriques (de taille, de forme et de texture) se focalisant sur l'aspect des noyaux et des cytoplasmes (112). Ils ont montré une corrélation entre certaines de ces features et la survie des patients dans les groupes adénocarcinomes et carcinomes épidermoïdes. Il est intéressant de noter dans cette étude que les features les plus importantes étaient différentes selon la tâche de prédiction à accomplir (60 features pour les adénocarcinomes et 15 features pour les carcinomes épidermoïdes). De plus, il semblerait que ces features soient difficilement identifiables à l'œil nu malgré leur réalité informatique. Un travail similaire a été réalisé par l'équipe de Wang et al. (113) avec une sélection de features plus restreintes et ne concernant que la taille et la forme des régions tumorales préalablement segmentées. Sur les 22 features analysées, 15 étaient significativement associées à la survie, à partir desquelles ils ont ensuite pu développer un modèle prédictif de la survie.

Dans les tumeurs du sein, Veta et al. (114) ont analysé la valeur pronostique de 2 paramètres morphométriques des noyaux que sont la moyenne et l'écart-type de l'aire du noyau, en les mesurant sur 101 tumeurs du sein chez l'homme. Ils ont ainsi pu mettre en évidence que la moyenne de l'aire des noyaux est un facteur pronostique indépendant, alors que l'écart-type ne l'est pas. Dans une autre étude évaluant un paramètre morphométrique, Tambasco et al. (115) ont montré que la complexité morphologique des tumeurs du sein, mesurée par la dimension fractale, était un

facteur pronostique indépendant de la survie à 10 ans. Ils mesuraient la dimension fractale sur des TMA (tissu microarray) de tumeurs du sein colorées par immunohistochimie anti-pancytokératine. Dans un niveau de complexité supérieur, C. Lu et al. (116) ont extrait et comparé des caractéristiques morphologiques de 276 tumeurs du sein. Sur les 615 features analysées, 15 étaient plus significatives et ont été retenues pour créer le classifieur, qui séparaient les tumeurs en 2 groupes (bas et haut risque) avec une précision de 65%, une sensibilité de 57,41% et une spécificité de 73,91%. Un point remarquable dans cette étude est que la majorité des features retenues et celles qui sont les plus discriminantes, sont des features liées à l'orientation des noyaux. Yuan et al. (117) ont quantifié différentes caractéristiques de l'hétérogénéité tumorale de tumeurs du sein en se concentrant sur les noyaux des cellules tumorales, stromales, épithéliales bénignes, et des lymphocytes. Avec ces données, ils ont pu affiner les résultats des tests moléculaires, mettre en évidence que l'infiltration lymphocytaire est prédictive d'une meilleure survie dans les cas RE négatifs et qu'une certaine distribution spatiale des cellules stromales était un facteur pronostic indépendant dans cette même catégorie de tumeur.

Ces approches en ML permettent d'analyser la valeur de critères morphologiques déjà utilisés en les quantifiant plus précisément. Cependant, ces approches demandent souvent beaucoup de travail d'annotations, de sélection des features les plus pertinentes, et les analyses qui en découlent se concentrent généralement uniquement sur des détails à petite échelle (cellule par cellule) sans prendre en compte la globalité.

1.4.4 - Le deep learning (DL)

Le Deep Learning est un processus de Machine Learning qui s'appuie sur l'utilisation de réseaux de neurones artificiels ou neural networks (NN). Les NN sont des systèmes informatiques inspirés de l'organisation histologique du cerveau, dans lesquels de multiples unités (neurones artificiels) sont liées entre elles et organisées en couches successives. Chaque neurone artificiel est une fonction mathématique qui reçoit une valeur d'entrée (input), la pondère avec des coefficients, et produit une valeur de sortie (output) qui sera transmise à un autre neurone. Il existe différentes architectures de réseaux de neurones. La plus utilisée en pathologie digitale est celle des Convolutional Neural Networks (CNN), ayant montré des performances supérieures dans les tâches de classification (118, 119). Un CNN va effectuer une opération de convolution de

l'image et l'efficacité de celle-ci est liée au fait que les neurones sont liés de proche en proche afin que l'ensemble puisse détecter des motifs locaux. Dans ces réseaux, les données non structurées de l'image (les pixels) sont analysées successivement par les différentes couches, qui en extraient des motifs (ou features) qui seront eux-mêmes analysés par d'autres couches. Ces features sont déterminées et sélectionnées de manière autonome par l'algorithme au fil de l'apprentissage, en fonction de leur importance pour la tâche à réaliser.

Dans les mêmes tâches de routine évoquées plus haut, les techniques de DL ont montré de très bons résultats comme dans la segmentation et la classification de tissus (120), la détection de tumeur (121 - 124), la détection de mitoses (125) ou l'automatisation de l'analyse en immunohistochimie (126). Et dans certaines situations, les algorithmes utilisés avaient une meilleure performance qu'un expert humain comme par exemple dans la détection de métastases dans des ganglions lymphatiques, où les chercheurs ont réussi à atteindre une sensibilité de 92,4% avec l'algorithme, contre une sensibilité de 73,2% pour un pathologiste (127). Dans un niveau de complexité supérieur et pour des tâches effectuées en routine par un pathologiste, des algorithmes de grading de cancer de prostate ont également été développés, atteignant des performances similaires à l'humain (128, 129).

En pathologie mammaire, l'utilisation du DL a permis d'atteindre de bonnes performances dans des tâches de prédiction de classes bénin vs. malin avec des accuracy d'au moins 95% ; et également dans des tâches de prédiction multi-classes de différentes lésions bénignes avec des accuracy d'au moins 90% (130 - 132).

Mais si ces approches présentent des avantages évidents par rapport aux techniques de ML, elles présentent certaines limites. Par exemple, les performances d'un algorithme DL sont dépendantes du nombre de données en apprentissage : elles seront médiocres sur peu de données et augmenteront avec la taille du dataset (133).

Par ailleurs, du fait d'une plus grande autonomie laissée à l'algorithme notamment dans la sélection de features pertinentes, et du fait d'une moindre possibilité de contrôle externe sur ces processus, il en résulte un effet "boîte noire" menant à des difficultés de maîtrise des résultats. J. Su et al. dans leur étude (134) révèle la fragilité de ces outils avec l'exemple qu'un changement d'un seul pixel dans une image peut altérer la classification de plusieurs modèles de DL différents. Dans l'article de Seyyed-

Kalantari et al. (135) les auteurs montrent que les erreurs diagnostiques faites par l'algorithme sur des radiographies de thorax, concernent en majorité les patients issus de population défavorisée ayant un accès restreint aux soins. Ces erreurs sont liées aux biais présents au début, dans le dataset d'apprentissage qui servira à construire le modèle. Si le dataset d'apprentissage est biaisé par un déséquilibre de classes, les prédictions de l'algorithme le seront également. Ce type d'erreur concerne plus largement le développement de tout nouveau dispositif médical qui sera testé et approuvé principalement sur la classe dominante (les hommes blancs) comme montré dans cet article paru dans The Guardian (136).

Un autre problème se pose lors de l'utilisation du modèle entraîné sur un set de données indépendant. Malgré un nombre important de travaux publiés, rares sont les modèles qui sont utilisés en pratique courante. Le manque de généralisation des résultats en est une des raisons principales (137). Différents facteurs variant d'un centre à l'autre peuvent influencer le processus d'entraînement et d'apprentissage de l'algorithme, comme par exemple les différentes étapes techniques anatomopathologiques qui auront un impact sur les caractéristiques de l'image scannées (variation de coloration, épaisseur de coupe, modèle de scanner ...). Pour éviter ces désagréments, certains auteurs ont conçu des cadres de travail mêlant des combinaisons de techniques informatiques pour améliorer la généralisation (138). L'absence de déploiement d'outil d'IA à grande échelle en anatomopathologie depuis 2012, l'année de la première victoire d'un challenge de classification (ImageNet Large Scale Visual Recognition Competition) par un réseau de neurones (AlexNet), reflète bien ces difficultés d'adaptation de ces méthodes à la pratique courante.

En raison de ces biais potentiels, certains chercheurs préfèrent se baser sur des méthodes de ML dans lesquelles les features sont extraites au préalable. C'est notamment le cas de l'équipe de Diao et al. qui, dans leur étude (139), ont combiné une annotation de WSI (whole slide images) à grande échelle par DL, à des analyses de features par technique de ML. La première étape de leur étude consistait à entraîner un algorithme de DL à reconnaître et classer les différents éléments présents sur les lames colorées en hémalun-éosine (fibroblastes, plasmocytes, lymphocytes, macrophages, cellules tumorales, tissu cancéreux, nécrose, stroma associé à la tumeur) de 5 types de tumeurs (mélanome, cancer gastrique, cancer du sein, adénocarcinome pulmonaire et carcinome épidermoïde pulmonaire). Cet algorithme était ensuite

capable d'annoter à grande échelle les WSI. Grâce à ces annotations, il était possible d'extraire 607 Human Interpretable Features (HIFs) qui étaient divisés en 6 catégories : comptes et densités cellulaires, caractéristiques des clusters cellulaires, interactions entre les différents types cellulaires, aire et quantification des types de tissus, paramètres architecturaux des différents tissus et caractéristiques morphométriques de ces tissus. Les analyses de ces features ont montré des corrélations avec le type tumoral (sauf pour la différenciation adénocarcinome et carcinome épidermoïde du poumon), avec l'expression de certains checkpoints immunitaires (PD-1, PD-L1, CTLA4 et TIGI) et avec certains phénotypes moléculaires. Ces features ont l'avantage d'être biologiquement pertinentes et compréhensibles pour l'œil humain, ce qui permettait une meilleure interprétabilité des résultats. Ce genre d'approche permet de dépasser les compétences humaines classiques tout en évitant les désagréments de l'effet "boîte noire" en restant dans un champ d'analyse restreint et compréhensible.

Ces différents travaux montrent que l'extraction de caractéristiques morphométriques et leur analyse par des méthodes d'IA peuvent être utiles dans la prédiction du pronostic.

1.4.5 - Prédiction des altérations moléculaires

A l'heure d'un intérêt croissant pour les analyses "omiques" des caractéristiques tumorales (génomique, transcriptomique, radiomique...) les images histologiques, de part la multitude d'images sous-visuelles qui les composent, deviennent des ensembles de données propices à cette approche. Et dans l'hypothèse que les caractéristiques moléculaires des tumeurs ont une expression morphologique, certains chercheurs ont étudié la prédiction des altérations moléculaires à l'aide de méthodes d'IA.

Dans les tumeurs pulmonaires, N. Coudray et al. (140) ont par exemple entraîné un réseau de neurones à reconnaître des mutations de 6 différents gènes (STK11, EGFR, FAT1, SETBP1, KRAS et TP53) sur des lames colorées en hémalun-éosine (HE), avec des résultats d'AUC allant de 0.733 à 0.856.

Dans les tumeurs du sein, H.D. Couture et al. (141) ont entraîné un réseau de neurones à prédire le statut moléculaire basal-like contre non basal-like ainsi que le risque de récurrence ROR-PT, obtenant une accuracy de 75 à 80%. Dans une autre étude, Jaber et al. (142) ont entraîné un réseau de neurones à prédire les 4 sous-types intrinsèques

donnés par le test Prosigna® PAM50 (luminal A, luminal B, HER2-enrichi et Basal-like), obtenant une accuracy sur lame entière d'environ 66%. Les erreurs de classification dans cette étude étaient plus fréquentes pour les tumeurs lumineales B qui étaient classées en lumineales A ou en autre catégorie. Pour les catégories de tumeurs lumineales, on observait une accuracy d'environ 54,6% pour la distinction des sous-types A et B.

D'autres équipes, celles de Basavanhally, Romo-Bucheli et Madabhushi, ont travaillé sur la prédiction du risque de récurrence donné par la signature moléculaire Oncotype DX®, combinant des approches de ML et de DL. La signature Oncotype DX® étant, semble-t-elle, fortement corrélée au grade histopronostique (143 - 145) ils ont comparé la valeur prédictive des différentes composantes de ce grade. En 2009 (146), ils ont développé un algorithme de ML s'appuyant sur 12 features nucléaires préalablement extraites, qui séparaient les tumeurs du sein de bas et haut risque de récurrence selon Oncotype DX®, avec une précision d'environ 84%. En 2016 (147) un algorithme DL était entraîné à identifier les noyaux et à différencier ceux formant des tubules, pour effectuer un ratio nombre de noyaux formant des tubules/nombre de noyaux total nommé TFI (Tubule Formation Indicator). Ce ratio était corrélé au score obtenu avec Oncotype DX®, avec un ratio plus élevé dans les tumeurs de bas risque. En 2017 (148) l'algorithme de ML se basait sur le compte mitotique effectué à l'aide d'une détection automatisée des mitoses par un algorithme de DL. Il séparait les tumeurs en bas et haut risque de récurrence avec une précision d'environ 83%. Individuellement, ces features morphologiques montrent déjà des résultats encourageants, et l'utilisation combinée de features différents semble également augmenter la précision de prédiction du risque de récurrence (149).

1.5 - Objectifs de l'étude

L'objectif de notre étude était d'évaluer et de comparer les performances de prédiction du sous-type moléculaire (luminal A et B) à partir d'images microscopiques de cancers du sein en coloration Hémalum-éosine (HE). Une approche par deep learning pure était comparée à une approche par machine learning se basant sur différentes caractéristiques cliniques et pathologiques, ainsi que sur des features morphologiques extraites au préalable, afin d'évaluer l'apport de ces features morphologiques en terme de performances pronostiques et d'interprétabilité des résultats obtenus.

2 - Matériel et méthodes

2.1 - Populations de l'étude

Pour cette étude, nous avons procédé de manière classique pour l'élaboration de modèles d'IA avec l'utilisation de 2 cohortes : une première cohorte pour l'entraînement et la validation des modèles, et une cohorte de test constituée de cas indépendants pour valider les performances du modèle sélectionné en entraînement. L'ensemble des patientes de ces cohortes étaient atteintes d'un cancer du sein ayant été analysé à l'IUCT Oncopole de Toulouse et pour lesquelles une signature d'expression génique Prosigna® PAM50 a été réalisée.

2.1.1 - Cohorte d'entraînement

Pour cette cohorte, nous nous sommes intéressés aux cas de tumeurs du sein testées avec la signature Prosigna® PAM50 à l'IUCT Oncopole de Toulouse, entre décembre 2018 et janvier 2021. Cette cohorte fait partie d'un projet à plus grande échelle nommé APRIORICS et piloté par le Dr Franchet, dont l'objectif est de créer une base de données de lames numérisées de tumeurs du sein, annotées par des marquages immunohistochimiques ciblant différentes structures histologiques (membrane nucléaire, mitoses, lymphocytes ...). Ce projet est financé par le Health Data Hub. Les patientes dont la tumeur du sein était susceptible d'être incluse dans le projet recevaient une information spécifique et un formulaire de non opposition à l'utilisation de leurs données, qu'elle complétait et renvoyait en cas d'opposition à l'utilisation de leurs prélèvements à visée de recherche. Sur les 88 patientes ayant bénéficié d'une analyse moléculaire durant cette période, 8 se sont opposées à l'utilisation de leurs données et ont donc été exclues de l'étude. Par la suite, au moment de la numérisation des lames pour la partie analyse d'image de ce travail de thèse, 14 cas n'avaient pas encore été pris en charge dans le cadre d'APRIORICS et ont donc été exclus. La cohorte d'entraînement comprenait donc 80 tumeurs, avec 66 lames HE disponibles pour l'analyse d'image.

2.1.2 - Cohorte de test

Pour cette cohorte, nous nous sommes intéressés aux cas de tumeurs du sein ayant bénéficié du test moléculaire Prosigna® PAM50 à l'IUCT Oncopole de Toulouse, entre février 2021 et février 2022. Cette cohorte était indépendante du projet APRIORICS. Sur les 59 tumeurs qui étaient susceptibles d'être incluses dans la cohorte, 32 ont été exclues en raison soit d'une absence de la lame HE dans les archives du service d'anatomopathologie, soit de l'absence du recueil de non opposition à l'utilisation des données personnelles. Ultérieurement, une WSI a été exclue car elle ne comportait que 3 patches analysables du fait d'une décoloration artéfactuelle. La cohorte de test comprenait donc 26 lames HE.

2.1.3 - Description des cohortes

Pour ces cohortes, les critères clinico-pathologiques traditionnels ainsi que les résultats des tests moléculaires étaient extraits des comptes rendus DIAMIC. Les caractéristiques de ces 2 populations sont présentées ci-dessous (tableau 4). Certains critères ne sont pas présentés comme le statut HER2 et les TILS puisqu'aucune tumeur de la cohorte ne surexprimait HER2 et que l'évaluation des TILS est un facteur pronostique dont l'intérêt n'est démontré que dans les tumeurs triple négatives ou les tumeurs surexprimant HER2. Également, les tumeurs testées en Prosigna® PAM50 étant toutes positives aux récepteurs aux oestrogènes, le statut des récepteurs est présenté en 2 catégories : RE+/RP- et RE+/RP+ avec une positivité du RP définie comme $RP \geq 20\%$ selon les chiffres de l'algorithme de St-Gallen.

	Cohortes d'entraînement		Cohorte de test
	totale	pour l'analyse d'image	(n = 26)
	(n = 80)	(n = 66)	
Sous-type intrinsèque – no. (%)			
• luminal A	37 (46)	29 (44)	12 (46)
• luminal B	43 (54)	37 (56)	14 (54)
Âge (classe) – no. (%)			
• ≤45 ans	15 (19)	12 (18)	0(0)
• >45 ans	65 (81)	54 (82)	26 (100)
Taille tumorale en mm; moyenne [min ; max]	18,35 [9 ; 56]	18,8 [9 ; 56]	18,19 [11 ; 28]
Taille tumorale (pT) – no. (%)			
• pT1b	1 (1)	1 (1)	0 (0)
• pT1c	54 (68)	42 (64)	19 (73)
• pT2	23 (29)	21 (32)	7 (27)
• pT3	2 (3)	2 (3)	0 (0)
Statut ganglionnaire (pN) – no. (%)			
• pNo	61 (76)	47 (71)	23 (88)
• pNo(i+)	1 (1)	1 (2)	1 (4)
• pN1mi	5 (6)	5 (8)	1 (4)
• pN1a	12 (15)	12 (8)	1 (4)
• pN2	1 (1)	1 (2)	0 (0)
Type histologique – no. (%)			
• nst	66 (83)	55 (83)	21 (80,8)
• lobulaire	5 (6)	5 (8)	0 (0)
• mucineux	3 (4)	3 (4,5)	1 (3,9)
• autre	1 (1)	1 (1,5)	0 (0)
• mixte	5 (6)	2 (3)	4 (15,3)
Grade de Nottingham – no. (%)			
• 1	2 (2,5)	2 (3)	0 (0)
• 2	76 (95)	62 (94)	27 (100)
• 3	2 (2,5)	2 (3)	0 (0)
Détail grade : architecture – no. (%)			

• 1	0 (0)	0 (0)	1 (3,9)
• 2	29 (36)	22 (33)	11 (42,3)
• 3	51 (64)	44 (67)	14 (53,8)
Détail grade : atypies nucléaires – no. (%)			
• 1	0 (0)	0 (0)	0 (0)
• 2	72 (90)	59 (89)	23 (88,5)
• 3	8 (10)	7 (11)	3 (11,5)
Détail grade : mitoses – no. (%)			
• 1	37 (46)	31 (47)	6 (23)
• 2	34 (43)	30 (45)	18 (69)
• 3	9 (11)	5 (8)	2 (8)
Détail grade : mitoses / mm² – moyenne [min ; max]	4,3 [0,3 ; 15,9]	4,1 [0,3 ; 15,9]	5 [0,6 ; 12]
Invasion vasculaire – no. (%)			
• oui	14 (17,5)	11 (17)	3 (11,5)
• non	66 (82,5)	55 (83)	23 (88,5)
Récepteurs hormonaux – no. (%)			
• RE+/RP+	62 (77,5)	50 (75,8)	18 (69)
• RE+/RP-	18 (22,5)	16 (24,2)	8 (31)
Risk Of Recurrence (ROR) – moyenne [min ; max]	48,2 [14 ; 82]	49,3 [14 ; 82]	49,8 [26 ; 72]
Probabilité de récurrence à 10 ans – moyenne [min ; max]	13,2 [3 ; 49]	14 [5 ; 49]	12 [4 ; 25]
• données manquantes	1	1	0

Tableau 4 : caractéristiques clinico-pathologiques et moléculaire des populations de l'étude

On remarque que les 3 populations sont homogènes et qu'elles présentent des répartitions sensiblement identiques des sous-types intrinsèques. Les différences notables sont pour l'âge au diagnostic, avec 18 et 19% de patientes en dessous de 45 ans pour les cohortes d'entraînement, contre aucune dans la cohorte de test ; et le nombre de mitoses par mm² qui est plus élevé dans la cohorte de test que dans les cohortes d'entraînement, avec près de 70% des tumeurs classées en grade 2 dans la cohorte de test, contre environ 45% de grade 1 et 45% de grade 2 dans les cohortes d'entraînement. Les cohortes d'entraînement et de test sont également différentes sur le statut ganglionnaire, avec 76% des patientes ne présentant aucune métastase ganglionnaire (pNo) dans la cohorte d'entraînement totale, contre 88% dans la cohorte test.

2.2 - Méthodes

Nous avons tout d'abord évalué la prédiction du sous-type intrinsèque uniquement sur la base des CCPT, par des approches de ML, pour comparer les performances avec les algorithmes proposés par St Gallen et Maisonneuve. Ensuite, nous avons évalué la prédiction par analyse d'image d'une approche en DL seule. Puis, nous avons évalué les performances de prédiction d'algorithmes de ML se basant sur des features extraites au préalable. Enfin, nous avons évalué les performances des combinaisons de ces différentes prédictions.

2.2.1 - Préparation des whole slide images

Les analyses d'images par IA s'appuyaient sur des lames HE entières et numérisées de tumeur du sein (WSI) comme présenté en figure 1.

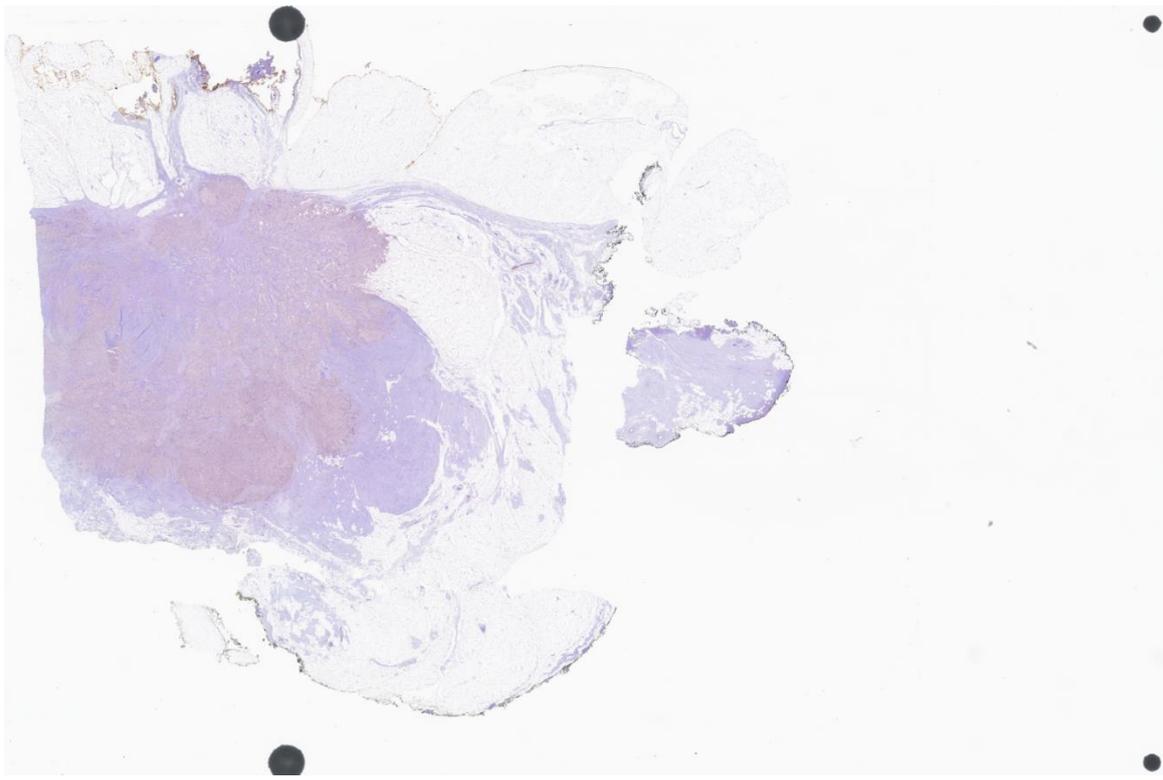


Figure 1 : exemple de WSI utilisée pour nos modèles

Pour la cohorte d'entraînement, après le désarchivage des blocs fixés en paraffine des 66 tumeurs du sein, le bloc le plus représentatif de la tumeur était sélectionné (généralement, celui ayant servi à l'analyse moléculaire). Dans le processus de l'étude

APRIORICS, une lame HE était refaite à partir du bloc fixé en paraffine. Après cette étape, la lame était numérisée sur scanner de lames Aperio GT 450DX au grossissement x40, soit une résolution de 0,263 $\mu\text{m}/\text{pixel}$. Les WSI ainsi générées étaient stockées au format sv5 sur des serveurs dédiés et étaient visualisables avec différents viewers de lames virtuelles (Aperio ImageScope, QuPath, Cytomine). Des vérifications manuelles de la qualité technique étaient opérées par le technicien chargé du projet, à chaque étape du processus.

Pour la cohorte de test, les tumeurs sélectionnées ne faisaient pas partie de l'étude APRIORICS. Les lames HE représentatives des tumeurs étaient désarchivées, nettoyées et scannées avec le même scanner et les mêmes paramètres que pour la première cohorte. Il n'y avait pas de recréation de lame HE à partir du bloc fixé en paraffine. Les WSI ainsi obtenues étaient transférées sur le serveur. Elles suivaient ultérieurement les mêmes traitements d'image que pour la première cohorte, avant l'application des algorithmes d'IA.

2.2.2 - Prédiction à partir des IHC et d'autres CCPT

La première étape de notre étude consistait à développer un algorithme capable de prédire les sous-types intrinsèques luminal A et B à partir des IHC et d'autres CCPT, à la manière des algorithmes décisionnels proposés dans la conférence de consensus de St Gallen et dans l'étude de Maisonneuve vues précédemment. Cette étape avait pour objectifs d'évaluer la concordance des résultats avec ceux de la littérature et d'étudier la pertinence potentielle d'autres CCPT non utilisés actuellement. Pour développer ces algorithmes, nous avons utilisé le modèle XG boost. Pour cette approche, nous n'avons pas effectué de pré-traitement des données.

Dans un premier temps, nous avons entraîné un modèle sur les valeurs des pourcentages de positivité des RE, RP et Ki67, ainsi que sur l'intensité du marquage des RE et RP.

Dans un second temps, nous avons augmenté le nombre de données d'entrée avec la valeur des différentes composantes du grade de Nottingham, la taille tumorale, le nombre de ganglion positif, le nombre de ganglion en rupture capsulaire, le statut ganglionnaire pN et la valeur de l'index mitotique.

2.2.3 - Prédiction par analyse d'image en DL

La deuxième étape de notre étude consistait à développer un classifieur DL capable de prédire les 2 classes lumineales A et B sur les images histologiques des tumeurs. Pour procéder à des analyses d'image sur ces WSI, il est nécessaire de fragmenter l'image en plus petits patches (figure 2). Nous avons choisi de travailler sur des patches au grossissement de niveau 1 et de dimension 512 x 512 pixels. Ce choix nous semblait être un bon compromis entre le fait de garder du contexte d'analyse et le nombre d'images à traiter par l'algorithme. Ces patches étaient ensuite triés pour ne garder que ceux contenant un minimum de tissu. Ce dataset d'image RGB (Red Green Blue) était ensuite normalisé pour obtenir des valeurs comprises entre 0 et 1.

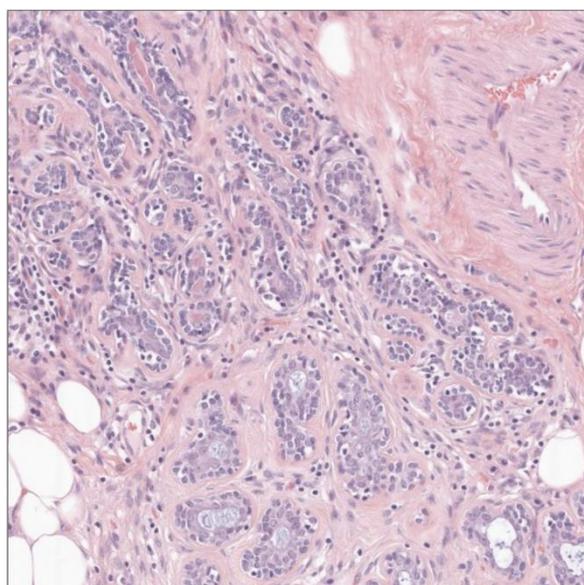


Figure 2 : exemple de patch 512 x 512 pixels

Sur une WSI de tumeurs du sein, certains patches ne contiennent pas de tumeur et ne sont donc pas informatifs pour la prédiction du sous-type moléculaire. Nous avons donc annoté les zones tumorales dans chaque WSI (figure 3) pour ne sélectionner que des patches contenant l'information. Les patches de ces zones étaient ensuite labellisés avec les 2 classes de prédiction, luminal A ou luminal B en fonction du résultat de l'analyse moléculaire.

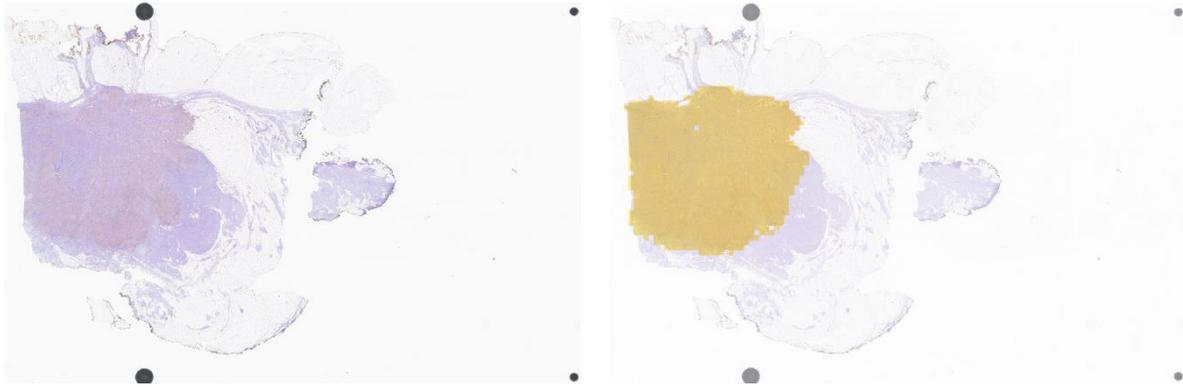


Figure 3 : exemple d'annotation de la zone tumorale (en jaune)

Après quelques essais et le constat que le modèle accordait parfois de l'importance à des secteurs ne présentant pas de tumeur (car certains patchs des zones tumorales annotées ne contiennent pas de cellule tumorale et ne sont donc pas informatifs), nous avons décidé d'introduire une 3ème classe de prédiction nommée "autre", qui correspond au tissu non tumoral en dehors des zones tumorales annotées, et de relancer les apprentissages sur l'ensemble des patchs, tumoraux ou non, avec ces 3 classes de prédiction. Cette étape avait pour but d'identifier des patchs ne contenant pas de cellules tumorales à l'intérieur des zones tumorales annotées. Ainsi, pour la distinction luminal A et luminal B, le réseau avait moins de risque de se focaliser sur des zones peu informatives.

Nous avons par la suite procédé à une augmentation des données (figure 4) en appliquant des modifications aux patchs (changement de couleurs, de contrastes, de luminosité, retournement, floutage, cut out, inversion et variation dans les canaux de couleur).

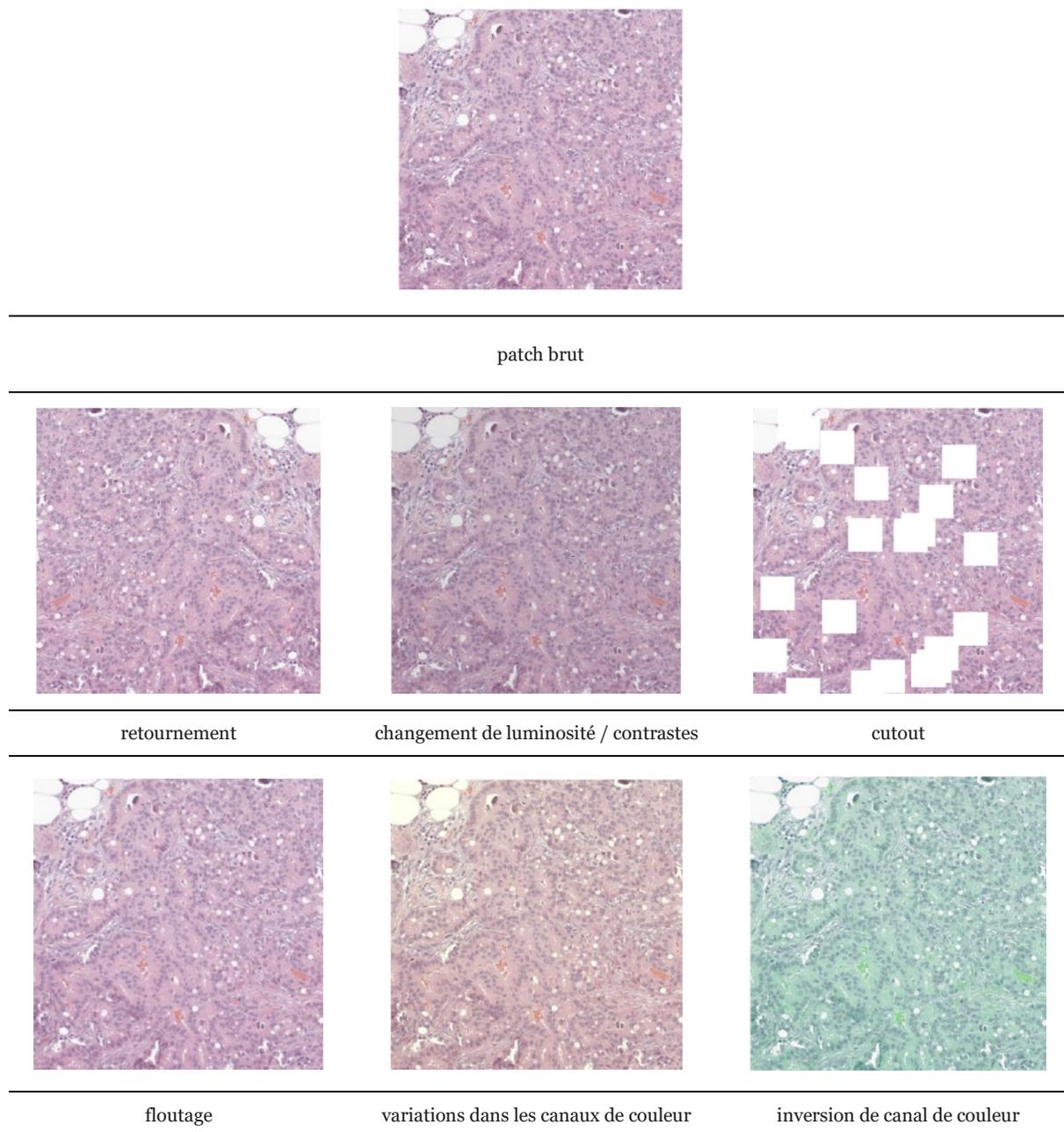


Figure 4 : exemples d'augmentation de données

Après essais de différents modèles de la littérature, nous avons choisi d'utiliser le réseau ResNet 50 que nous avons par la suite affiné, auquel nous avons rajouté une couche entièrement connectée. Les calculs des algorithmes étaient effectués sur un GPU Tesla p100. Les hyper paramètres utilisés pour l'entraînement du modèle sont présentés ci-dessous (tableau 5).

Hyperparamètres	Resnet50 fine-tuned
classifieur drop out rate	0,5
loss	cross-entropy
epochs	30
batch size	32
learning rate	0,01
weight decay	0,0001
scheduler	ReduceLROnPlateau

Tableau 5 : hyperparamètres utilisés lors de l'entraînement de nos modèles

Le jeu de données était équilibré en terme de nombre de WSI pour chaque classe mais déséquilibré en terme de nombre de patch pour chaque classe. Pour contrebalancer ce déséquilibre lors de l'apprentissage et la sélection aléatoire des patches, des poids différents étaient attribués aux patches en fonction de sa classe (la classe "autre", majoritaire, avait un poids plus faible que les 2 autres).

Le jeu de données était divisé en un set d'entraînement et un set de validation avec des proportions de 80 et 20 %. Cette partition du jeu de données était répétée 5 fois pour obtenir 5 modèles de prédiction différents permettant d'augmenter la fiabilité du modèle généré (méthode de cross-validation en 5 plis ou 5 fold cross-validation).

Les modèles fournissent des prédictions à l'échelle du patch. Pour obtenir une prédiction à l'échelle de la lame entière, nous conservons uniquement les patches où les prédictions luminal A ou B sont prédominantes et nous excluons les patches prédit de classe "autre" (figure 5), étant donnée l'excellente accuracy observée pour cette tâche. Ensuite nous réalisons une agrégation des données de ces patches en testant 2 méthodes : une moyenne des prédictions et un vote à la majorité.

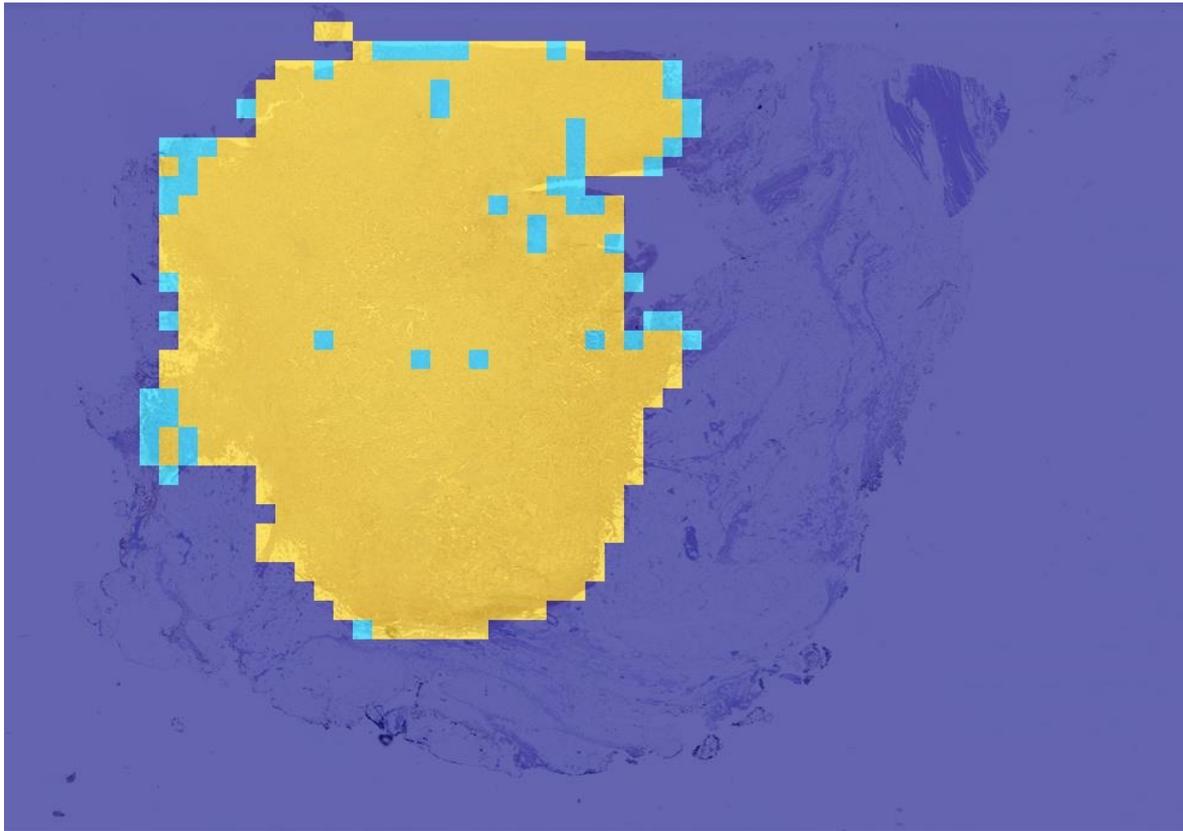


Figure 5 : exemple de représentation des prédictions de l'ensemble des patches de la lames (bleu = luminal A ; jaune = luminal B ; violet = autre)

2.2.4 - Prédiction par analyse d'image en ML à partir des features nucléaires

La troisième étape de notre étude consistait à développer un classifieur ML capable de prédire les 2 classes lumineales A et B en se basant sur des features nucléaires extraits préalablement. Pour se faire, nous avons commencé par segmenter les noyaux (figure 6) dans les zones tumorales annotées, en utilisant le réseau de neurones de segmentation d'instance HoVer-Net (150).

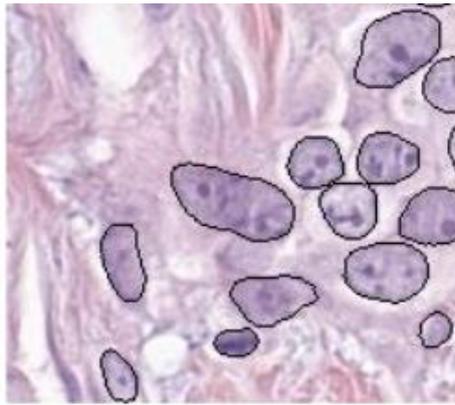


Figure 6 : exemple de segmentation par HoVer-Net

Cette segmentation n'étant pas parfaite (figure 7), il était nécessaire d'ajuster les résultats en éliminant des objets segmentés qui ne sont pas des noyaux. Un algorithme était entraîné pour reconnaître les mauvaises segmentations. Un critère de taille était également mis en place pour ne garder que les seuls objets mesurant plus de 100 pixels².

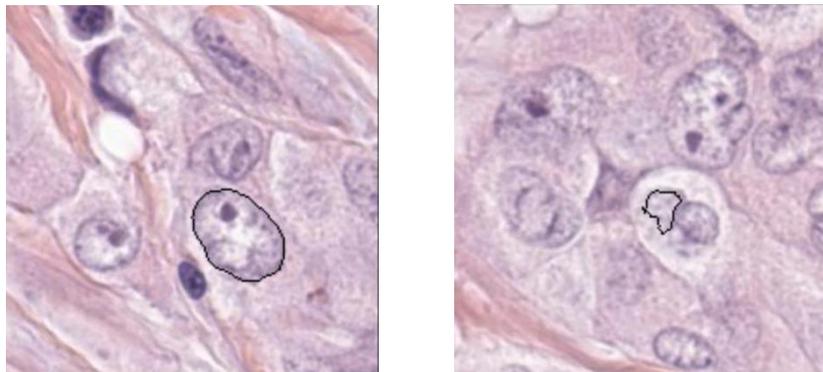


Figure 7 : exemple de bonne segmentation (à gauche) et de mauvaise segmentation (à droite)

A partir de cette segmentation, il était possible d'extraire différentes catégories de features nucléaires liées à la géométrie, la texture ou la coloration des noyaux. Cette segmentation nous permettait également de classer les noyaux en 2 catégories en fonction de leur taille et de leur coloration : noyau de lymphocyte et noyau de cellule tumorale (figure 8). Ce qui nous permettait de quantifier par la suite, des features de distribution et d'interaction de ces 2 populations.

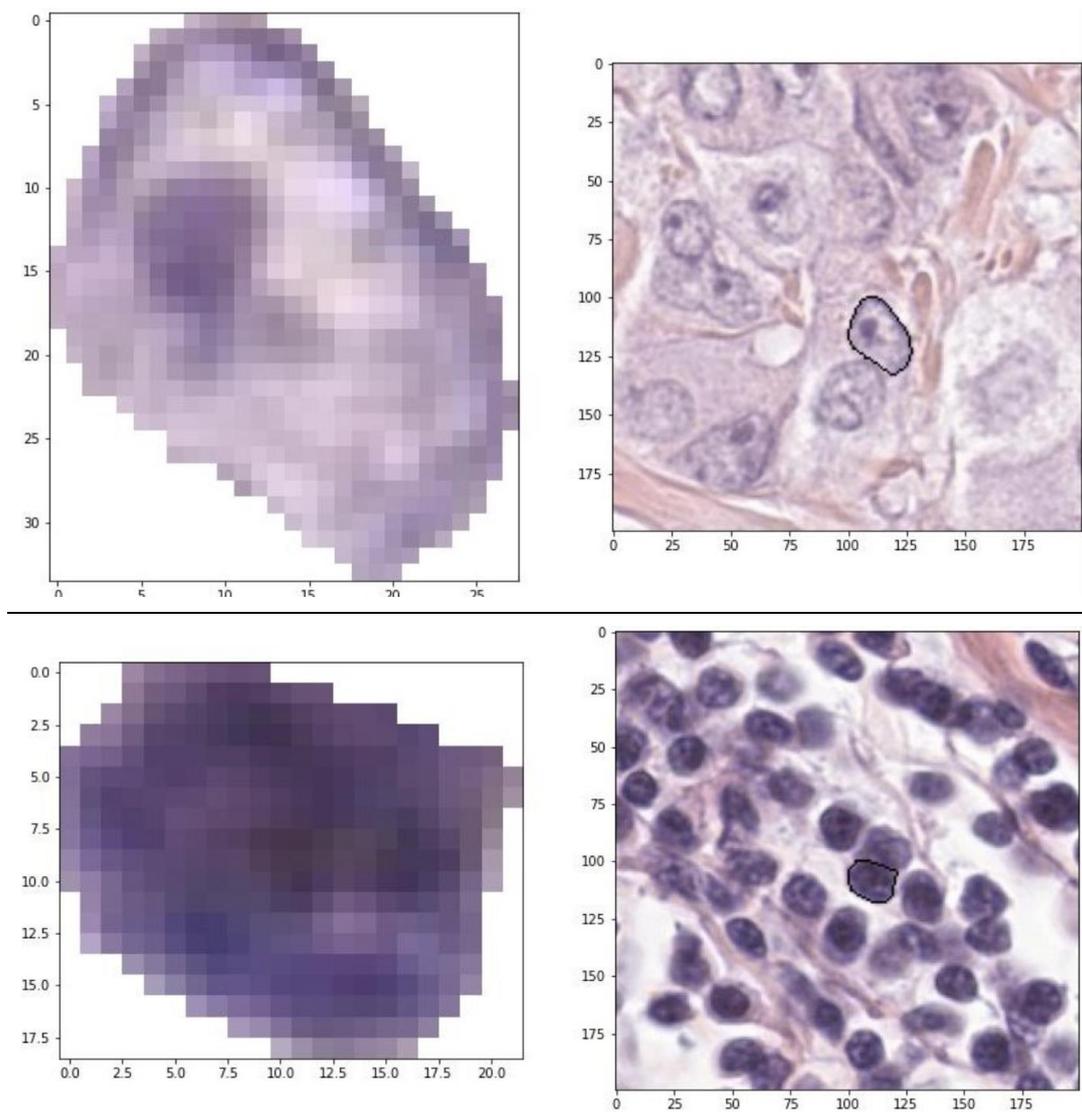


Figure 8 : distinction des cellules tumorales (haut) et des lymphocytes (bas) sur les features

Au total, 103 features étaient extraites. L'ensemble des features est présenté et détaillé dans le tableau 6. Les valeurs de ces features étaient extraites pour chaque noyau d'une tumeur, puis des moyennes (mean) ou des déviations standards (std) étaient calculées à l'échelle de la tumeur. Les valeurs étaient collectées dans un tableau.

Features de géométrie	area_mean	area_std
	gd_ax_mean	gd_ax_std
	pt_ax_mean	pt_ax_std
	length_mean	length_std
	compactness_mean	compactness_std
	elliptic_fit_mean	elliptic_fit_std
	enclosing_circle_overlap_mean	enclosing_circle_overlap_std
	length_width_ratio_mean	length_width_ratio_std
	Eccentricity_mean	Eccentricity_std
	smoothness_mean	smoothness_std
	convex_area_mean	convex_area_std
iou_mean	iou_std	
Features de coloration	normalized_red_intensity_mean	normalized_red_intensity_std
	normalized_green_intensity_mean	normalized_green_intensity_std
	normalized_blue_intensity_mean	normalized_blue_intensity_std
	h_mean_mean	h_mean_std
	e_mean_mean	e_mean_std
	h_std_mean	h_std_std
	e_std_mean	e_std_std
	hue_mean_mean	hue_mean_std
	hue_std_mean	hue_std_std
	saturation_mean_mean	saturation_mean_std
	value_hsv_mean_mean	value_hsv_mean_std
	saturation_std_mean	saturation_std_std
	value_hsv_std_mean	value_hsv_std_std
	hu_0_mean	hu_0_std
	hu_1_mean	hu_1_std
	hu_2_mean	hu_2_std
	hu_3_mean	hu_3_std
hu_4_mean	hu_4_std	
hu_5_mean	hu_5_std	
hu_6_mean	hu_6_std	
Features de texture	Angular Second Moment_mean	Angular Second Moment_std
	Contrast_mean	Contrast_std
	Correlation_mean	Correlation_std
	Sum of Squares: Variance_mean	Sum of Squares: Variance_std
	Inverse Difference Moment_mean	Inverse Difference Moment_std
	Sum Average_mean	Sum Average_std
	Sum Variance_mean	Sum Variance_std
	Sum Entropy_mean	Sum Entropy_std
	Entropy_mean	Entropy_std
	Difference Variance_mean	Difference Variance_std
	Difference Entropy_mean	Difference Entropy_std
	Information Measure of Correlation 1_mean	Information Measure of Correlation 1_std
Information Measure of Correlation 2_mean	Information Measure of Correlation 2_std	
Features de répartition cellulaire	cellularite_tumoral	
	cellularite_tumoral_big	
	cellularite_tumoral_mid	
	cellularite_tumoral_small	

part_tum_petit_0
part_tum_petit_1
part_tum_mid_0
part_tum_mid_1
part_tum_big_0
part_lymph_0
part_lymph_1
cellularite_lympho
lymph_v_tum

Tableau 6 : liste des 103 features extraites

Une première analyse de variance par ANOVA f-test était effectuée pour sélectionner parmi ces features, les plus pertinentes. Cette même analyse était réitérée à chaque fois, pour les prédictions combinées avec les IHC, les CCPT et les features nucléaires (l'ensemble des données étaient trié à nouveau par ANOVA f-test).

Pour la prédiction du sous-type intrinsèque ensuite, différents modèles ont été testés (XG boost, Random Forest, régression logistique et support vector machine ou SVM). Les meilleures performances étaient obtenues avec le Random Forest et le XG boost, qui donnaient des résultats proches. Nous avons donc gardé les performances du Random Forest pour les résultats.

2.2.5 - Combinaisons des différentes méthodes

Enfin, nous avons combiné les résultats de ces différentes approches pour évaluer leur apport dans la prédiction du sous-type intrinsèque.

Pour la prédiction en DL, nous avons utilisé le même logiciel XGboost que pour la prédiction à partir des CCPT, en ajoutant aux données la valeur de la prédiction faite par le DL. A noter que dans la cohorte d'entraînement totale, 14 patientes n'avaient pas pu bénéficier de la prédiction par DL, et pour ces cas, la valeur était traitée par l'algorithme comme une valeur manquante.

Pour la prédiction à partir des features en combinaison avec les IHC et les CCPT, nous avons gardé une approche par Random Forest.

2.2.6 - Evaluation des performances et de l'interprétabilité des modèles

2.2.6.1 - Evaluation des performances

Pour chacune de ces méthodes, les performances de prédiction étaient évaluées par l'accuracy et l'AUC (Area Under the ROC Curve). Dans une matrice de confusion comme présentée dans la figure 9, VP désigne les vrais positifs, VN les faux négatifs, FP les faux positifs et FN les faux négatifs.

		Réalité	
		luminal A	luminal B
Prédiction	luminal A	VP	FP
	luminal B	FN	VN

Figure 9 : exemple de matrice de confusion

L'accuracy est le taux de bonnes réponses données par l'algorithme. Elle est calculée comme suit : $Accuracy = (VP + VN) / (VP + VN + FP + FN)$

La courbe ROC est une courbe représentant le taux de vrais positifs en fonction du taux de faux positifs pour différents seuils de positivité. L'AUC reflète donc la probabilité que le modèle classe un exemple positif aléatoire au-dessus d'un exemple négatif aléatoire. La valeur de l'AUC est comprise entre 0 et 1, un modèle ayant 100% de prédiction juste aura une AUC de 1.

2.2.6.2 - Evaluation de l'interprétabilité

Certains outils permettent de faire des représentations visuelles des résultats sous forme de graphiques ou d'images. Pour les algorithmes de DL, nous avons utilisé la méthode Grad-CAM (Gradient-weighted Class Activation Mapping) qui permet, à l'échelle d'un patch, de créer une carte thermique en utilisant les gradients de la couche convolutive finale, pour visualiser les régions d'intérêt pour la classification de l'algorithme. A l'échelle de la lame entière pour les algorithmes de DL, nous réalisons des Heatmaps, en attribuant à chaque patch une couleur allant du bleu au rouge, qui traduit son score de prédiction en fonction du sous-type prédit (pour une tumeur prédite luminale B, les patchs classés luminal B seront rouges et les patchs classés

luminal A seront bleus), permettant ainsi d'étudier, à l'échelle de la lame, l'hétérogénéité de la classification en fonction des zones. Ces 2 méthodes sont présentées dans la figure 10.

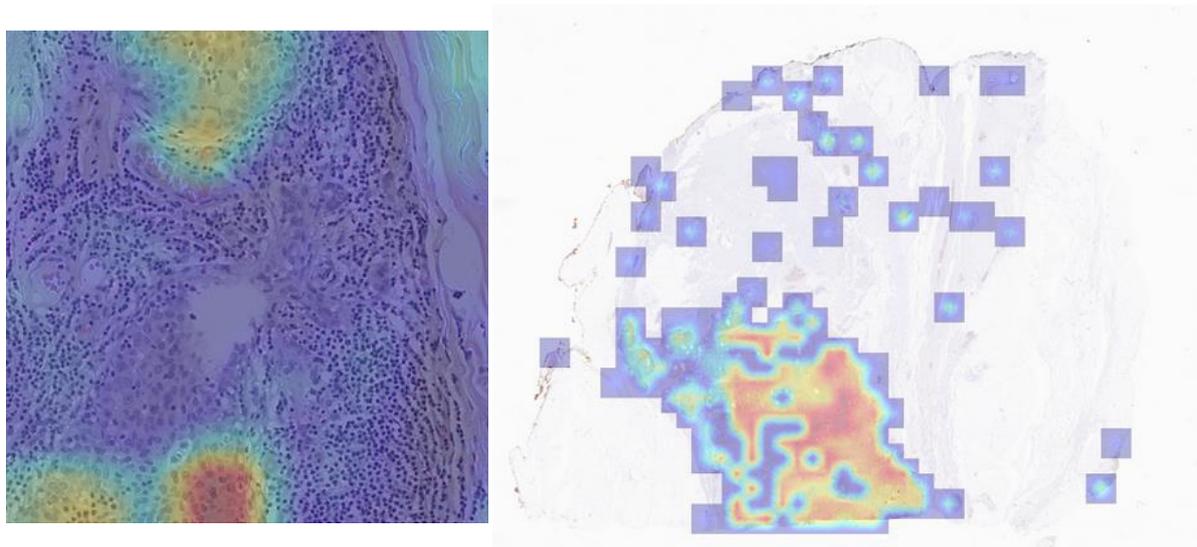


Figure 10 : exemple de Grad-CAM pour un patch (à gauche) et de Heatmap à l'échelle d'une WSI (à droite)

Pour la prédiction à partir des CCPT, des features nucléaires ou des combinaisons des différentes méthodes, nous réalisons des SHAP value qui sont des représentations des différentes données d'entrée, classées selon leur importance pour la tâche à réaliser par l'algorithme (figure 11). Chaque point représente une valeur de la donnée d'entrée, allant du bleu (faible) au rouge (haute). Elles possèdent une valeur d'impact sur le modèle (en abscisse) négative, nulle ou positive et sont réparties de part et d'autre de l'axe zéro en fonction de cette valeur. Dans les SHAP value qui seront présentées dans nos résultats, nous avons conçu les algorithmes de telle sorte que les points ayant des valeurs négatives (à gauche de l'axe) sont des valeurs qui font prédire le sous-type luminal A, et les points ayant des valeurs positives (à droite) sont des valeurs qui tendent à faire prédire le sous-type luminal B.

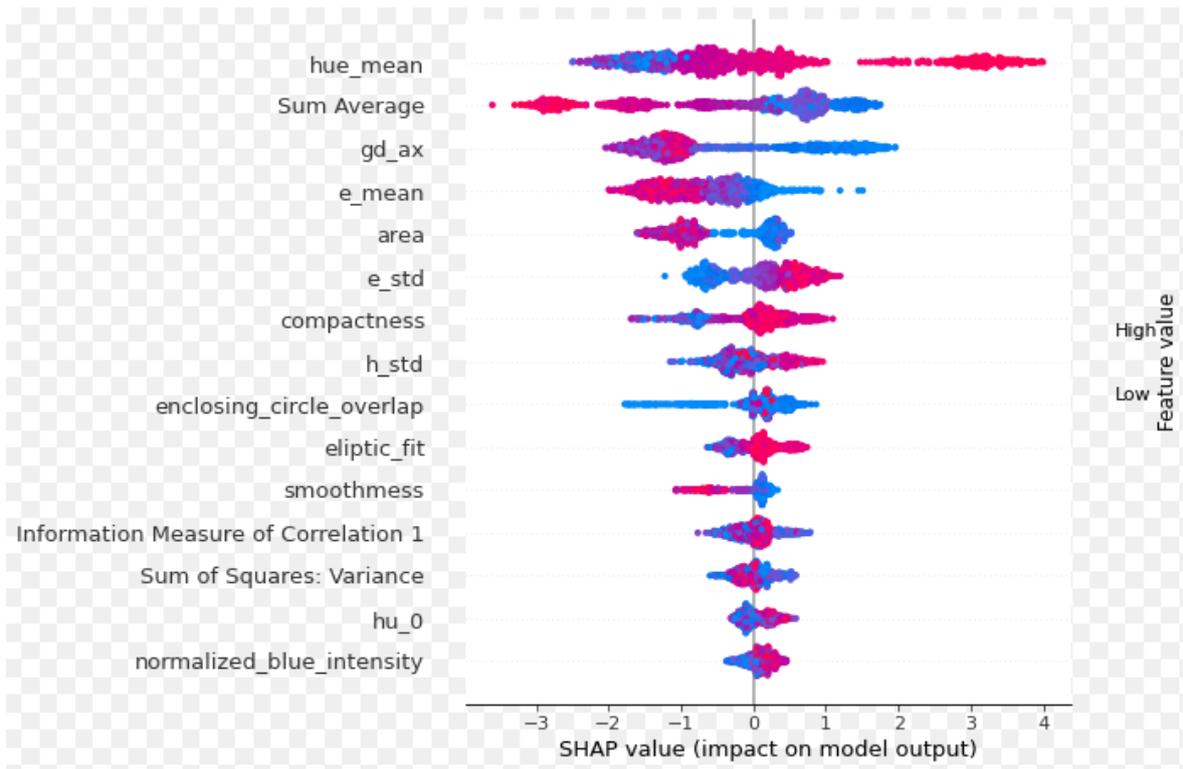


Figure 11 : exemple de SHAP Value, avec les données d'entrée triées du haut vers le bas, par ordre d'importance pour l'algorithme, les plus importantes étant en haut

3 - Résultats et discussion

3.1 - Performances de la prédiction à partir des IHC et d'autres CCPT dans un dataset monocentrique (Oncopole)

3.1.1 - Selon St Gallen et Maisonneuve

	cohorte d'entraînement totale	cohorte de test
Accuracy avec St Gallen	0,625	0,692
Accuracy avec Maisonneuve	0,660	0,731

Tableau 7 : performances de prédiction du sous-type intrinsèque sur les IHC proposés par St Gallen et Maisonneuve

On observe ici des performances variant de 62,5 à 73,1 % d'accuracy pour la prédiction du sous-type intrinsèque avec les seuils recommandés dans la conférence de St Gallen et retravaillés par Maisonneuve. Ceux de Maisonneuve semblent conférer une amélioration des performances. Ces résultats sont concordants avec ceux obtenus dans la littérature (152).

3.1.2 - Prédiction et interprétabilité par ML avec XG boost

	cohorte d'entraînement totale		cohorte de test	
	average accuracy	auc	accuracy	auc
RE/RP/Ki67	0,80	0,80	0,73	0,73
RE/RP/Ki67 + autres CCPT	0,85	0,847	0,846	0,84

Tableau 8 : performances de prédiction du sous-type intrinsèque par ML (XGboost) sur les IHC et les CCPT

Avec les seules valeurs de Ki67, RE et RP, on constate que la prédiction est sensiblement identique à celle observée avec l'algorithme de Maisonneuve et al. En revanche, l'ajout d'autres CCPT semble améliorer de 10% les performances, avec des résultats plus stables entre l'entraînement et le test.

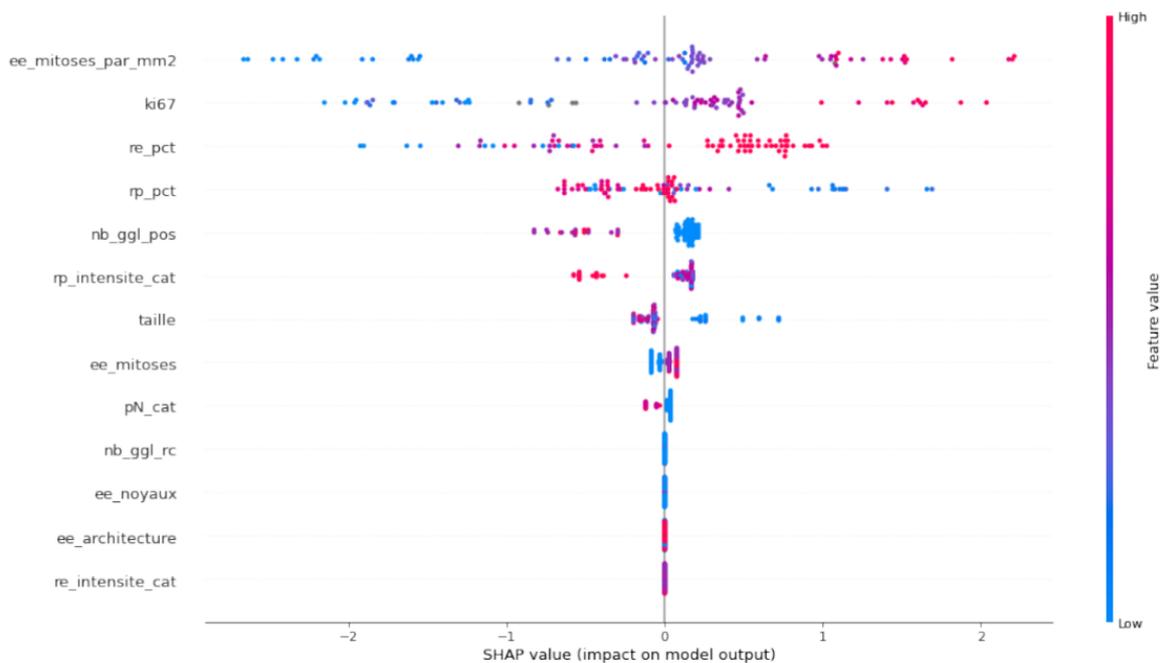


FIGURE 15 – SHAP value summary plot ki re/rp ccpu.

Figure 12 : SHAP value de la prédiction du sous-type intrinsèque par ML (XGboost) sur la combinaison des IHC et des CCPT

Concernant l'interprétabilité des résultats, la figure 12 montre que la donnée la plus importante pour l'algorithme est l'index mitotique exprimé en nombre de mitoses par mm², suivie par le Ki67. L'importance de la prolifération dans la distinction des sous-types intrinsèques va dans le sens de ce qui est attendu compte tenu de la grande part des gènes impliqués dans la prolifération dans les gènes analysés en Prosigna® PAM50. Ces 2 données présentent des répartitions de valeurs relativement superposables, ce qui pourrait traduire une corrélation entre ces 2 variables. Néanmoins, l'étude de corrélation des variables par les graphiques présentés en annexe 2, montre au contraire que ces 2 variables ne sont pas corrélées dans notre cohorte d'entraînement. On remarque également que le grade mitotique est moins utile à la prédiction que l'index mitotique, ce qui signifie qu'il existe une perte d'information dans le passage de l'index mitotique au grade.

Les pourcentages de positivité aux RE et RP arrivent ensuite, avec des valeurs faibles/modérées des RE plutôt en faveur d'un sous-type luminal A, et des valeurs fortes plutôt en faveur d'un luminal B. Ce résultat paraît contre-intuitif de par le fait que, biologiquement, les tumeurs lumineales A expriment fortement les gènes des

récepteurs, à l'inverse des luminales B. La distribution des valeurs de RP quant à elle suit la même logique que l'algorithme de Maisonneuve, avec des valeurs élevées plutôt en faveur d'un sous-type luminal A.

La donnée suivante, le nombre de ganglions positifs, interpelle car il semblerait qu'une valeur élevée soit en faveur d'un sous-type luminal A. La même tendance est observée de façon moins marquée pour le stade ganglionnaire pN, étant donné que ces 2 variables sont corrélées. Cette observation semble contre-intuitive pour le sous-type étant le moins agressif. Ce résultat nous semblerait plutôt lié à un biais de sélection des patientes. En effet, la cohorte utilisée correspondant à des situations de "vraie vie", les situations de discordance (agressivité tumorale vs envahissement ganglionnaire, index mitotique vs index de prolifération Ki67...) étaient fréquentes car elles justifiaient la prescription d'une signature moléculaire. L'analyse de la répartition des valeurs de cette variable pour chaque sous-type dans la cohorte d'entraînement confirme en effet un déséquilibre, avec la majorité des cas présentant au moins 1 ganglion atteint étant de sous-type luminal A.

Un autre résultat surprenant est l'intensité de marquage du RP, qui semble être utile pour la prédiction du sous-type luminal A lorsque celui-ci est élevé. Intuitivement, cette donnée étant peu informative en pratique courante, ce résultat nous semble provenir plus d'un biais dans les données des cohortes.

La taille tumorale a également une répartition étonnante étant donné que l'ensemble des valeurs élevées apparaît à gauche dans la catégorie luminal A. Cette répartition des valeurs semble plus provenir d'un biais de sélection que d'une réelle différence intrinsèque aux sous-types, comme le montre l'étude de répartition de ces variables (tableau 9).

	Cohorte d'entraînement totale	
	luminal A	luminal B
	(n = 37)	(n = 43)
Taille tumorale en mm – moyenne [min ; max]	19,4 [10 ; 55]	17,4 [9 ; 56]
Détail grade : mitoses / mm ² – moyenne [min ; max]	2,9 [0,3 ; 9,7]	5,4 [1 ; 15,9]
Nombre de ganglions envahis – moyenne [min ; max]	0,5 [0 ; 4]	0,2 [0 ; 2]
Taux de positivité de RE – moyenne [min ; max]	89,7 [10 ; 100]	97,3 [80 ; 100]
Taux de positivité de RP – moyenne [min ; max]	61 [0 ; 100]	62 [0 ; 100]

Tableau 9 : répartition selon les sous-types intrinsèques des valeurs de la taille tumorale, du nombre de mitoses / mm², du nombre de ganglions envahis, et du pourcentage de positivité aux RE et RP dans la cohorte d'entraînement

Le reste des données (nombre de ganglions en rupture capsulaire, grade nucléaire, grade architectural, et intensité du marquage du RE) ne semble pas influencer la décision du modèle, les valeurs de l'impact sur le modèle étant nulles pour chaque donnée des variables.

3.1.3 - Conclusion préliminaire sur la prédiction par les IHC et les CCPT

La prédiction basée sur les CCPT montre une performance satisfaisante et la combinaison de ces critères aux données des immunohistochimies RE/RP/Ki67 semble améliorer la prédiction. Dans celle-ci, la prolifération cellulaire, reflétée par le Ki67 et l'index mitotique semble jouer un rôle important. Cependant l'étude de corrélation de ces 2 variables montre qu'elle est faible, proche de zéro. Ce qui pourrait signifier que l'évaluation de l'une ou des deux variables est encore imparfaite. La cellularité tumorale pourrait jouer un rôle : le Ki67 évalue un pourcentage de cellules tumorales et n'est donc que peu ou pas impacté par la cellularité tumorale, tandis que l'index mitotique, étant un nombre de mitoses par mm², est influencé par la cellularité tumorale.

Toutefois, ce modèle a également sélectionné des variables peu informatives en pratique courante ou dont la répartition dans la population semble biaisée. La prédiction sans ces variables devrait donc être différente de celle observée. Mais les résultats étant relativement identiques pour la cohorte test ne présentant pas ces biais (tableau 10), il semblerait que la prédiction en soit relativement indépendante.

	Cohorte de test	
	luminal A	luminal B
	(n = 12)	(n = 14)
Taille tumorale en mm – moyenne [min ; max]	18,1 [13 ; 28]	18,3 [11 ; 28]
Détail grade : mitoses / mm ² – moyenne [min ; max]	4,1 [1,7 ; 7]	5,8 [0,6 ; 12]
Nombre de ganglions envahis – moyenne [min ; max]	0,08 [0 ; 1]	0,14 [0 ; 1]
Taux de positivité de RE – moyenne [min ; max]	98,75 [90 ; 100]	98,57 [95 ; 100]
Taux de positivité de RP – moyenne [min ; max]	51,1 [0 ; 90]	66,6 [0 ; 100]

Tableau 10 : répartition selon les sous-types intrinsèques des valeurs de la taille tumorale, du nombre de mitoses / mm², du nombre de ganglions envahis, et du pourcentage de positivité aux RE et RP dans la cohorte de test

3.2 – Performances de la prédiction par analyse d’image en DL

3.2.1 - Performances sur patch

Cohorte d’entraînement			Cohorte de test		
fold	accuracy	auc	fold	accuracy	auc
0	0,71	0,59	0	0,6	0,59
1	0,47	0,48	1	0,71	0,49
2	0,68	0,72	2	0,68	0,72
3	0,60	0,50	3	0,59	0,51
4	0,57	0,58	4	0,57	0,58

Tableau 11 : performances des prédictions du sous-type intrinsèque par DL seul sur patch

Le tableau 11 présente des performances de prédiction assez faibles à l’échelle du patch, avec 0,57 à 0,71 d’accuracy en test. Ces valeurs sont relativement similaires entre chaque fold et entre l’entraînement et la validation, ce qui montre une certaine fiabilité des résultats, sans surapprentissage.

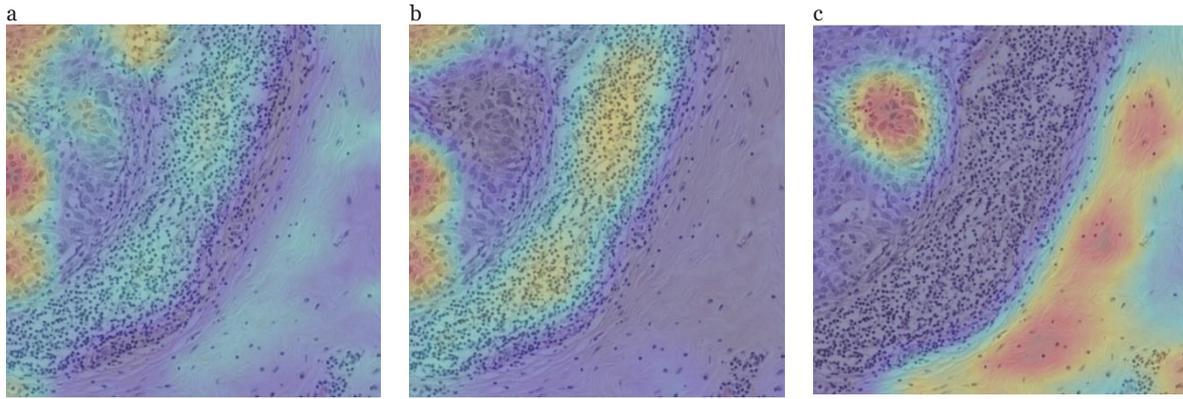


Figure 13 : exemple de Gradcam pour un patch, avec les zones d'intérêt en rouge pour l'algorithme (a = prédiction luminal A ; b = prédiction luminal B ; c = prédiction autre)

En ce qui concerne l'interprétabilité des résultats, l'étude des GradCam des patches était surprenante (figure 13). Il était parfois observé une tendance de l'algorithme à attacher de l'importance à des zones riches en lymphocytes pour le diagnostic de luminal B comme le montre la figure 13. Cependant, dans d'autres patches, l'algorithme attachait de l'importance à des zones qui ne paraissaient pas pertinentes à première vue : des zones ne présentant pas de cellule tumorale pouvaient avoir de l'importance pour affirmer un sous-type luminal A ou B, et des zones riches en cellules tumorales pouvaient apparaître importantes pour affirmer la catégorie "autre".

3.2.2 - Performances sur lame entière

Cohorte d'entraînement			Cohorte de test		
fold	accuracy	auc	fold	accuracy	auc
0	0,76	0,55	0	0,62	0,64
1	0,46	0,46	1	0,54	0,66
2	0,60	0,60	2	0,42	0,49
3	0,48	0,48	3	0,54	0,73
4	0,76	0,76	4	0,65	0,71
average				0,69	0,73

Tableau 12 : performances des prédictions du sous-type intrinsèque par DL seul sur lame entière en effectuant une moyenne des prédictions

Cohorte d'entraînement			Cohorte de test		
fold	acc	auc	fold	acc	auc
0	0,69	0,68	0	0,5	0,55
1	0,57	0,33	1	0,54	0,63
2	0,77	0,66	2	0,46	0,47
3	0,69	0,48	3	0,54	0,68
4	0,77	0,79	4	0,62	0,66
average				0,65	0,64

Tableau 13 : performances des prédictions du sous-type intrinsèque par DL seul sur lame entière en effectuant un vote à la majorité

Sur ces tableaux 12 et 13, on observe tout d'abord que la performance est globalement meilleure en lame entière que sur patch. Ceci semble cohérent avec la méthode diagnostique effectuée par un pathologiste qui peut rarement effectuer un diagnostic sur un patch d'une lame entière, du fait d'un manque de contexte pour cette image. Il semblerait également que, malgré les résultats assez proches, on obtient une meilleure prédiction pour la méthode de moyenne des patches que pour la méthode de vote à la majorité. Ce qui pourrait être le reflet d'une faible étendue des scores de prédiction par patch, proches de 0,5 (les valeurs varient entre 0 et 1, un score en faveur d'un luminal B est proche de 1 et un score en faveur d'un luminal A proche de 0). Dans cette méthode de moyenne, l'accuracy est sensiblement au même niveau que la performance à partir des IHC, et semble moins bonne que celle combinant les IHC aux CCPT.

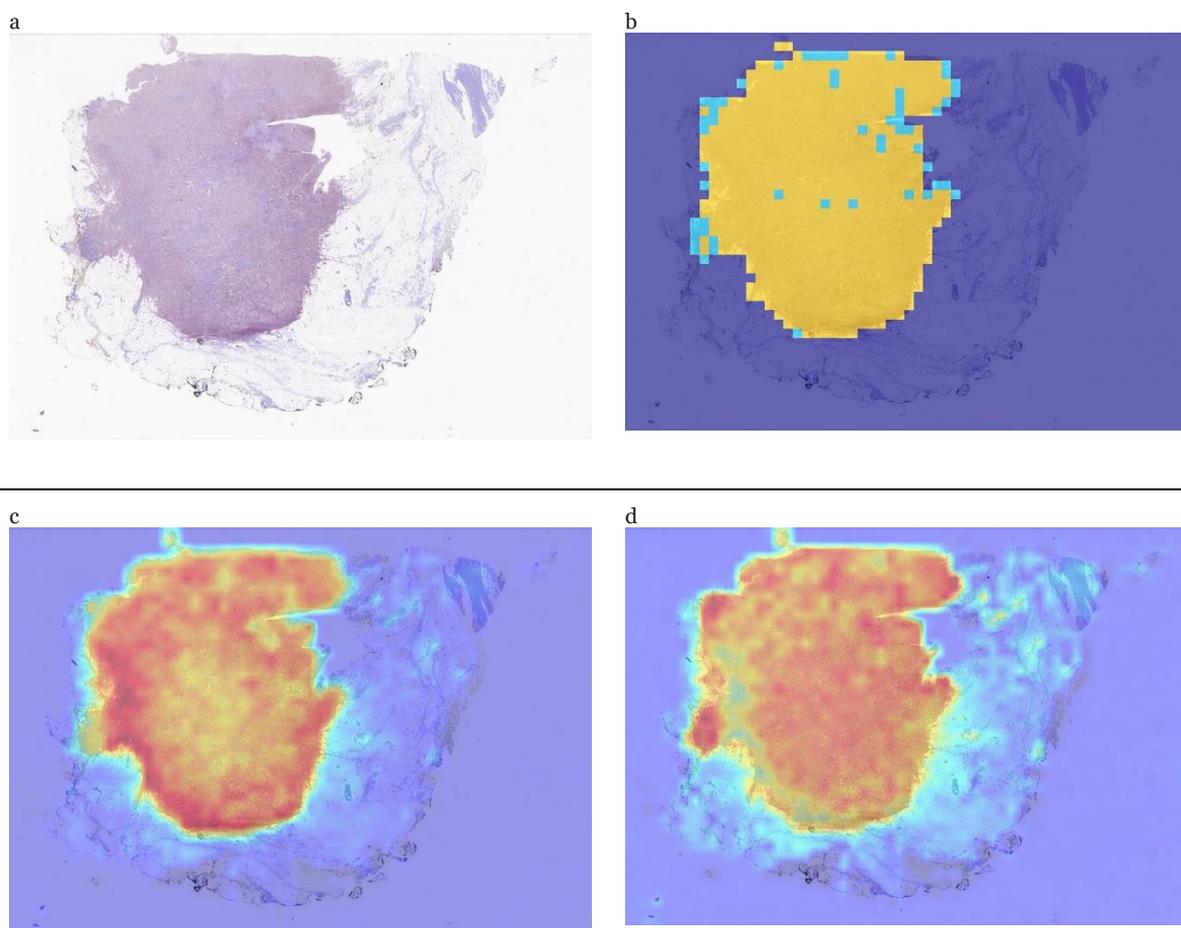


Figure 14 : exemple de carte de prédiction et de heatmaps à l'échelle d'une lame entière. Ici, il s'agit d'une tumeur classée
 luminaire B par le modèle.
 (a = WSI brute ; b = cartographie des prédictions par patch avec les patches prédits luminal A en bleu, les patches prédits luminal B en jaune et les patches prédits autre en violet ; c = heatmaps pour la prédiction luminal B ; d = heatmaps pour la prédiction luminal A)

Concernant l'interprétabilité (figure 14), à l'échelle des lames les tumeurs montraient un profil mixte, faites de patches classés luminal A et de patches classés luminal B. Les heatmaps montraient ensuite les zones classées fortement luminal A/B ou autre. La recherche de caractéristiques morphologiques dans ces zones A ou B ne permettait pas de mettre en évidence de différence morphologique franche sur les lames HE.

3.2.3 - Apport du DL à la prédiction par les IHC et les CCPT

	cohorte d'entraînement totale		cohorte de test	
	average accuracy	auc	accuracy	auc
RE/RP/Ki67	0,80	0,80	0,73	0,71
RE/RP/Ki67 + autres CCPT	0,85	0,847	0,846	0,84
RE/RP/Ki67 + prédiction DL	0,78	0,90	0,77	0,89
RE/RP/Ki67 + autres CCPT + prédiction DL	0,81	0,81	0,73	0,86

Tableau 14 : performances de prédiction du sous-type intrinsèque par ML (XGboost) sur la combinaison des IHC, des CCPT et du DL

Sur ce tableau 14, la prédiction du DL semble apporter une augmentation de l'AUC avec peu ou pas d'amélioration de l'accuracy, quand elle est combinée à la prédiction par les IHC combinées ou non avec les autres CCPT. L'amélioration de l'AUC est de même ordre que celle observée avec l'ajout des autres CCPT à la prédiction par les IHC.

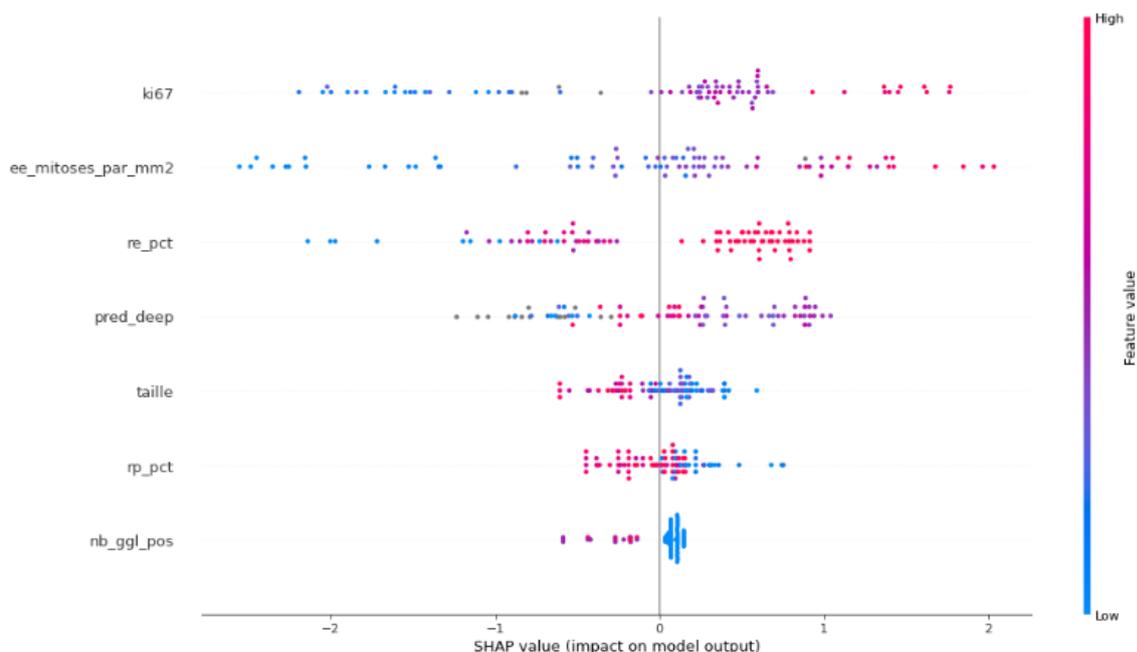


Figure 15 : SHAP value de la prédiction du sous-type intrinsèque par ML (XGboost) sur la combinaison des IHC, des CCPT et du DL

En ce qui concerne l'interprétabilité des résultats, sur la figure 15, on observe que, par rapport à la SHAP value de la prédiction basée sur les IHC combinées aux autres CCPT, l'ajout de la valeur de prédiction du DL modifie les variables sélectionnées. Ainsi, le statut ganglionnaire pN, le nombre de ganglions en rupture capsulaire, le grade mitotique, et le grade nucléaire ne font plus partie de la sélection. La valeur de prédiction du DL vient se placer après les indicateurs de la prolifération (le Ki67 et l'index mitotique) et du pourcentage de positivité des RE.

Cependant, on constate que le nuage de valeurs de prédiction du DL qui oriente vers un sous-type luminal A (à gauche de l'axe 0) comporte une part importante de valeurs manquantes (couleur grise) des 14 cas qui n'ont pu bénéficier de l'analyse d'image, ce qui n'est pas satisfaisant. De plus, il n'est pas observé de valeur manquante dans le nuage de point opposé (luminal B) ce qui n'est pas corrélé à la répartition des sous-types dans les cas n'ayant pas bénéficié de l'analyse d'image (8 luminaux A et 6 luminaux B). Par ailleurs, le nuage de points du côté luminal B comporte en majorité des valeurs moyennes (violette), et les valeurs fortes (rouges) sont réparties de chaque côté. Cette répartition peu discriminante des valeurs reflète et explique en partie les performances modérées de prédiction du DL. Il semblerait sur ce graphique que les valeurs basses ou moyennes de la prédiction du DL soient assez fiables pour prédire respectivement un sous-type luminal A ou luminal B, et que des valeurs hautes soient peu contributives.

3.2.4 - Conclusions préliminaires sur la prédiction par DL

La prédiction du sous-type intrinsèque par le DL donne donc une performance légèrement moins bonne que celle observée pour les CCPT. La combinaison au DL montre également de moins bonnes performances que pour les immunohistochimies combinées aux autres CCPT. Néanmoins, il convient de rappeler que les performances du DL croissent habituellement avec le nombre de données. Or dans notre étude, le nombre de cas total pour les analyses en DL est de 92, ce qui est peu. Au vue de ces performances, les 2 sous-types seraient donc difficile à distinguer sur lame HE. Ces 2 sous-types sont certainement proches morphologiquement, avec un continuum lésionnel entre. Dans notre étude, il semblerait qu'une faible valeur du score de prédiction par le DL soit assez fiable pour affirmer un sous-type luminal A, alors que

les valeurs plus élevées semblent moins catégoriques. L'étude des Gradcam et des Heatmaps laisse à penser que l'infiltration lymphocytaire pourrait aider dans la distinction de ces 2 sous-types, mais aucune étude dans la littérature scientifique ne permet d'appuyer cette hypothèse. Par ailleurs, les résultats obtenus dans notre étude sont meilleurs que ceux observés dans l'étude de Jaber et al. (142) dans laquelle ils obtenaient une accuracy d'environ 54,6% pour la distinction des sous-types lumineux A et B. Cette différence est peut-être due à la méthode utilisée pour construire leur modèle, entraîné à prédire 4 sous-types intrinsèques, ce qui ne peut que favoriser un taux d'erreur plus important.

3.3 – Performances de la prédiction par analyse d'image en ML à partir de features extraites au préalable

3.3.1 - Performances sur les features seules

	cohorte d'entraînement	cohorte de test	
	accuracy	accuracy	auc
Toutes les features	0,71	0,77	0,79
Features géométriques	0,68	0,73	0,76
Features de distribution cellulaire	0,65	0,73	0,75
Features de couleur et texture	0,66	0,54	0,65

Tableau 15 : performances des prédictions du sous-type intrinsèque par ML (random forest) sur les features seules

Le tableau 15 montre que la prédiction faite à partir de l'ensemble des features obtient une performance de 0,77 d'accuracy et 0,79 d'AUC ; des résultats légèrement meilleurs que ceux observés pour la prédiction avec le DL seul sur lame entière en effectuant une moyenne (0,69 d'accuracy et 0,73 d'AUC). Concernant les différentes catégories de features analysées, on observe que les features géométriques et de distribution cellulaire donnent des performances sensiblement équivalentes, tandis que les features de couleur et de texture donnent une performance plus basse. Il semblerait donc que ces dernières aient moins d'importance pour la prédiction du sous-type intrinsèque. Ce qui corréle avec la pratique courante, comme le montre l'évaluation du grade nucléaire compris dans le grade de Nottingham, qui se focalise sur la taille et la forme des noyaux plus que sur leur variation de couleur et de texture.

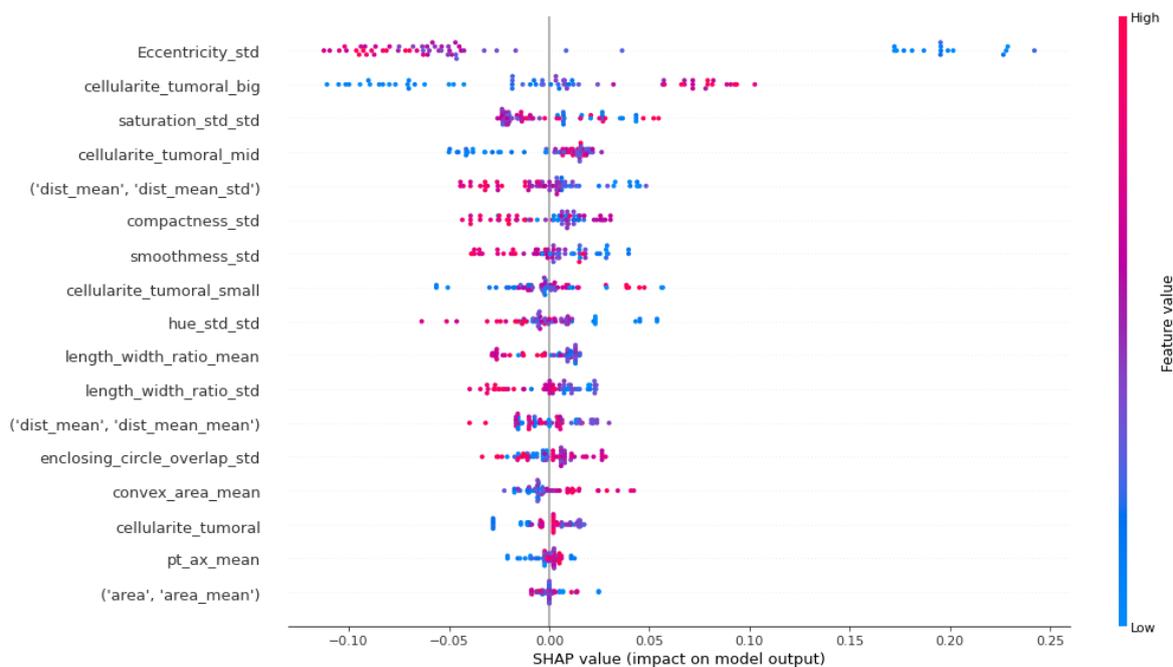


Figure 16 : SHAP value de la prédiction du sous-type intrinsèque par ML (random forest) sur les features seules

En ce qui concerne l'interprétabilité, sur la figure 16 on observe que le modèle de Random Forest a retenu 17 features discriminantes pour la prédiction du sous-type intrinsèque.

Parmi ces features, 6 sont des features de répartition cellulaire reflétant toutes la cellularité tumorale :

- Les features `cellularite_tumoral_big/mid/small` et `cellularite_tumoral` sont mesurées par le nombre de noyau de cellules tumorales totales ou par catégorie de grande/moyenne/petite taille, rapporté à l'aire de la surface tumorale annotée. Selon ce graphique, il semblerait que la plus discriminante soit la cellularité des grandes cellules avant celle des moyennes, des petites ou de l'ensemble des cellules tumorales. Pour toutes ces features, on observe que des cellularités élevées orientent vers un sous-type luminal B.
- Les features `distance_mean_mean/std` sont une mesure de la distance moyenne d'un noyau à ses plus proches voisins, exprimée en moyenne ou déviation standard à l'échelle d'une lame. Elle reflète donc indirectement la cellularité tumorale. Pour ces features, la déviation standard semble plus informative que la moyenne, avec des valeurs hautes pour ces features qui semblent orienter plutôt vers un sous-type luminal A. Ce résultat peut paraître

un peu contre-intuitif mais pourrait s'expliquer par le fait que plus la cellularité augmente, plus les cellules sont jointives, et donc plus la déviation standard est réduite. A l'inverse d'une tumeur peu cellulaire dans laquelle les noyaux sont espacés, parfois séparés par un stroma plus abondant, pour lesquelles la variation de distance inter-nucléaire a de forte chance d'être variable. Or les tumeurs lumineales A seraient moins cellulaires que les B.

9 autres features sont des features morphométriques :

- La `length_width_ratio_mean` est la moyenne des ratios de la plus grande dimension sur la plus petite dimension du noyau. Cette feature reflète donc le caractère elliptique des noyaux (un noyau rond aura des valeurs basses alors qu'un noyau elliptique aura des valeurs hautes). Ici, on remarque que des valeurs hautes de ces features font tendre vers le sous-type luminal A, ce qui signifie que les noyaux des tumeurs lumineales A seraient plus elliptiques et les noyaux des lumineales B plus ronds.
- L'`eccentricity_std` est la déviation standard d'un paramètre caractéristique d'une courbe conique. Un cercle a une excentricité de 0, et une ellipse a une excentricité entre 0 et 1. La `compactness_std` est la déviation standard d'une mesure de la rondeur de l'objet, donnant la valeur minimale de 1 pour un cercle. La `length_width_ratio_std` est la déviation standard du ratio de la plus grande dimension sur la plus petite dimension du noyau. Ces 3 features ne sont plus des reflets du caractère elliptique du noyau, mais plutôt des indicateurs du pléomorphisme nucléaire, étant donné que ce sont les déviations standard de ces features. Or, dans ce graphique, on observe que des valeurs hautes font tendre le modèle vers un sous-type luminal A. Ce résultat paraît contre-intuitif avec l'expérience de terrain, étant donné qu'en pratique courante, un pléomorphisme nucléaire est plutôt associé à une prolifération agressive.
- La `smoothness_std` est une déviation standard de la mesure de la régularité du contour du noyau. Ici on observe que des valeurs hautes sont plutôt associées aux lumineales A. Ce qui reflète encore un pléomorphisme des aspects plus marqué pour ces tumeurs. En revanche, cette feature ne nous renseigne pas sur le degré de régularité des noyaux. Il aurait fallu avoir la feature `smoothness_mean` pour l'évaluer.

- L'area_mean est la moyenne des aires des noyaux d'une tumeur. La convex_area_mean est une moyenne des aires des formes géométriques minimales sans concavité entourant chaque noyau. Ces 2 features sont complémentaires et permettent d'évaluer la taille des noyaux. Et la comparaison de ces 2 features montre une étendue de valeur plus grande pour la convex_area_mean, et une distribution également plus dichotomique avec des valeurs hautes pour les tumeurs lumineuses B, et des valeurs basses pour les tumeurs lumineuses A. Tandis que l'area_mean semble moins informative. Ce résultat laisse à penser que les noyaux des tumeurs lumineuses B sont globalement plus grands et qu'ils présentent plus de concavités que les noyaux des lumineuses A et donc, plus d'irrégularité de contour.
- L'enclosing_circle_overlap_std est une déviation standard de la mesure de la dimension du cercle qui contient tous les pixels du noyau. C'est un reflet de la variation de taille du noyau mais cette feature semble avoir une distribution peu informative sur le graphique. Ceci peut s'expliquer par le fait que cette feature se base sur une approximation de la taille du noyau, qui peut entraîner une perte de précision de l'information. Elle devrait également être corrélée à la feature gd_ax_mean, mais celle-ci n'apparaît pas dans la sélection de feature.
- Le pt_ax_mean est la mesure du petit axe du noyau, avec la moyenne sur une lame. Dans ce tableau, des valeurs basses semblent être en faveur d'une tumeur lumineuse A. Ce qui pourrait plutôt être dû au caractère elliptique qu'à la taille des noyaux étant donné que la feature gd_ax_mean n'apparaît pas ici, et que la feature area_mean apparaît en dernier dans le graphique

Les 2 dernières features sont des features de texture :

- La hue_std_std est la déviation standard de la mesure de la déviation standard de la teinte de chaque pixel au sein d'un noyau.
- La saturation_std_std est la déviation standard de la mesure de la déviation standard de la saturation de chaque pixel au sein d'un noyau. Ces features reflètent l'hétérogénéité de texture des noyaux d'une tumeur. Une valeur haute signifie une forte variation de la texture des noyaux. Ici, des valeurs modérées/hautes pour ces features semblent orienter vers un sous-type luminal A. Ce sous-type présenterait donc un pléomorphisme plus marqué de la texture

des noyaux. Comme les résultats pour les features reflétant le pléomorphisme nucléaire, ce résultat semble également contre-intuitif avec l'expérience de terrain, mais apporte un argument supplémentaire pour le plus grand pléomorphisme nucléaire des tumeurs lumineales A.

3.3.2 - Apport des features sur la prédiction par les IHC et les CCPT

	cohorte d'entraînement		cohorte de test	
	accuracy	auc	accuracy	auc
RE/RP/Ki67	0,80	0,80	0,73	0,71
RE/RP/Ki67 + autres CCPT	0,85	0,847	0,846	0,84
RE/RP/Ki67 + meilleures features	0,78	X	0,77	0,86
RE/RP/Ki67 + autres CCPT + meilleures features	0,80	X	0,77	0,78

Tableau 16 : performances des prédictions du sous-type intrinsèque par ML (random forest) sur la combinaison des features aux IHC et aux CCPT

Sur le tableau 16, on observe que le bénéfice qu'apporte les features à la prédiction est sensiblement de même ordre pour la prédiction par les IHC ou pour la prédiction par les IHC combinées aux autres CCPT. Cette augmentation semble moins importante que celle que l'on observe quand on ajoute les CCPT aux IHC. On note toutefois une meilleure AUC pour la prédiction avec les immunohistochimies combinées aux CCPT.

L'apport des features sur ces prédictions combinées est donc ici sensiblement du même ordre que celui observé pour le DL. Il semble également surprenant qu'en ajoutant des features au modèle, l'accuracy diminue.

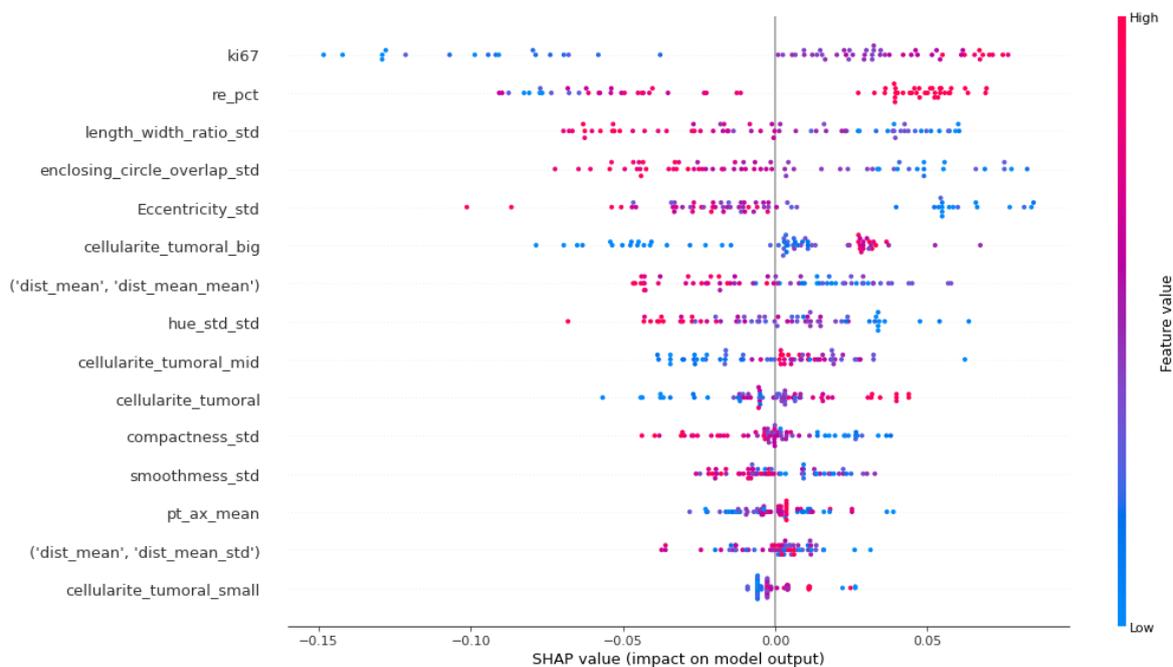


Figure 17 : SHAP value de la prédiction du sous-type intrinsèque par ML (random forest) sur la combinaison des features aux IHC

Concernant l'interprétabilité, sur la figure 17 on observe que les features nucléaires sélectionnées pour cette prédiction, sont des features retrouvées précédemment, ce qui démontre une certaine fiabilité du poids de ces features. Cependant, l'importance de chaque feature et sa place dans la SHAP value varient. Dans les features nucléaires, 3 semblent à nouveau avoir un fort impact sur le modèle : `eccentricity_std`, `cellularite_tumoral_big` et `hue_std_std`. Ce qui vient confirmer l'importance de ces features qui reflètent la cellularité tumorale et le pléomorphisme nucléaire (du moins pour la rondeur et la teinte des noyaux). Certaines features ont au contraire disparu de la sélection : `saturation_std_std`, `length_width_ratio_mean`, `area_mean` et `convex_area_mean`, alors que 2 d'entre elles semblaient avoir un impact important dans la précédente SHAP value. Par ailleurs, Les valeurs du Ki 67 et du RE sont ici les plus discriminantes pour le modèle, devant les features nucléaires. En revanche, la valeur du RP n'apparaît plus dans ce graphique.

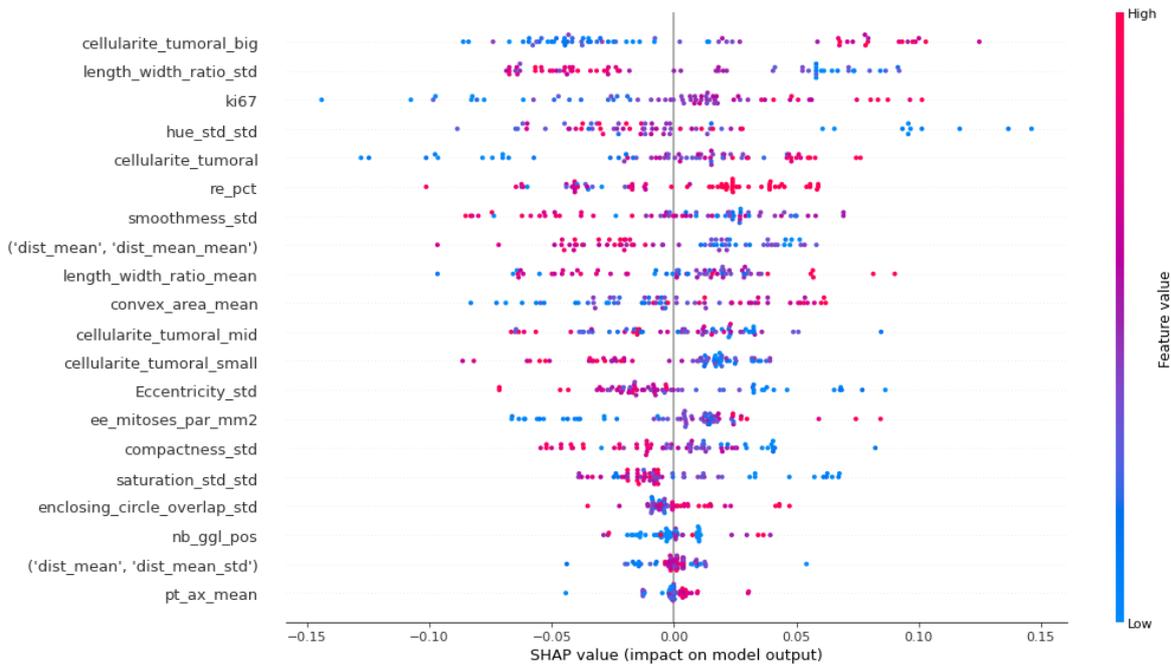


Figure 18 : SHAP value de la prédiction du sous-type intrinsèque par ML (random forest) sur la combinaison des features aux IHC et aux CCPT

Sur la figure 18, on observe à nouveau les mêmes features nucléaires sélectionnées dans la prédiction avec les features seules, ce qui vient appuyer une nouvelle fois l'impact de ces features. Ici encore, l'importance de chaque feature et sa place dans la SHAP value varient. Dans les features nucléaires ayant un fort impact, on retrouve de nouveau la `cellularite_tumoral_big` et la `hue_std_std`. D'autres features reflétant le pléomorphisme comme la `length_width_ratio_std` et la `smoothness_std` semblent également avoir un fort impact dans ce modèle. La feature `area_mean` n'apparaît plus dans la sélection.

Contrairement à la prédiction précédente, les valeurs du Ki67 et du RE semblent ici moins discriminantes pour le modèle, par rapport à certaines features. La valeur du RP n'apparaît toujours pas dans ce graphique. On remarque également que le nombre de mitose par mm² arrive plus tard dans l'ordre d'importance, après le Ki67 et le RE, alors qu'il apparaissait en tête de liste dans la prédiction par les IHC combinées aux CCPT.

A la suite de ces résultats, nous avons cherché à évaluer les performances d'un algorithme pour lequel on aurait effectué un tri dans les features en éliminant celles corrélées entre elles. Pour ce faire, nous avons utilisé les tables de corrélation des différentes variables (annexe 2) pour éliminer manuellement des features nucléaires corrélées. Nous avons éliminé les features `enclosing_circle_overlap_std`,

pt_ax_mean, convex_area_mean, smoothness_std, cellularite_tumoral_mid, length_width_ratio_std, cellularite_tumoral_small, ee_mitoses, pN_cat des variables de l'algorithme. Les performances sont présentées ci-dessous dans le tableau 17.

	cohorte d'entraînement		cohorte de test	
	accuracy	auc	accuracy	auc
RE/RP/Ki67	0,80	0,80	0,73	0,71
RE/RP/Ki67 + autres CCPT	0,85	0,847	0,846	0,84
RE/RP/Ki67 + autres CCPT + meilleures features	0,80	X	0,77	0,78
RE/RP/Ki67 + autres CCPT + meilleures features après élimination des features corrélées	0,83	0,85	0,77	0,86

Tableau 17 : performances des prédictions du sous-type intrinsèque par ML (random forest) sur la combinaison des features aux IHC et aux CCPT, après élimination des features corrélées

On remarque que la performance est légèrement meilleure en termes d'AUC avec le tri des features corrélées. En revanche l'accuracy reste identique à la précédente prédiction sans tri des features.

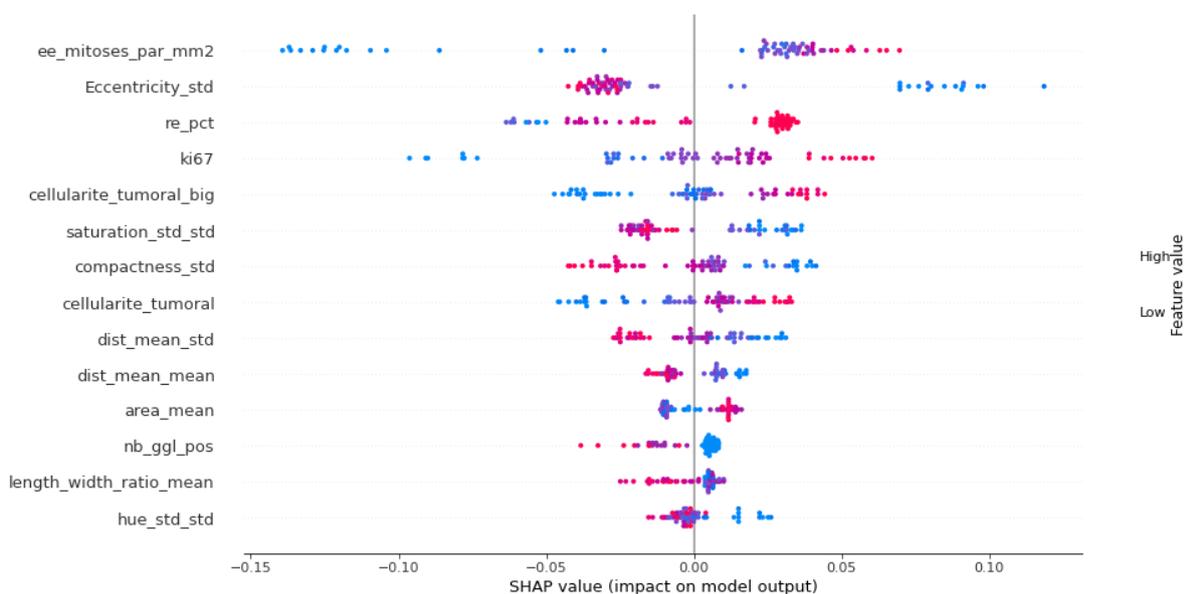


Figure 19 : SHAP value de la prédiction du sous-type intrinsèque par ML (random forest) sur la combinaison des features aux IHC et aux CCPT après élimination des features corrélées

En ce qui concerne l'interprétabilité, la figure 19 montre une répartition des features plus équilibrée que précédemment, avec les indicateurs de prolifération (compte

mitotique et Ki67) ayant plus d'importance que précédemment. Les features qui semblent importantes sont à nouveau l'`eccentricity_std`, la `cellularite_tumoral_big` et la `saturation_std_std`, couvrant des caractéristiques morphologiques tumorales variées.

3.3.3 - Conclusion préliminaire sur la prédiction par les features nucléaires

Les résultats obtenus pour la prédiction à partir des features seules, ou en combinaison aux CCPT, sont du même ordre que ceux obtenus pour le DL. Cette première observation nous fait penser que les CCPT extraits par le pathologiste en pratique courante contiennent déjà les informations pertinentes pour la prédiction du sous-type intrinsèque, et par conséquent de la survie.

La différence entre ces 2 méthodes d'analyse d'image réside donc essentiellement dans l'interprétabilité des résultats, qui étaient beaucoup plus tangibles pour les features, et qui ont soulevé plus de questions que ceux du DL. En effet, les premiers résultats de cette étude montrent que ces 2 sous-types pourraient se distinguer sur la cellularité tumorale, les tumeurs lumineuses B étant plus cellulaire que les A. Mais une autre caractéristique était inattendue : les tumeurs lumineuses A présenteraient un plus grand pléomorphisme nucléaire que les tumeurs lumineuses B. Ce qui va à l'encontre de ce que l'on attendrait avec le grade nucléaire. Les investigations ultérieures que nous avons effectuées pour tenter de comprendre cette différence, ont révélé un biais potentiel qui est la considération des noyaux de fibroblastes ou des noyaux des cellules endothéliales comme étant des noyaux tumoraux. Dans cette hypothèse, la cellularité étant plus faible dans les tumeurs lumineuses A, il y aurait plus de fibroblastes et de capillaires dans la zone tumorale, et les valeurs des features de ces noyaux viendraient alors créer artificiellement un pléomorphisme. Cependant, si aucun biais n'est mis en évidence, ce serait une donnée intéressante qui viendrait mettre en lumière le manque de pertinence de l'évaluation du grade nucléaire, et le besoin de peut-être le redéfinir.

Les écarts relativement étroits des valeurs des features retrouvées dans les SHAP value viennent consolider l'hypothèse de proximité morphologique des 2 sous-types, observée dans les performances du DL. L'exiguïté de cet écart était confirmée lorsque nous analysons certaines features en détail comme l'`eccentricity_std`, pour laquelle les valeurs minimales et maximales étaient en fait très proches, variant de 0,14 à 0,17 pour

des valeurs comprises entre 0 et 1. Cette étroitesse des valeurs des features peut légitimement nous faire questionner la significativité de ces résultats, qu'il conviendrait d'évaluer en réitérant l'étude sur un plus grand nombre de cas. Par ailleurs, nous avons cherché à visualiser ces différences en examinant les patchs tumoraux qui présentaient des valeurs extrêmes d'*eccentricity_std* (figure 20). La différence visuelle observée dans ces patchs rassure par rapport aux chiffres obtenus, puisqu'on observe une réelle différence entre ces 2 images, avec une impression de forte densité et d'homogénéité nucléaire dans le sous-type luminal B, avec des noyaux qui semblent être plus volumineux et plus ronds.

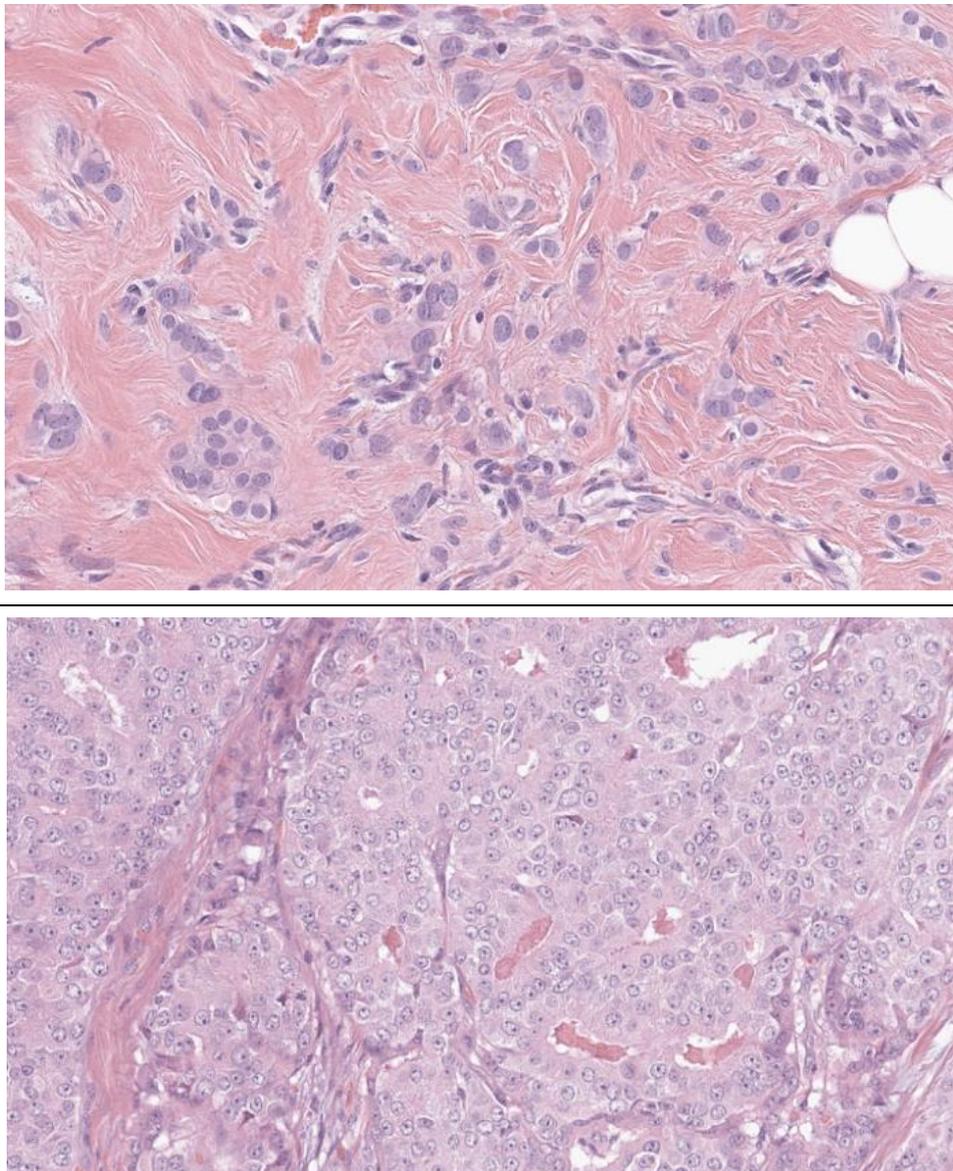


Figure 20 : analyse de patchs présentant des valeurs d'*eccentricity_std* extrêmes
(en haut: sous-type luminal A avec une *eccentricity_std* haute (0,17) ; en bas : sous-type luminal B avec une *eccentricity_std* basse (0,14))

L'interprétation des résultats pour les features a néanmoins demandé beaucoup de travail, conjointement avec les ingénieurs en charge du projet, pour comprendre ce que les features signifiaient, comment interpréter les méthodes d'explication (SHAP value) et évaluer les potentiels biais à éviter au long du processus. Également, la création de l'algorithme et le codage des différentes features à la main ont nécessité beaucoup plus de travail que l'utilisation de l'algorithme de DL.

Par ailleurs, le ML basé sur les features s'est retrouvé être la méthode où le surapprentissage était fréquent, et qui nécessitait des ajustements pour l'éviter. Cela était dû au ratio du nombre de features par rapport au nombre de tumeurs analysées, qui était trop élevé quand nous ne faisons pas assez de sélection dans les features.

Un autre point intéressant dans l'étude des features est que certaines, qui semblaient pertinentes, n'ont pourtant pas été sélectionnées pour les prédictions des algorithmes. Il s'agit en particulier des features de répartition cellulaire concernant les lymphocytes. Le DL ayant montré une tendance à porter un intérêt aux zones riches en lymphocytes, nous nous attendions à voir quelques-unes de ces features dans la sélection de features importantes. Cette discordance pourrait signifier que l'infiltrat lymphocytaire est trop inconstant parmi les cas pour mettre en évidence une différence significative, ou que les features codées pour l'évaluer ne sont pas optimaux. Ces résultats soulèvent néanmoins des questions quant à l'intérêt des TILS dans la catégorie des tumeurs lumineuses, qui jusqu'à présent n'a pas démontré d'intérêt. Nous n'avons pas eu le temps de mettre au point certaines features présentes dans la littérature comme l'orientation des noyaux, qui semble être un critère important dans l'étude de C. Lu et al. (116).

L'interprétabilité des résultats nous permet ici de montrer des tendances (un sous-type semble être plus cellulaire qu'un autre par exemple), et il serait intéressant de pouvoir visualiser des seuils pour les valeurs des features pertinentes. Mais cette extraction n'est pas simple à réaliser et elle nécessite un travail approfondi sur l'algorithme que nous n'avons pas eu le temps d'effectuer.

4 - Discussion des limites, forces et perspectives de l'étude

4.1 - Limites de l'étude

Une première limite de notre étude se situe dans l'hypothèse de départ. Nous aurions pu nous poser la question si un pathologiste pouvait prédire le sous-type intrinsèque sur des lames HE ou non. Puis établir de cette réponse, une ou plusieurs hypothèses préalables qui auraient pu orienter davantage les méthodes utilisées (calcul du nombre de cas en fonction de la différence attendue par exemple).

Une autre limite de notre étude réside dans la création des cohortes. Tout d'abord, ces 2 cohortes d'entraînement et de test sont séparées temporellement. Ce qui peut être à l'origine de biais liés à des changements de pratique (indications des tests moléculaires par exemple) ou liés à des variations dans les processus techniques en anatomopathologie (changement de réactifs, colorations, automates...). Il aurait été souhaitable d'inclure tous les cas au même moment et de créer les cohortes à posteriori, de manière randomisée. Par ailleurs, le nombre de cas dans nos cohortes est faible notamment pour des analyses d'image, et d'autant plus que la différence morphologique entre les 2 sous-types semble être étroite. Enfin, le monocentrisme de cette étude rend probablement les algorithmes peu utilisables sur un jeu de données extérieur, en particulier pour les algorithmes de DL. Le ML sur les features nucléaires devrait être plus facilement utilisable étant donné que leur développement repose sur une segmentation d'instance qui semble reproductible.

Ensuite, en ce qui concerne le développement des algorithmes, nous avons vu dans les résultats que les tumeurs semblaient être mixtes, constituées de zones ayant des caractéristiques lumineuses A et de zones ayant des caractéristiques lumineuses B. Or dans l'étiquetage des patches, aucune distinction n'était faite et tous les patches d'une tumeur étaient étiquetés avec le même sous-type. Ce qui n'est pas optimal et qui pourrait entraîner une baisse de performance des algorithmes. Il serait sans doute plus optimal d'entraîner les algorithmes sur les cas dont le test moléculaire est franchement en faveur d'un luminal A ou d'un luminal B (car ils devraient théoriquement être plus homogènes), et d'évaluer les performances sur les cas possédant un score plus intermédiaire.

Enfin, quelques difficultés persistent dans l'interprétation des résultats. Tout d'abord, il se pose la question de la significativité des différences observées entre chaque méthode. On souhaiterait avoir des intervalles de confiance pour chaque valeur, mais les expériences de cette étude n'ayant pas été reproduites plusieurs fois, il n'est pas possible pour le moment de calculer ces intervalles de confiance. On ne peut donc qu'observer des tendances d'amélioration ou de baisse des performances pour chaque méthode. Par ailleurs, pour le DL, il aurait pu être intéressant de faire une cohorte témoin avec des cas volontairement mal étiquetés, pour obtenir un résultat de prédiction de base, théoriquement proche de 50 %. Cela aurait été un garant d'une fiabilité de nos résultats car si ce résultat avait été de l'ordre de ceux obtenus dans notre étude (environ 70 %), nous aurions été beaucoup plus sceptiques concernant la valeur de nos résultats. De plus, pour les comparaisons entre les prédictions basées sur les CCPT et les prédictions basées sur l'analyse d'image, les cohortes ne comprennent pas le même nombre de cas, ce qui altère la comparabilité des résultats.

4.2 - Forces de l'étude

A notre connaissance, il s'agit de la première étude à évaluer la prédiction du sous-type intrinsèque luminal A ou B par des techniques d'analyse d'image. Les autres études qui se sont intéressées à la prédiction des altérations moléculaires ont souvent cherché à prédire les catégories de sous-types (luminal, basal-like, HER2+...) qui peuvent déjà être distinguées par immunohistochimie ou morphologie. Or, les analyses moléculaires sont demandées, en France, pour distinguer les 2 sous-types luminaux A et B, dont la seule approximation est celle proposée par la conférence de St Gallen et revisitée par Maisonneuve.

Un autre avantage de notre étude est d'avoir réalisé les analyses d'images sur les lames entières des tumeurs qui provenaient des blocs ayant servi à faire les tests moléculaires, plutôt que sur des TMA. De ce fait, les images analysées étaient fortement corrélées à l'analyse moléculaire. De plus, le fait d'avoir pu extraire les features sur la totalité de la tumeur permettait d'avoir des valeurs plus représentatives de l'ensemble de cette tumeur, que celles qui auraient été obtenues avec des biopsies ou des TMA.

Par ailleurs, à l'inverse des inconvénients présentés précédemment, le monocentrisme de notre étude présente un avantage pour l'analyse d'image ; celui de s'affranchir de certains biais dûs à des différences de techniques (coloration, épaisseur de coupe, fixation) ou de recrutement entre les centres. Cependant, si une étude multicentrique avait été menée, les cohortes auraient pu être constituées en randomisant sur ces caractéristiques.

Enfin, l'interaction fréquente avec les ingénieurs en charge du projet tout au long du développement des algorithmes, a permis d'optimiser continuellement le processus en identifiant les informations pertinentes et en évitant au maximum les biais d'apprentissage des algorithmes. Ce travail a également abouti à la création d'une interface numérique pour visualiser plus facilement les résultats, qui servira dans les autres projets d'intelligence artificielle qui suivront.

4.3 - Perspectives

Nous avons pu constater dans notre étude que les meilleures performances de prédiction étaient obtenues pour les IHC en combinaison aux CCPT. Et les variables qui étaient le plus discriminantes étaient les variables continues qui reflétaient la prolifération tumorale (Ki67 et nombre de mitoses). Or l'évaluation de ces caractéristiques est actuellement "manuelle" et est donc potentiellement source d'erreur et de variation. Il serait intéressant de réitérer cette évaluation en utilisant des CCPT objectivés de manière plus précise par analyse d'image, et exprimés en valeurs continues.

Dans cette étude nous avons commencé à visualiser concrètement les différences de valeurs dans les features qui influencent la prédiction, mais nous n'avons pas eu le temps d'approfondir cette recherche. Or ce processus nous semble essentiel pour pouvoir appliquer les résultats à la pratique courante. Toutefois, la dispersion étroite des valeurs n'aura peut-être que peu de traduction à l'échelle visuelle de l'homme et restera dépendante d'une analyse d'image par IA. Dans cette éventualité, il conviendra de se poser la question de la pertinence de critères ayant des valeurs sous-visuelles.

Dans notre étude, les performances de prédiction pour chaque méthode étaient proches avec des accuracy autour de 70-80%. Il serait intéressant de regarder en détail

les prédictions afin d'identifier les tumeurs mal classées, qui sont peut-être les mêmes dans chaque méthode, et si tel est le cas, d'essayer de comprendre les raisons potentielles. Car il se pourrait qu'une partie des tumeurs soit facilement prédictible et qu'une autre ne puisse obtenir un score discriminant. Cette hypothèse rappelle les différences de performances de prédiction observées par Jaber et al. dans leur étude (142), où les performances des modèles sur les tumeurs dites "low-confidence" qui présentaient un score moléculaire plus ambigu, étaient plus basse que sur les tumeurs non sélectionnées (56,73% contre 65,92% d'accuracy). Dans cette hypothèse, on pourrait tester une première fois les tumeurs en analyse d'image, et proposer l'analyse moléculaire uniquement aux cas dont la prédiction est équivoque.

Nous avons également constaté dans cette étude que des biais dans la sélection des noyaux effectués par la segmentation par DL, pouvaient être à l'origine des différences de valeurs observées pour certaines features. Cela laisse à penser que l'extraction de ces données présente probablement encore quelques imperfections qu'il faudrait corriger afin d'affiner les résultats.

Il serait également intéressant d'évaluer la prédiction du ROR et du pourcentage de récurrence à 10 ans après 5 ans d'hormonothérapie, ces valeurs étant plus informatives dans la pratique courante que le sous-type intrinsèque. Les premiers essais que nous avons réalisés pour cette tâche ont cependant montré des résultats très faibles, probablement dûs au fait que ce type d'analyse nécessiterait un plus grand nombre de cas. Toutefois, le développement d'un algorithme de prédiction du résultat de l'analyse moléculaire n'est peut-être pas le plus pertinent, et il serait préférable de développer un algorithme de prédiction à partir des données brutes de survie. Car l'analyse moléculaire donne une prédiction de la survie qui présente une marge d'erreur. Or l'algorithme basé sur l'analyse d'image pourrait très bien surpasser les performances de l'analyse moléculaire, mais si nous l'évaluons sur la prédiction du résultat de cette analyse, les performances seront très probablement moins bonnes.

L'analyse des features nucléaires a montré qu'elle pouvait permettre de distinguer 2 types de noyaux. Il serait donc intéressant de développer cette technique pour extraire davantage de features liées aux différentes populations cellulaires (histiocytes, fibroblastes, cellules endothéliales ...).

Dans les analyses par DL que nous avons réalisées, nous n'avons pas utilisé de technique de multi-scaling, qui permettent d'inclure les différents niveaux de zoom de la WSI. Ces techniques permettent de capter des caractéristiques globales (forme et contour de la tumeur par exemple) et des caractéristiques plus petites (forme et contour des noyaux). On élargirait ainsi les features que le DL peut traiter, ce qui devrait augmenter les performances de prédiction.

A l'image de ce qu'ont réalisé Wang et al. (151) avec leur algorithme de grading Deep Grade, il serait intéressant d'évaluer les performances d'un tel modèle, entraîné à classer les tumeurs en bas et haut grade, sur la prédiction du sous-type intrinsèque. Il se pourrait qu'un tel modèle sépare ces tumeurs en 2 catégories de la même manière que les analyses moléculaires.

Comme évoqué précédemment, notre étude présente l'avantage de se baser sur les mêmes blocs ayant servi aux analyses moléculaires. L'utilisation de la transcriptomique spatiale, qui permet de mesurer l'expression de gènes dans différentes localisations du prélèvement, permettrait de resserrer les liens entre les données moléculaires et les features.

Dans notre étude, les 2 méthodes d'intelligence artificielle ont été traitées séparément, puis comparées. Mais les analyses en DL semblaient révéler des zones d'intérêt qui auraient pu servir à faire une extraction de features ciblée par la suite. Dans une même logique de combinaison des 2 méthodes, on pourrait également se servir des features extraites par les couches profondes du DL, comme substrat d'apprentissage pour des modèles de machine learning. Les performances surpasseraient peut-être celles observées pour le modèle de DL pur.

5 - Conclusion

L'intelligence artificielle gagne en intérêt dans l'anatomopathologie notamment pour la prédiction de la survie des patients à partir de l'analyse d'images microscopiques de leur cancer. Deux méthodes s'opposent dans ce domaine en fonction de l'autonomie laissée à la machine : le deep learning et le machine learning plus classique (méthodes tabulaires).

Dans cette étude, nous avons comparé les performances de prédiction du sous-type intrinsèque luminal A vs luminal B de cancers du sein RH+ HER2- donné par la signature Prosigna® PAM50, premièrement par des modèles de machine learning s'appuyant sur les résultats des immunohistochimies RE, RP et Ki67, puis des autres critères clinico-pathologiques traditionnels ; et secondairement, par des méthodes d'analyse d'image utilisant le deep learning d'une part, ou le machine learning s'appuyant sur des features nucléaires extraites au préalable d'autre part. Nous avons également combiné ces méthodes entre elles afin d'évaluer l'apport de chacune sur la prédiction en pratique courante effectuée sur la base des seuls résultats d'immunohistochimie. A notre connaissance, c'est la première étude qui s'intéresse à ce sujet.

Les résultats obtenus pour les tâches de prédiction dans ce dataset monocentrique montrent des performances de prédiction sensiblement équivalentes pour chaque méthode, prise séparément, obtenant des accuracy de l'ordre de 70 %. La combinaison des différentes méthodes de machine learning ou de deep learning avec les critères clinico-pathologiques traditionnels semblent améliorer légèrement la prédiction de base faite sur l'immunohistochimie, avec des améliorations sensiblement identiques pour les 2 méthodes d'analyse d'image (deep learning et machine learning avec extraction de features au préalable).

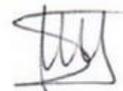
L'analyse de l'interprétabilité des résultats montrent que certains critères ou features semblent plus importants que d'autres pour la prédiction du sous-type intrinsèque. En effet, dans les critères clinico-pathologiques traditionnels, ce sont le nombre de mitoses par mm² et la prolifération cellulaire évaluée par le Ki67 qui semblent surpasser les autres critères pour la prédiction du sous-type intrinsèque. L'algorithme de deep learning semble par ailleurs porter un intérêt aux zones riches en lymphocytes pour le sous-type luminal B. L'analyse des features nucléaires semblent montrer que

les tumeurs lumineuses B correspondent à des tumeurs ayant une cellularité augmentée, des noyaux globalement plus arrondis mais aux contours plus irréguliers, et un pléomorphisme plus faible ; contrairement aux tumeurs lumineuses A qui présenteraient un pléomorphisme plus marqué, des noyaux plus elliptiques, et une densité cellulaire moindre.

Ce travail montre que les modèles de machine learning basés sur l'extraction de features permettent d'obtenir des résultats sensiblement identiques à ceux obtenus par deep learning pour la tâche de classification choisie. En outre, leur caractère explicable permet d'avoir une approche critique du modèle en ce qui concerne les données connues de la science, et d'ouvrir des pistes de recherche sur les critères plus inattendus. A noter que ce travail est la première brique d'un projet plus vaste (projet APRIORICS) qui fera intervenir l'extraction de features, non seulement des noyaux, mais de dizaines d'objets élémentaires des tumeurs, et non plus sur quelques dizaines d'images microscopiques mais sur un millier. Les travaux réalisés dans le cadre de mon projet de thèse ont permis d'aborder et de résoudre un certain nombre de questions qui profiteront à l'ensemble du projet APRIORICS.


Professeur Jeanick SELVES
Département Anatomie et Cytologie Pédiatriques
INSTITUT UNIVERSITAIRE DU CANCER
1 avenue F-870 Joliot-Curie
31059 TOULOUSE CEDEX 9
Tél. 05 31 15 63 22

Vu et permis d'imprimer
Le Président de l'Université Toulouse III – Paul Sabatier
Faculté de Santé
Par délégation,
La Doyenne-Directrice
Du Département de Médecine, Maïeutique, Paramédical
Professeure Odile RAUZY



Bibliographie

- (1) Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021 May;71(3):209-249.
- (2) Defossez G, le Guyader-Peyrou S, Uhry Z, Grosclaude P, Colonna M, Dantony E, et al. Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine entre 1990 et 2018. Volume 1 – Tumeurs solides. Saint-Maurice (Fra): Santé publique France, 2019. 372 p.
- (3) Amin MB, Edge SB, Greene FL, et al, eds. *AJCC Cancer Staging Manual*. 8th ed. New York: Springer; 2017.
- (4) Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet Lond Engl.* 14 mai 2005;365(9472):1687-717.
- (5) Rossi L, Stevens D, Pierga JY, Lerebours F, Reyat F, Robain M, Asselain B, Rouzier R. Impact of Adjuvant Chemotherapy on Breast Cancer Survival: A Real-World Population. *PLoS One.* 2015 Jul 27;10(7):e0132853.
- (6) Early Breast Cancer Trialists' Collaborative Group: Polychemotherapy for early breast cancer: An overview of the randomized trials. *Lancet* 352:930-942, 1998
- (7) Early Breast Cancer Trialists' Collaborative Group: Tamoxifen for early breast cancer: An overview of randomised trials. *Lancet* 351:1451-1467, 1998
- (8) Adami HO, Malke B, Holmberg L, Persson I, Stone B. The relation between survival and age at diagnosis in breast cancer. *N Engl J Med.* 28 août 1986;315(9):559-63.
- (9) Kollias J, Elston CW, Ellis IO, Robertson JF, Blamey RW. Early-onset breast cancer histopathological and prognostic considerations. *Br J Cancer.* 1997;75(9):1318-23.
- (10) Botteri E, Bagnardi V, Rotmensz N, Gentilini O, Disalvatore D, Bazolli B, et al. Analysis of local and regional recurrences in breast cancer after conservative surgery. *Ann Oncol Off J Eur Soc Med Oncol.* avr 2010;21(4):723-8.
- (11) Kim E-K, Noh WC, Han W, Noh D-Y. Prognostic significance of young age (<35 years) by subtype based on ER, PR, and HER2 status in breast cancer: a nationwide registry-based study. *World J Surg.* juin 2011;35(6):1244-53.
- (12) Kroman N, Jensen MB, Wohlfahrt J, Mouridsen HT, Andersen PK, Melbye M. Factors influencing the effect of age on prognosis in breast cancer: population based study. *BMJ.* 19 févr 2000;320(7233):474-8.
- (13) Fredholm H, Eaker S, Frisell J, Holmberg L, Fredriksson I, Lindman H. Breast cancer in young women: poor survival despite intensive treatment. *PLoS One.* 2009 Nov 11;4(11):e7695.
- (14) Sabiani L, Houvenaeghel G, Heinemann M, Reyat F, Classe JM, Cohen M, Garbay JR, Giard S, Charitansky H, Chopin N, Rouzier R, Daraï E, Coutant C, Azuar P, Gimbergues P, Villet R, Tunon de Lara C, Lambaudie E. Breast cancer in young women: Pathologic features and molecular phenotype. *Breast.* 2016 Oct;29:109-16.
- (15) Louwman WJ, Janssen-Heijnen MLG, Houterman S, Voogd AC, van der Sangen MJC, Nieuwenhuijzen G a. P, et al. Less extensive treatment and inferior prognosis for breast cancer patient with comorbidity: a population-based study. *Eur J Cancer Oxf Engl* 1990. mars 2005;41(5):779-85.

- (16) Fitzgibbons PL, Page DL, Weaver D et al. Prognostic factors in breast cancer. College of American Pathologists Consensus Statement 1999. *Arch Pathol Lab Med.* 2000;124: 966-78.
- (17) Sant M, Allemani C, Berrino F et al. Breast carcinoma survival in Europe and the United States. *Cancer.* 2004;100: 715-22.
- (18) Fung F, Cornacchi SD, Vanniyasingam T, Dao D, Thabane L, Simunovic M, et al. Predictors of 5-year local, regional, and distant recurrent events in a population-based cohort of breast cancer patients. *Am J Surg.* févr 2017;213(2):418-25.101
- (19) Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer.* 1 janv 1989;63(1):181-7.
- (20) Sant M, Allemani C, Berrino F et al. Breast carcinoma survival in Europe and the United States. *Cancer.* 2004;100: 715-22.
- (21) Fisher B, Bauer M, Wickerham DL, Redmond CK, Fisher ER, Cruz AB, et al. Relation of number of positive axillary nodes to the prognosis of patients with primary breast cancer. An NSABP update. *Cancer.* 1 nov 1983;52(9):1551-7.
- (22) WHO Classification of Tumours Editorial Board. Breast tumours. Lyon (France): International Agency for Research on Cancer; 2019. (WHO classification of tumours series, 5th ed.; vol 2).
- (23) Bloom HJ, Richardson WW: Histological grading and prognosis in breast cancer: A study of 1409 cases of which 359 have been followed for 15 years. *Br J Cancer* 11:359-377, 1957
- (24) Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology.* nov 1991;19(5):403-10.
- (25) Rakha EA, El-Sayed ME, Lee AHS, Elston CW, Grainge MJ, Hodi Z, et al. Prognostic Significance of Nottingham Histologic Grade in Invasive Breast Carcinoma. *J Clin Oncol.* juill 2008;26(19):3153-8.
- (26) Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Res Treat.* 1992;22: 207-19.
- (27) Balslev I, Axelsson CK, Zedeler K, et al: The Nottingham Prognostic Index applied to 9,149 patients from the studies of the Danish Breast Cancer Cooperative Group (DBCG). *Breast Cancer Res Treat* 32:281-290, 1994
- (28) Okugawa H, Yamamoto D, Uemura Y, et al: Prognostic factors in breast cancer: The value of the Nottingham Prognostic Index for patients treated in a single institution. *Surg Today* 35:907-911, 2005
- (29) Frierson HF, Wolber RA, Berean KW, Franquemont DW, Gaffey MJ, Boyd JC, et al. Interobserver reproducibility of the Nottingham modification of the Bloom and Richardson histologic grading scheme for infiltrating ductal carcinoma. *Am J Clin Pathol* 1995;103:195– 8.
- (30) Robbins P, Pinder S, de Klerk, Dawkins H, Harvey J, Sterrett G, et al. Histologic grading of breast carcinomas: a study of interobserver agreement. *Hum Pathol* 1995;26:873–9.
- (31) Dalton LW, Page DL, DuPont WD. Histologic grading of breast carcinoma: a reproducibility study. *Cancer* 1994;73: 2765–70.
- (32) Harvey JM, De Klerk NH, Sterrett GF. Histological grading in breast cancer: interobserver agreement, and relation to other prognostic variables including ploidy. *Pathology* 1992; 24:63– 8.
- (33) Theissig F, Kunze KD, Haroske G, Meyer W. Histologic grading of breast cancer: interobserver reproducibility and prognostic significance. *Pathol Res Pract* 1990;186:732– 6.

- (34) Pinder SE, Ellis IO, Galea M, et al. Pathological prognostic factors in breast cancer. III. Vascular invasion: relationship with recurrence and survival in a large study with long-term follow-up. *Histopathology* 1994;24:41–7.
- (35) Lauria R, Perrone F, Carlomagno C, et al. The prognostic value of lymphatic and blood vessel invasion in operable breast cancer. *Cancer* 1995;76:1772–8.
- (36) Lee AHS, Pinder SE, Macmillan RD, Mitchell M, Ellis IO, Elston CW, et al. Prognostic value of lymphovascular invasion in women with lymph node negative invasive breast carcinoma. *Eur J Cancer Oxf Engl* 1990. févr 2006;42(3):357-62.
- (37) Schoppmann SF, Bayer G, Aumayr K, Taucher S, Geleff S, Rudas M, et al. Prognostic value of lymphangiogenesis and lymphovascular invasion in invasive breast cancer. *Ann Surg.* août 2004;240(2):306-12.
- (38) Orbo A, Stalsberg H, Kunde D. Topographic criteria in the diagnosis of tumor emboli in intramammary lymphatics. *Cancer* 1990;66:972–7.
- (39) Galon J, Costes A, Sanchez-Cabo F et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 2006; 313: 1960–1964.
- (40) Loi S, Sirtaine N, Piette F et al. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. *J Clin Oncol* 2013; 31: 860–867.
- (41) Loi S, Michiels S, Salgado R et al. Tumor infiltrating lymphocytes is prognostic and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. *Ann Oncol* 2014; 25(8): 1544–1550.
- (42) Denkert C, von Minckwitz G, Darb-Esfahani S, Lederer B, Heppner BI, Weber KE, Budczies J, Huober J, Klauschen F, Furlanetto J, Schmitt WD, Blohmer JU, Karn T, Pfitzner BM, Kümmel S, Engels K, Schneeweiss A, Hartmann A, Noske A, Fasching PA, Jackisch C, van Mackelenbergh M, Sinn P, Schem C, Hanusch C, Untch M, Loibl S. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol.* 2018 Jan;19(1):40-50.
- (43) Loi S, Drubay D, Adams S, Pruneri G, Francis PA, Lacroix-Triki M, Joensuu H, Dieci MV, Badve S, Demaria S, Gray R, Munzone E, Lemonnier J, Sotiriou C, Piccart MJ, Kellokumpu-Lehtinen PL, Vingiani A, Gray K, Andre F, Denkert C, Salgado R, Michiels S. Tumor-Infiltrating Lymphocytes and Prognosis: A Pooled Individual Patient Analysis of Early-Stage Triple-Negative Breast Cancers. *J Clin Oncol.* 2019 Mar 1;37(7):55
- (44) Salgado R, Denkert C, Demaria S, Sirtaine N, Klauschen F, Pruneri G, Wienert S, Van den Eynden G, Baehner FL, Penault-Llorca F, Perez EA, Thompson EA, Symmans WF, Richardson AL, Brock J, Criscitiello C, Bailey H, Ignatiadis M, Floris G, Sparano J, Kos Z, Nielsen T, Rimm DL, Allison KH, Reis-Filho JS, Loibl S, Sotiriou C, Viale G, Badve S, Adams S, Willard-Gallo K, Loi S; International TILs Working Group 2014. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol.* 2015 Feb;26(2):259-71. doi: 10.1093/annonc/mdu450. Epub 2014 Sep 11.
- (45) Domagala W , Markiewski M , Harezga B , Dukowicz A , Osborn M . Prognostic significance of tumor cell proliferation rate as determined by the MIB-1 antibody in breast carcinoma: its relationship with vimentin and p53 protein . *Clin Cancer Res.* 1996 ; 2 (1) : 147 – 154 .
- (46) Trihia H , Murray S , Price K , et al . Ki-67 expression in breast carcinoma: its association with grading systems, clinical parameters, and other prognostic factors — a surrogate marker? *Cancer .* 2003 ; 97 (5) : 1321 – 1331 .

- (47) de Azambuja E, Cardoso F, de Castro G Jr, et al. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients. *Br J Cancer*. 2007; 96(10): 1504–1513.
- (48) Criscitiello C, Disalvatore D, De Laurentiis M, Gelao L, Fumagalli L, Locatelli M, Bagnardi V, Rotmensz N, Esposito A, Minchella I, De Placido S, Santangelo M, Viale G, Goldhirsch A, Curigliano G. High Ki-67 score is indicative of a greater benefit from adjuvant chemotherapy when added to endocrine therapy in luminal B HER2 negative and node-positive breast cancer. *Breast*. 2014 Feb;23(1):69-75.
- (49) Pichon MF, Broet P, Magdelenat H, Delarue JC, Spyrtatos F, Basuyau JP, et al. Prognostic value of steroid receptors after long-term follow-up of 2257 operable breast cancers. *Br J Cancer*. juin 1996;73(12):1545-51.
- (50) Prat A, Cheang MC, Martin M et al. Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer. *J Clin Oncol*. 2013;31: 203-9.
- (51) Arpino G, Weiss H, Lee AV et al. Estrogen receptor-positive, progesterone receptor-negative breast cancer: association with growth factor receptor expression and tamoxifen resistance. *J Natl Cancer Inst*. 2005;97: 1254-61.
- (52) Rakha EA, Reis-Filho JS, Ellis IO. Basal-like breast cancer: a critical review. *J Clin Oncol*. 2008;26: 2568-81.
- (53) Lippman ME, Allegra JC. Quantitative estrogen receptor analyses: the response to endocrine and cytotoxic chemotherapy in human breast cancer and the disease-free interval. *Cancer*. 1980;46: 2829-34.
- (54) Ravdin PM, Green S, Dorr TM et al. Prognostic significance of progesterone receptor levels in estrogen receptor-positive patients with metastatic breast cancer treated with tamoxifen: results of a prospective Southwest Oncology Group study. *J Clin Oncol*. 1992;10: 1284-91
- (55) Osborne CK, Yochmowitz MG, Knight WA, 3rd, McGuire WL. The value of estrogen and progesterone receptors in the treatment of breast cancer. *Cancer*. 1980;46: 2884-8.
- (56) Arpino G, Weiss H, Lee AV et al. Estrogen receptor-positive, progesterone receptor-negative breast cancer: association with growth factor receptor expression and tamoxifen resistance. *J Natl Cancer Inst*. 2005;97: 1254-61.
- (57) Rubin I, Yarden Y. The basic biology of HER2. *Ann Oncol*. 2001;12 Suppl 1: S3-8.
- (58) Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*. 9 janv 1987;235(4785):177-82.
- (59) Joensuu H, Isola J, Lundin M et al. Amplification of erbB2 and erbB2 expression are superior to estrogen receptor status as risk factors for distant recurrence in pT1NoMo breast cancer: a nationwide population-based study. *Clin Cancer Res*. 2003;9: 923-30.
- (60) Chia S, Norris B, Speers C et al. Human epidermal growth factor receptor 2 overexpression as a prognostic factor in a large tissue microarray series of node-negative breast cancers. *J Clin Oncol*. 2008;26: 5697-704.
- (61) Slamon DJ, Godolphin W, Jones LA et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science*. 1989;244: 707-12.
- (62) Dowsett M, Houghton J, Iden C, Salter J, Farndon J, A'Hern R, et al. Benefit from adjuvant tamoxifen therapy in primary breast cancer patients according oestrogen receptor, progesterone receptor, EGF receptor and HER2 status. *Ann Oncol Off J Eur Soc Med Oncol*. mai 2006;17(5):818-26.
- (63) Dowsett M, Allred C, Knox J, Quinn E, Salter J, Wale C, et al. Relationship between quantitative estrogen and progesterone receptor expression and human epidermal

growth factor receptor 2 (HER-2) status with recurrence in the Arimidex, Tamoxifen, Alone or in Combination trial. *J Clin Oncol Off J Am Soc Clin Oncol*. 1 mars 2008;26(7):1059-65.

(64) Piccart-Gebhart MJ, Procter M, Leyland-Jones B et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med*. 2005;353:1659-72.129

(65) Romond EH, Perez EA, Bryant J et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med*. 2005;353:1673-84.

(66) Fekih M, Petit T, Zarca D, Guinebretière JM, André F, Pierga JY, et al. Utilisation de référentiels et hétérogénéité décisionnelle des indications de chimiothérapie adjuvante dans les cancers du sein exprimant les récepteurs hormonaux, HER2-négatifs: résultats d'un sondage national en France. *Bull Cancer* 2014;101(10):918-24.

(67) Haute Autorité de Santé, Rapport d'évaluation technologique Utilité clinique des signatures génomiques dans le cancer du sein de stade précoce, 2019.

(68) Joensuu H, Pylkkänen L, Toikkanen S. Long-term survival in node-positive breast cancer treated by locoregional therapy alone. *Br J Cancer*. sept 1998;78(6):795-9.

(69) Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B, Desmedt C, Larsimont D, Cardoso F, Peterse H, Nuyten D, Buyse M, Van de Vijver MJ, Bergh J, Piccart M, Delorenzi M. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst*. 2006 Feb 15;98(4):262-72.

(70) Bachelot T, Bouzid K, Delozier T, Lortholary A, Petit T. État des lieux et mise à jour des traitements systémiques adjuvants. *Oncologie* 2011;13:698-702.

(71) Strasser-Weippl K, Goss PE. Competing risks in low-risk breast cancer. *Am Soc Clin Oncol Educ Book* 2013:32- 9.

(72) Campbell HE, Taylor MA, Harris AL, Gray AM. An investigation into the performance of the Adjuvant! Online prognostic programme in early breast cancer for a cohort of patients in the United Kingdom. *Br J Cancer*. 2009 Oct 6;101(7):1074-84.

(73) Mook S, Schmidt MK, Rutgers EJ, van de Velde AO, Visser O, Rutgers SM, Armstrong N, van't Veer LJ, Ravdin PM. Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: a hospital-based retrospective cohort study. *Lancet Oncol*. 2009 Nov;10(11):1070-6.

(74) Rakha EA, Soria D, Green AR, Lemetre C, Powe DG, Nolan CC et al (2014) Nottingham prognostic index plus (NPI+): a modern clinical decision making tool in breast cancer. *Br J Cancer* 110(7):1688–1697.

(75) Green AR, Soria D, Stephen J, Powe DG, Nolan CC, Kunkler I et al (2016) Nottingham prognostic index plus: validation of a clinical decision making tool in breast cancer in an independent series. *J Pathol: Clin Res* 2(1):32–40

(76) Green AR, Soria D, Powe DG, Nolan CC, Aleskandarany M, Szász MA, et al. Nottingham prognostic index plus (NPI+) predicts risk of distant metastases in primary breast cancer. *Breast Cancer Res Treat*. 2016;157(1):65-75.

(77) Cuzick J, Dowsett M, Pineda S, Wale C, Salter J, Quinn E, et al. Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. *J Clin Oncol Off J Am Soc Clin Oncol*. 10 nov 2011;29(32):4273-8.

(78) Dowsett M, Sestak I, Regan MM, Dodson A, Viale G, Thurlimann B, Colleoni M, Cuzick J (2018) Integration of clinical variables for the prediction of late distant recurrence in patients with Estrogen receptor-positive breast cancer treated with 5 years of endocrine therapy: CTS5. *J Clin Oncol* 36(19):1941–1948

- (79) Richman J, Ring A, Dowsett M, Sestak I. Clinical validity of clinical treatment score 5 (CTS5) for estimating risk of late recurrence in unselected, non-trial patients with early oestrogen receptor-positive breast cancer. *Breast Cancer Res Treat.* 2021 Feb;186(1):115-123.
- (80) Lakhanpal R, Sestak I, Shadbolt B, Bennett GM, Brown M, Phillips T, et al. IHC4 score plus clinical treatment score predicts locoregional recurrence in early breast cancer. *Breast Edinb Scotl.* oct 2016;29:147-52.
- (81) Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al. The eighth edition AJCC cancer staging manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin* 2017;67(2):93-9.
- (82) Institut national du cancer. Cancer du sein infiltrant non métastatique. Questions d'actualité. Boulogne Billancourt: INCA; 2012.
- (83) Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 2013;10(2):e1001381.
- (84) Vickers AJ, Kattan MW, Daniel S. Method for evaluating prediction models that apply the results of randomized trials to individual patients. *Trials* 2007;8:14.
- (85) Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162(1):W1-73.
- (86) Perou CM, Sorlie T, Eisen MB et al. Molecular portraits of human breast tumours. *Nature.* 2000;406: 747-52.
- (87) Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A.* 11 sept 2001;98(19):10869-74.
- (88) Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A.* 8 juill 2003;100(14):8418-23.
- (89) van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature.* 2002;415:530-536.
- (90) Buyse, M. et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J. Natl. Cancer Inst.* 98, 1183-1192 (2006).
- (91) Cronin M, Sangli C, Liu ML, et al. Analytic validation of the Oncotype Dx genomic diagnostic test for recurrence prognosis and therapeutic response prediction in node-negative, estrogen receptor-positive breast cancer. *Clin Chem.* 2007;53:1084-91.
- (92) Partin J, Mamounas E. Impact of the 21-gene recurrence score assay compared with standard clinicopathologic guidelines in adjuvant therapy selection for node-negative, estrogen receptor-positive breast cancer. *Ann Surg Oncol.* 2011;18(12):3399-406.
- (93) Parker JS, Mullins M, Cheang MC et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol.* 2009;27: 1160-7.
- (94) Rapport d'évaluation technologique : utilité des signatures génomiques dans le cancer du sein de stade précoce, 2019.
- (95) Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, et al. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res BCR.* 2008;10(4):R65.

- (96) Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, Senn HJ; Panel members. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol.* 2013 Sep;24(9):2206-23.
- (97) Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, Ellis M, Henry NL, Hugh JC, Lively T, McShane L, Paik S, Penault-Llorca F, Prudkin L, Regan M, Salter J, Sotiriou C, Smith IE, Viale G, Zujewski JA, Hayes DF; International Ki-67 in Breast Cancer Working Group. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst.* 2011 Nov 16;103(22):1656-64.
- (98) Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ; Panel members. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol.* 2011 Aug;22(8):1736-47.
- (99) Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, Senn HJ; Panel members. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol.* 2013 Sep;24(9):2206-23.
- (100) Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, Thürlimann B, Senn HJ; Panel Members. Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann Oncol.* 2015 Aug;26(8):1533-46.
- (101) Nielsen TO, Leung SCY, Rimm DL, Dodson A, Acs B, Badve S, Denkert C, Ellis MJ, Fineberg S, Flowers M, Kreipe HH, Laenkholm AV, Pan H, Penault-Llorca FM, Polley MY, Salgado R, Smith IE, Sugie T, Bartlett JMS, McShane LM, Dowsett M, Hayes DF. Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. *J Natl Cancer Inst.* 2021 Jul 1;113(7):808-819.
- (102) Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, et al. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res BCR.* 2008;10(4):R65.
- (103) Maisonneuve P, Disalvatore D, Rotmensz N, Curigliano G, Colleoni M, Dellapasqua S, Pruneri G, Mastropasqua MG, Luini A, Bassi F, Paganì G, Viale G, Goldhirsch A. Proposed new clinicopathological surrogate definitions of luminal A and luminal B (HER2-negative) intrinsic breast cancer subtypes. *Breast Cancer Res.* 2014 Jun 20;16(3):R65.
- (104) Mendelsohn ML, Kolman WA, Perry B, Prewitt JM. Computer analysis of cell images. *Postgrad Med.* 1965 Nov;38(5):567-73.
- (105) Pantanowitz L, Valenstein PN, Evans AJ, Kaplan KJ, Pfeifer JD, Wilbur DC, et al. Review of the current state of whole slide imaging in pathology. *J Pathol Inform* 2011;2:36.
- (106) Madabhushi, A. & Lee, G. Image analysis and machine learning in digital pathology: challenges and opportunities. *Med. Image Anal.* 33, 170–175.
- (107) Basavanahally, A. et al. Incorporating domain knowledge for tubule detection in breast histopathology using O'Callaghan neighborhoods. *Proc. SPIE* 7963, 796310 (2011).
- (108) Khan, A. M., & Sirinukunwattana, K., & Rajpoot, N. A global covariance descriptor for nuclear atypia scoring in breast histopathology images. *IEEE J. Biomed. Heal. Inform.* 19, 1637–1647 (2015).

- (109) Veta, M. et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Med. Image Anal.* 20, 237–248 (2015).
- (110) Diamond, J., Anderson, N. H., Bartels, P. H., Montironi, R. & Hamilton, P. W. The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia. *Hum. Pathol.* 35, 1121–1131 (2004).
- (111) Kong J, Cooper LAD, Wang F, Gutman DA, Gao J, Chisolm C, Sharma A, Pan T, Van Meir EG, Kurc TM, Moreno CS, Saltz JH, Brat DJ. Integrative, multimodal analysis of glioblastoma using TCGA molecular data, pathology images, and clinical outcomes. *IEEE Trans Biomed Eng.* 2011;58(12):3469–74.
- (112) Yu, K.-H. et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat. Commun.* 7:12474
- (113) Wang, S., Chen, A., Yang, L. et al. Comprehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome. *Sci Rep* 8, 10393 (2018).
- (114) Veta M, Kornegoor R, Huisman A, Verschuur-Maes AH, Viergever MA, Pluim JP, van Diest PJ. Prognostic value of automatically extracted nuclear morphometric features in whole slide images of male breast cancer. *Mod Pathol.* 2012 Dec;25(12):1559-65.
- (115) Tambasco M, Eliasziw M, Magliocco AM. Morphologic complexity of epithelial architecture for predicting invasive breast cancer survival. *J Transl Med.* 2010 Dec 31;8:140.
- (116) Lu C, Romo-Bucheli D, Wang X, Janowczyk A, Ganesan S, Gilmore H, Rimm D, Madabhushi A. Nuclear shape and orientation features from H&E images predict survival in early-stage estrogen receptor-positive breast cancers. *Lab Invest.* 2018 Nov;98(11):1438-1448.
- (117) Yuan, Y. et al. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Sci. Transl. Med.* 4, 157ra143 (2012).
- (118) Xu, J., Luo, X., Wang, G., Gilmore, H. & Madabhushi, A. A Deep convolutional neural network for segmenting and classifying epithelial and stromal
- (119) Hou, L. et al. Patch-based convolutional neural network for whole slide tissue image classification. in *Proc. CVPR*, pp. 2424-2433 (2016).
- (120) Xu, J., Luo, X., Wang, G., Gilmore, H. & Madabhushi, A. A Deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. *Neurocomputing* 191, 214–223 (2016).
- (121) Cruz-Roa, A. A., Ovalle, J. E. A., Madabhushi, A. & Gonzalez, F. A. O. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. *Lecture Notes in Computer Science*, 8150 (2013).
- (122) Cruz-Roa A, Basavanhally A, González F, Gilmore H, Feldman M, Ganesan S, Shih N, Tomaszewski J, Madabhushi A. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks In: Gurcan MN, Madabhushi A, editors. *SPIE Medical Imaging*. San Diego, USA: SPIE; 2014. p. 904103.
- (123) Wang, D., Khosla, A., Gargeya, R., Irshad, H. & Beck, A. H. Deep learning for identifying metastatic breast cancer (2016). Preprint at <http://arxiv.org/abs/1606.05718>
- (124) Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA.* 2017;318(22):2199-2210.

- (125) Cireşan, D. C., Giusti, A., Gambardella, L. M. & Schmidhuber, J. Mitosis detection in breast cancer histology images with deep neural networks. *Proc. MICCAI 6*, 411–418 (2013).
- (126) Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. *Anal Quant Cytol Histol.* 2001;23(4):291–9.
- (127) Liu Y, Gadepalli K, Norouzi M, Dahl GH, Kohlberger T, Boyko A, et al. Detecting cancer metastases on gigapixel pathology images. *arXiv.* 2017. <https://arxiv.org/abs/1703.02442>.
- (128) Ström P, Kartasalo K, Olsson H, et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol.* 2020;21:222-232.
- (129) Bulten W, Pinckaers H, van Boven H, et al. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. *Lancet Oncol.* 2020;21:233-241.
- (130) Motlagh MH, Jannesari M, Aboulkheyr H, Khosravi P, Elemento O, Totonchi M, Hajirasouliha I: Breast cancer histopathological image classification: a deep learning approach. *bioRxiv* 2018:242818.
- (131) Xie J, Liu R, Luttrell J, Zhang C. Deep learning based analysis of histopathological images of breast cancer. 2019;10(80):1-19.
- (132) Jiang Y, Chen L, Zhang H, Xiao X. Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. *PLoS One.* 2019;14(3):e0214587.
- (133) Alom, Md. Zahangir & Taha, Tarek & Yakopcic, Chris & Westberg, Stefan & Sidike, Paheding & Nasrin, Mst & Hasan, Mahmudul & Essen, Brian & Awwal, Abdul & Asari, Vijayan. (2019). A State-of-the-Art Survey on Deep Learning Theory and Architectures. *Electronics.* 8.292.10.3390/electronics8030292.
- (134) J. Su, D. V. Vargas and K. Sakurai, "One Pixel Attack for Fooling Deep Neural Networks," in *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 5, pp. 828-841, Oct. 2019, doi: 10.1109/TEVC.2019.2890858.
- (135) Seyyed-Kalantari, L., Zhang, H., McDermott, M.B.A. et al. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. *Nat Med* 27, 2176–2182 (2021).
- (136) Nicolas Davis, From oximeters to AI, where bias in medical devices may lurk, 2021, *The Guardian*.
<https://www.theguardian.com/society/2021/nov/21/from-oximeters-to-ai-where-bias-in-medical-devices-may-lurk>
- (137) Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine.* 2019;17(1):1–9.
- (138) Khened, M., Kori, A., Rajkumar, H. et al. A generalized deep learning framework for whole-slide image segmentation and analysis. *Sci Rep* 11, 11579 (2021).
- (139) Diao, J.A., Wang, J.K., Chui, W.F. et al. Human-interpretable image features derived from densely mapped cancer pathology slides predict diverse molecular phenotypes. *Nat Commun* 12, 1613 (2021).
- (140) Coudray, N., Ocampo, P.S., Sakellaropoulos, T. et al. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nat Med* 24, 1559–1567 (2018).
- (141) Couture, H.D., Williams, L.A., Geradts, J. et al. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. *npj Breast Cancer* 4, 30 (2018).

- (142) Jaber, M.I., Song, B., Taylor, C. et al. A deep learning image-based intrinsic molecular subtype classifier of breast tumors reveals tumor heterogeneity that may affect survival. *Breast Cancer Res* 22, 12 (2020).
- (143) Khoury T, Huang X, Chen X, Wang D, Liu S, Opyrchal M. Comprehensive Histologic Scoring to Maximize the Predictability of Pathology-generated Equation of Breast Cancer Oncotype DX Recurrence Score. *Appl Immunohistochem Mol Morphol* 2016;24:703–711. [PubMed: 26808126]
- (144) Thaker NG, Hoffman KE, Stauder MC, Shaitelman SF, Strom EA, Tereffe W, Smith BD, Perkins GH, Huo L, Munsell MF, et al. The 21-gene recurrence score complements IBTR! Estimates in early-stage, hormone receptor-positive, HER2-normal, lymph node-negative breast cancer. *Springerplus* 2015;4:36. [PubMed: 25674496]
- (145) Acs G, Kiluk J, Loftus L, Laronga C. Comparison of Oncotype DX and Mammostrat risk estimations and correlations with histologic tumor features in low-grade, estrogen receptor-positive invasive breast carcinomas. *Mod Pathol* 2013;26:1451–1460. [PubMed: 23743933]
- (146) A. Basavanhally, Jun Xu, A. Madabhushi and S. Ganesan, "Computer-aided prognosis of ER+ breast cancer histopathology and correlating survival outcome with Oncotype DX assay," 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009, pp. 851-854.
- (147) Romo-Bucheli D, Janowczyk A, Gilmore H, Romero E, Madabhushi A. Automated Tubule Nuclei Quantification and Correlation with Oncotype DX risk categories in ER+ Breast Cancer Whole Slide Images. *Sci Rep.* 2016 Sep 7;6:32706.
- (148) Romo-Bucheli D, Janowczyk A, Gilmore H, Romero E, Madabhushi A. A deep learning based strategy for identifying and associating mitotic activity with gene expression derived risk categories in estrogen receptor positive breast cancers. *Cytometry A.* 2017 Jun;91(6):566-573.
- (149) Basavanhally, A. et al. Multi-field-of-view strategy for image-based outcome prediction of multi-parametric estrogen receptor-positive breast cancer histopathology: comparison to oncotype dx. *Journal of pathology informatics* 2, (2011).
- (150) Graham S, Vu QD, Raza SEA, Azam A, Tsang YW, Kwak JT, Rajpoot N. HoverNet: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med Image Anal.* 2019 Dec;58:101563.
- (151) Wang Y, Acs B, Robertson S, Liu B, Solorzano L, Wählby C, Hartman J, Rantalainen M. Improved breast cancer histological grading using deep learning. *Ann Oncol.* 2022 Jan;33(1):89-98.
- (152) Prat A, Cheang MC, Martín M, Parker JS, Carrasco E, Caballero R, Tyldesley S, Gelmon K, Bernard PS, Nielsen TO, Perou CM. Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer. *J Clin Oncol.* 2013 Jan 10;31(2):203-9.

Annexes

Annexe 1 : Exemple de compte rendu d'analyse moléculaire Prosigna® PAM50



INSTITUT UNIVERSITAIRE
DU CANCER DE TOULOUSE
Oncopole

LABORATOIRE DE BIOLOGIE MEDICALE ONCOLOGIE
N° enregistrement : 31.045
N° Agrément : 5054

BIOLOGISTE MEDICAL RESPONSABLE
Pr Gilles FAURE
Tel : 05 31 15 52 00

BIOLOGISTES MEDICAUX
Dr Virginie ARMAND-LABIT
Tel : 05 31 15 52 12
Pr Stéphanie CHATELUT
Tel : 05 31 15 52 30
Dr Samir DAHMEENE
Tel : 05 31 15 52 18
Dr Laure HESLER
Tel : 05 31 15 52 14
Dr Laurence MALARD
Tel : 05 31 15 52 15
Dr Fabienne THOMAS
Agréée en Pharmacopéutiques
Tel : 05 31 15 52 19

PRATICIENS AGRES EN
ONCOGENETIQUE MOLECULAIRE
Pr Gilles FAURE
Tel : 05 31 15 52 00
Dr Christine TOULAS
Tel : 05 31 15 52 20

SECRETARIAT
Tel : 05 31 15 52 01
Fax : 05 31 15 52 38
act@iuct-oncopole.fr

Date : 17/12/2021

À la demande de la RCP Sénologie du

Patient

Nom d'usage :
Nom de naissance :
Prénom :
Date de naissance :
N° IPP :

Prélèvement

Nature de l'échantillon reçu : ARN
Numéro de l'examen anapath. :
Date de l'extraction :
Date de réception dans le laboratoire :
Numéro laboratoire :
Renseignements Anato-mo-pathologiques :
Taille de la tumeur :
Nombre de ganglions positifs :

Test Prosigna™ Breast Cancer Prognostic Gene Signature Assay

Mesure de l'expression d'une signature moléculaire de 50 gènes (PAM50) pour identifier le sous type et évaluer un score de risque de récurrence à distance.

RESULTATS

Risque de récurrence (score ROR, échelle 0-100) : 55

Catégorie de risque : Risque Intermédiaire

Sous type intrinsèque : Luminal B

Probabilité du risque de récurrence à 10 ans : 12%

Patient
Tumor Size:
Lymph Nodes:

Specimen
ID #:
Date Reported: December

Run Set ID: 20211216-Fb
Comments:



ID #: Tumor Size: Lymph Nodes:
Assay Description: The Prosigna[®] breast cancer gene signature assay measures the expression of 50 different genes to identify subtype and report a Risk of Recurrence Score (ROR), which is used to assign the patient to a predefined risk group. These results are derived from a proprietary algorithm based on the PAM50 gene signature, intrinsic subtype, and clinical variables including tumor size and nodal status.

Risk of Recurrence*:

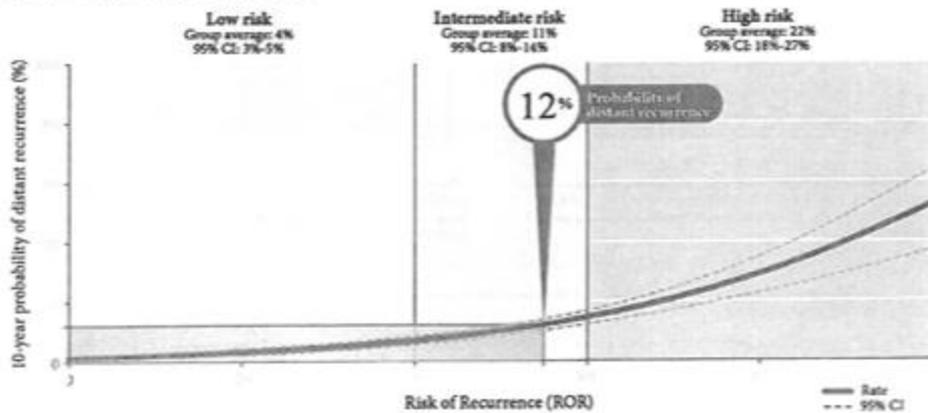


* The ROR ranges from 0 through 100 and correlates with the probability of distant recurrence (DR) in the tested patient population. The risk classification is provided to guide the interpretation of the ROR using cutoffs related to clinical outcome.

Probability of Distant Recurrence:

In the clinical validation studies, patients who were node-negative, luminal B subtype, with an ROR score of 55 were in the intermediate-risk group. This group averaged an 11% probability of distant recurrence at 10 years.

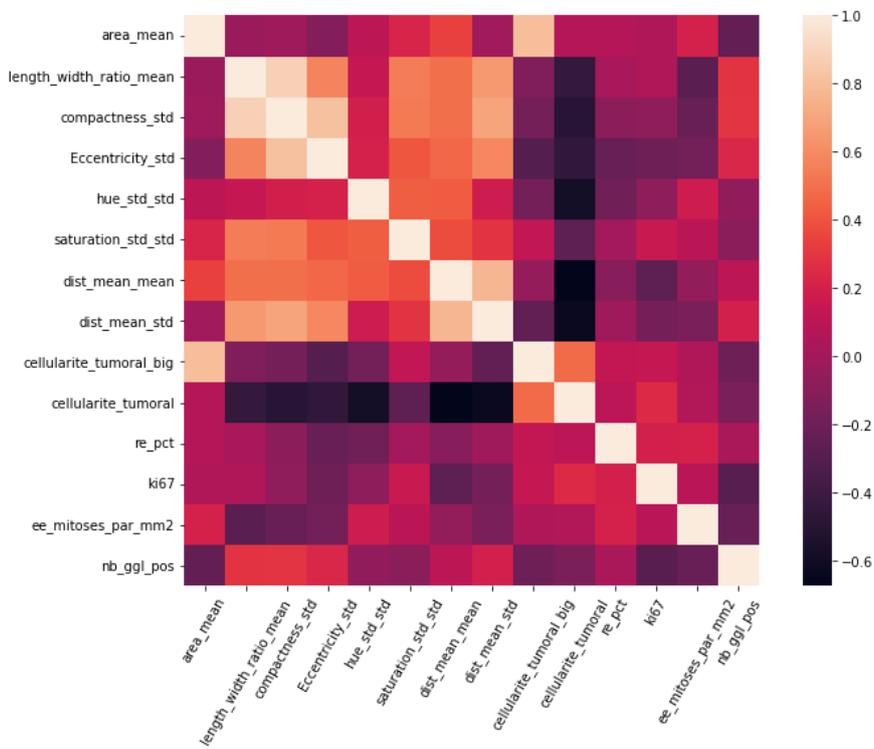
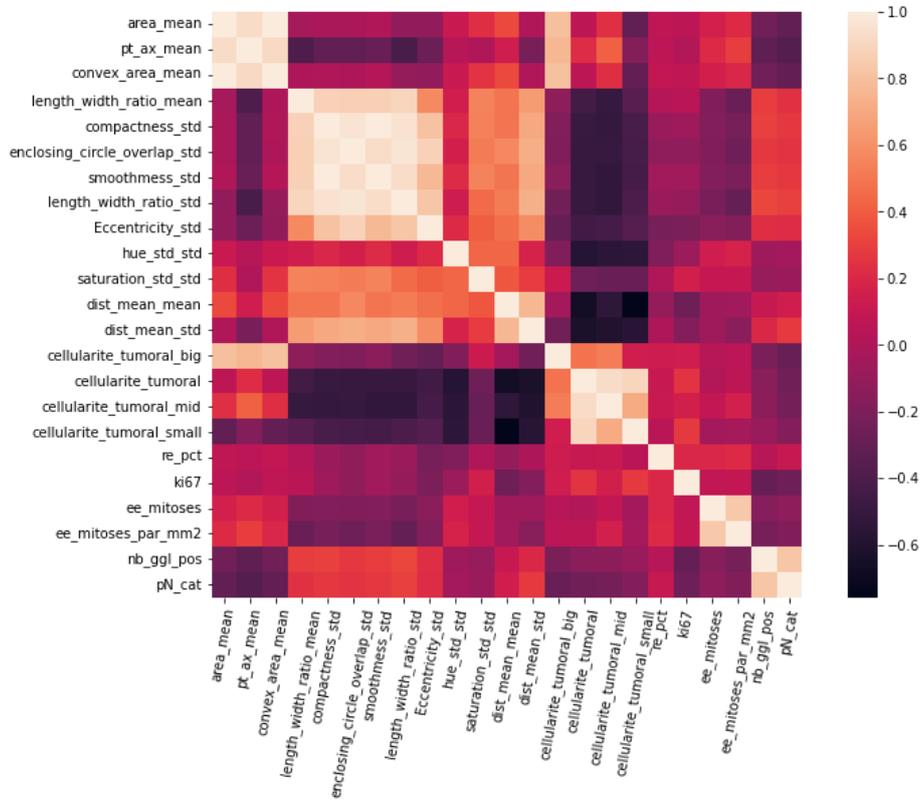
The Prosigna[®] algorithm has been validated by 2 randomized clinical trials including more than 2400 patients with varying rates of distant recurrence. An analysis of these 2 clinical validation studies shows that the probability of distant recurrence for the intermediate-risk population is 11%, while the high-risk population has a significantly greater probability of distant recurrence.†



For more information, visit PROSIGNA.com or e-mail info@prosigna.com

†Data apply to patients being treated with hormone therapy for 5 years as in the tested patient population. See Package insert for further information on therapy regimens and tested patient population. It is unknown whether these findings can be extended to other patient populations or treatment schedules. ©2014 2021 NanoString Technologies.

Annexe 2 : matrices de corrélation pour les différentes caractéristiques utilisées dans l'étude (en haut : caractéristiques non triées ; en bas : caractéristiques triées pour les modèles de ML)



APPORT DE LA SEGMENTATION PREALABLE DES OBJETS BIOLOGIQUES DANS DES IMAGES MICROSCOPIQUES DE CANCERS DU SEIN, POUR LA PREDICTION DE LEUR SOUS-TYPE MOLECULAIRE

RESUME :

La prise en charge thérapeutique des cancers du sein infiltrants localisés comprend un traitement locorégional associé ou non à un traitement systémique. Pour certaines tumeurs, le choix de recourir à une chimiothérapie adjuvante est difficile et une analyse moléculaire peut être réalisée afin d'étayer ce choix. L'objectif de notre étude était d'évaluer et de comparer les performances de prédiction du sous-type moléculaire donné par la signature moléculaire Prosigna® PAM50, à partir d'images microscopiques de cancers du sein. Deux méthodes étaient comparées : une se basant sur du deep learning et l'autre sur une analyse par machine learning de features morphologiques extraites au préalable. Nous avons pu mettre en évidence des performances de prédiction sensiblement équivalentes pour chaque méthode, avec des accuracy de l'ordre de 70 %. La combinaison de ces méthodes avec les CCPT semblent améliorer légèrement la prédiction, avec des améliorations sensiblement identiques pour ces 2 méthodes. L'analyse de l'interprétabilité des résultats montrent que certains critères ou features semblent plus importants que d'autres pour la prédiction du sous-type intrinsèque. En effet, le nombre de mitoses par mm² et le Ki67 semblent surpasser les autres critères. Le deep learning semble porter un intérêt aux zones riches en lymphocytes pour le sous-type luminal B. Et l'analyse des features nucléaires semblent montrer que les tumeurs lumineales B correspondent à des tumeurs avec une cellularité augmentée, des noyaux globalement plus arrondis mais aux contours plus irréguliers, et un pléomorphisme plus faible ; contrairement aux tumeurs lumineales A.

DISCIPLINE ADMINISTRATIVE : Médecine spécialisée clinique – Anatomie et cytologie pathologiques

MOTS-CLÉS : cancer du sein – intelligence artificielle – machine learning – deep learning – sous-type intrinsèque – analyse d'image – features – interprétabilité – pronostic – signature moléculaire - prédiction

INTITULÉ ET ADRESSE DE L'UFR OU DU LABORATOIRE :

Université Toulouse III-Paul Sabatier
Faculté de médecine Toulouse-Purpan,
37 Allées Jules Guesde 31000 Toulouse

Directeur de thèse : Camille FRANCHET
