



HAL
open science

Caractérisation du sous type histologique par analyse radiomique des IRM cérébrales de métastases

Céline Corcy

► **To cite this version:**

Céline Corcy. Caractérisation du sous type histologique par analyse radiomique des IRM cérébrales de métastases. Médecine humaine et pathologie. 2020. dumas-02983204

HAL Id: dumas-02983204

<https://dumas.ccsd.cnrs.fr/dumas-02983204v1>

Submitted on 29 Oct 2020

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

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



UNIVERSITÉ de PICARDIE JULES VERNE
Faculté de Médecine

Année 2020 N°2020-104

**THÈSE
POUR LE DIPLÔME D'ÉTAT
DE
DOCTEUR EN MÉDECINE,**

SPECIALITE IMAGERIE MEDICALE

Caractérisation du sous type histologique par analyse radiomique des IRM cérébrales de métastases

Présentée et soutenue publiquement
le Vendredi 18 Septembre 2020
Par

Céline CORCY

Née le 28/07/91, à Reims

Dirigée par le Docteur Myriam EDJLALI-GOUJON
et le Professeur Jean-Marc CONSTANS

Jury :

M. Le Professeur Jean-Marc CONSTANS, PU-PH en neuroradiologie, Président du Jury
M. Le Professeur Henri SEVESTRE, PU-PH en anatomopathologie
Mme Le Professeur Catherine LOK, PU-PH en dermatologie
M. Le Professeur Michel LEFRANC, PU-PH en neurochirurgie
Mme Le Docteur Charlotte ROBERT, MCU en physique médicale
Mme Le Docteur Myriam EDJLALI-GOUJON, PH en neuroradiologie.

A Maman, Papa, Elise, Romain et Marie-Laure qui m'ont toujours soutenue.

TABLE DES MATIERES

<u>LISTE DES FIGURES, TABLEAUX ET EQUATIONS</u>	4
<u>REMERCIEMENTS</u>	7
<u>LISTE DES ABREVIATIONS</u>	11
<u>INTRODUCTION</u>	16
<u>I. Contexte général</u>	16
<u>II. Que sait-on des voies de dissémination et de la localisation des lésions secondaires cérébrales ?</u>	18
<u>III. Démarche diagnostique devant la découverte d'une possible métastase cérébrale ?</u>	21
1. <u>Bilan IRM devant une possible localisation cérébrale.</u>	21
2. <u>Bilan à la recherche du primitif</u>	26
3. <u>Diagnostic histologique</u>	27
<u>IV. La radiomique, une nouvelle voie d'analyse d'image ?</u>	31
<u>V. Objectifs de l'étude</u>	33
<u>MATERIEL ET METHODES</u>	34
<u>I. Différentes étapes d'une analyse radiomique</u>	34
<u>II. Acquisition des images</u>	34
1. <u>Population</u>	34
2. <u>Données IRM</u>	36
<u>III. Segmentation du volume d'intérêt</u>	36
<u>IV. Pré-traitement des images</u>	38
1. <u>Correction de l'inhomogénéité du champ magnétique</u>	38
2. <u>Skull-stripping</u>	39
3. <u>Z-Score</u>	39
4. <u>Ré-échantillonnage spatial</u>	39
<u>V. Extraction des index radiomiques</u>	39
1. <u>Index de forme</u>	40
2. <u>Index d'intensité (ou index agnostiques de 1er ordre)</u>	40
3. <u>Index de texture (ou index agnostiques de 2ème ordre)</u>	42
<u>VI. Information spatiale</u>	46
<u>VII. Analyse statistique</u>	46
1. <u>Généralités</u>	46
2. <u>Analyse statistique</u>	48
3. <u>Ré-échantillonnage des populations</u>	48
4. <u>Principales mesures de performance</u>	49

5.	<u>Modèles testés</u>	51
6.	<u>Stratégie One versus One</u>	52
7.	<u>Stratégie One versus Others</u>	53
	<u>RESULTATS</u>	54
I.	<u>Population</u>	54
II.	<u>Résultats de la stratégie One versus One</u>	55
1.	<u>Métastases d'origine mammaire versus pulmonaire</u>	56
2.	<u>Métastases d'origine mammaire versus colique</u>	57
3.	<u>Métastases d'origine mammaire versus mélanome</u>	57
4.	<u>Métastase d'origine pulmonaire versus colique</u>	57
5.	<u>Métastase d'origine mélanome versus colique</u>	58
III.	<u>Résultats de la stratégie One versus Others</u>	59
	<u>DISCUSSION</u>	61
I.	<u>Rappel des principaux résultats et implication clinique</u>	61
II.	<u>Comparaison à la littérature</u>	62
III.	<u>Forces de l'étude</u>	68
IV.	<u>Limites de l'étude</u>	68
V.	<u>Perspectives</u>	70
	<u>CONCLUSION</u>	71
	<u>BIBLIOGRAPHIE</u>	72
	<u>ANNEXES</u>	77
	<u>ANNEXE 1 : Modèles définis comme optimaux pour répondre à la stratégie « One vs. One »</u>	77
	<u>ANNEXE 2 : Résultats portant sur la <i>precision</i> de la classification One versus One</u>	77
	<u>ANNEXE 3 : Résultats portant sur le <i>recall</i> de la classification One versus One</u>	78
	<u>ANNEXE 4 : Résultats portant sur l'<i>accuracy</i> de la classification One versus One</u>	78
	<u>ANNEXE 5 : Résultats portant sur l'apprentissage des scores roc-AUC et F1 de la classification One versus One</u>	79
	<u>ANNEXE 6 : Modèles définis comme optimaux pour répondre à la stratégie « One vs. Others »</u>	79
	<u>ANNEXE 7 : Résultats portant sur la <i>precision</i> de la classification One versus Others</u>	80
	<u>ANNEXE 8 : Résultats portant sur le <i>recall</i> de la classification One versus Others</u>	80
	<u>ANNEXE 9 : Résultats portant sur l'<i>accuracy</i> de la classification One versus Others</u>	81
	<u>ANNEXE 10 : Résultats portant sur l'apprentissage des scores roc-AUC et F1 de la classification One versus Others</u>	81
	<u>ANNEXE 11 : Caractéristiques Morphologiques des Métastases Cérébrales à l'IRM [86]</u>	81
	<u>RESUME et MOTS CLES</u>	83

LISTE DES FIGURES, TABLEAUX ET EQUATIONS

<u>Figure 1:Distribution des types de tumeurs primitives inconnues au moment de la découverte de la métastase [14]</u>	17
<u>Figure 2 : Différentes voies de dissémination des métastases cérébrales, responsables de différentes localisations métastatiques distinguées en IRM : atteinte leptoméningée (A), pachyméningée (B), miliaire métastatique sur extension hémotogène (C), extension périneurale par un foramen de la base : infiltration de la fosse ptérygopalatine (D)</u>	18
<u>Figure 3:Barrière hémato-encéphalique [14]</u>	19
<u>Figure 4:Dissémination des cellules tumorales métastatiques [4]</u>	20
<u>Figure 5:Séquence volumique 3DT1 EG après injection de gadolinium retrouvant un processus expansif occupant l'espace intra parenchymateux, frontal droit, correspondant à une métastase cérébrale. La périphérie se rehausse et le centre parait de signal liquidien nécrotique.</u>	23
<u>Figure 6:Métastase de mélanome pariétal gauche entourée d'un œdème périphérique de forte intensité FLAIR (pointe de flèche). Notez que l'œdème n'implique pas le cortex environnant [23]</u>	24
<u>Figure 7:Bilan IRM d'une lésion métastatique en séquences morphologiques.</u>	25
<u>Figure 8:Volume T1 EG après injection de gadolinium retrouvant des lésions métastatiques secondaires d'un colon (A), mélanome (B), poumon (C), sein (D).</u>	26
<u>Figure 9 : Foyer de mélanome métastatique impliquant le lobe frontal au niveau de la commissure antérieure avec hémorragie associée et œdème environnant [20]</u>	28
<u>Figure 10 : Mélanome métastatique (a) A faible grossissement, la démarcation est relativement nette, à l'exception d'un étalement périphérique limité le long des espaces périvasculaires. La substance blanche adjacente apparaît pâle et oedémateuse. (b) A plus fort grossissement : un grossissement plus élevé révèle une néoplasie épithélioïde et filiforme à pigmentation variable, avec des nucléoles proéminents, une activité mitotique et des mélanophages dispersés, typiques du mélanome [20]</u>	29

<u>Figure 11 :Carcinome pulmonaire métastatique à petites cellules.</u>	30
<u>Figure 12: Organigramme de processus de la radiomique et de l'utilisation de la radiomique dans l'aide à la décision.</u>	32
<u>Figure 13 : Critères d'inclusion et d'exclusion des patients.</u>	35
<u>Figure 14 : Interface du logiciel Mango, avec lequel la segmentation a été réalisée.</u>	37
<u>Figure 15: à gauche : image avec bias field et, à droite, image corrigée)[69]</u>	38
<u>Figure 16: Description d'un histogramme d'intensité</u>	41
<u>Figure 17: Schéma des 13 directions nécessaires pour connecter un voxel avec ses 26 voisins</u>	42
<u>Figure 18:Matrice de Co-occurrence des niveaux de gris (GLCM)[76]</u>	42
<u>Figure 19 : Matrices de texture usuelles et index de texture dérivés [78]</u>	45
<u>Figure 20 : Illustration de la méthodologie de la validation croisée</u>	47
<u>Tableau 1 : Index de Formes</u>	40
<u>Tableau 2 : Index de 1er ordre</u>	41
<u>Tableau 3:Index de texture</u>	44
<u>Tableau 4 : Classifieurs testés pour la construction des modèles.</u>	52
<u>Tableau 5:Performances de Classification pour la stratégie One vs One</u>	56
<u>Tableau 6: Performances de Classification pour la stratégie One vs Others</u>	59
<u>Tableau 7 : Tableau comparant la méthodologie et les résultats de notre étude aux deux études récentes similaires à la nôtre de la littérature</u>	62
<u>Équation 1:Formule du Z-Score</u>	39
<u>Équation 2: Formule de la precision</u>	50
<u>Équation 3:Formule du recall</u>	50
<u>Équation 4:Formule de l'accuracy</u>	50
<u>Équation 5:Formule du score F1</u>	50

REMERCIEMENTS

A Monsieur le Professeur Jean-Marc Constans:

Responsable du centre d'activité de Neuroradiologie
Coresponsable de l'équipe Universitaire CHIMERE
Coordonnateur du DES d'Imagerie Médicale
Responsable Médical et de la recherche clinique sur l'IRM de recherche du GIE Faire Face
Professeur des Universités-Praticien Hospitalier
(Radiologie et Imagerie médicale)

Pôle Imagerie

Merci de me faire l'honneur de présider le jury de ma thèse et de m'accompagner pour cette thèse. Vous m'avez fait aimer la neuroradiologie et donner l'envie de poursuivre dans ce domaine par la suite. De plus, vous m'avez fait découvrir la recherche et donné la motivation de continuer dans cette voie.

A Monsieur le Professeur Henri Sevestre :

Professeur d'Anatomie et de Cytologie Pathologiques à l'UFR de Médecine d'Amiens
Chef du Service d'Anatomie et Cytologie Pathologiques du CHU d'Amiens Picardie
Adjoint au chef de l'Oncopôle

Merci d'avoir accepté de participer au jury de ma thèse, de m'avoir si gentiment aidée à récolter ma base de données pour le CHU d'Amiens et d'avoir pris de votre temps pour m'accueillir au sein de votre service.

A Madame le Professeur Catherine Lok:

Professeur des Universités-Praticien Hospitalier
(Dermatologie-Vénérologie)
Membre du Conseil de l'UFR de Médecine
Chef du Service de Dermatologie

Je vous remercie chaleureusement d'avoir accepté de participer au jury de ma thèse. Merci de m'apporter votre point de vue de dermatologue dans ce travail sur les métastases cérébrales. Veuillez trouver ici l'expression de ma sincère reconnaissance.

A Monsieur le Professeur Michel Lefranc :

Professeur des Universités - Praticien Hospitalier
(Neurochirurgie)
Directeur du GRECO (groupement de recherche et d'études en chirurgie robotisée) UPJV

Je vous remercie profondément d'avoir accepté de participer au jury de ma thèse et de m'apporter votre point de vue de neurochirurgien dans ce travail sur les métastases cérébrales. Veuillez trouver ici l'expression de ma sincère reconnaissance.

A Madame le Docteur Myriam Edjlali-Goujon :

Docteur en Médecine - Praticien Hospitalier
(Neuroradiologie)

Je te remercie énormément pour toute cette direction de travail de thèse.

Merci pour cette idée de métastases cérébrales et de radiomique.

Grâce à toi, il s'agira de ma première publication et de mon entrée dans le monde de la recherche.

Merci pour ton investissement sans faille et ta profonde gentillesse à mon égard.

Merci pour tout ton apprentissage en neuroradiologie et en recherche. Je suis vraiment heureuse d'avoir pu croiser ta route et j'espère continuer après ce travail.

A Madame le Docteur Charlotte Robert :

Maître de conférences – Physicienne médicale

Merci de me faire l'honneur de participer au jury de ma thèse.

Je vous remercie énormément pour votre rigoureuse collaboration et votre compagnonnage très réactif pour cette étude. Merci d'avoir pris le temps de m'expliquer et par conséquent de comprendre tous les concepts de la radiomique. Merci pour votre grande disponibilité.

Et un immense Merci pour la correction très détaillée du manuscrit avec toutes vos explications !!

A Monsieur Alexandre Carré :

Merci beaucoup pour ton investissement dans cette étude et pour avoir répondu à toutes mes questions si nombreuses !!! Merci pour toutes tes explications et ton investissement dans la correction de cette thèse. Je te souhaite une très bonne continuation pour la suite.

A Monsieur le Docteur Cédric Renard :

Merci pour toute l'aide que tu m'as apportée tout au long de mon internat, de ton partage de connaissances illimitées, du temps (pas mal, il faut le dire!) que tu m'as accordé et de ce plaisir de faire aimer la radiologie.

A Madame Alix de Causans :

Je crois que je ne te dirai jamais assez merci pour toute ton aide pour cette thèse. Tu es vraiment une superbe personne sur qui on peut compter. J'ai hâte de te revoir

A Romain :

Merci de m'avoir supportée pendant tout ce long parcours, d'avoir toujours été là, si attentionné, de m'avoir aidée à résoudre tant de problèmes (surtout informatiques !), et de toujours m'aimer telle que je suis et d'accepter de passer toute ta vie avec moi ;).

A Elise :

Merci d'avoir toujours été là et de m'avoir tellement apporté, plein de joie et de bonheur, un soutien indescriptible, à toute épreuve. Tu es la meilleure sœur qu'on puisse espérer avoir.

Je suis tellement fière de toi et très heureuse de t'avoir à mes côtés.

A Papa et Maman:

Que dire ! Merci pour tout ! Je vous dois tellement. Merci de m'avoir donné le goût du travail, de la persévérance, de m'encourager à chaque épreuve, sans renoncer, et

de toujours rester à mes petits soins même à 29 ans! Merci de m'aimer comme je suis avec mes défauts et d'être toujours aussi généreux envers moi. Je ne l'oublierai jamais.

A Marie-Laure :

Merci pour ta bonne humeur, ta gentillesse, ta créativité et ton aide à tout niveau, que ce soit au niveau scolaire, artistique ou personnel!!! Merci pour ce que tu es, tout simplement !!! Je ne serais pas arrivée là sans ton soutien. Merci encore d'être une super tata extraordinaire !!!!

Et un immense merci pour ta si précieuse relecture !!!!

A toute ma famille, Grégoire, Jean-Arthur, Boris, Louis, Romain, Philippe, Erick, Sabine, Tonton, Annabel et à mes grands parents : Merci pour tous ces bons moments et votre soutien !Merci de m'avoir guidée dans cette voie.

A Sophie, Jacques, Estelle, Alexis et Céline :

Merci d'avoir toujours été là, de votre compréhension et de votre aide dans tout mon parcours.

A Amandine :

Merci pour ta gentillesse à mon égard, de ton génialissime compagnonnage durant tout cet internat et notamment durant le DU d'ORL. Que de fous rires et de bons moments!! Reste comme tu es et ne change rien !!! (« on n'est pas là pour être sympa » ;)). Mon seul regret est que tu ne sois pas ma co chef ! Ca aurait tellement envoyé du rêve !

A Julie :

Merci pour toute ton aide, de ta douce attention envers moi, et de ton extrême générosité. Tu es une superbe personne, en plus d'être une super maman ! Tu arrives à gérer tellement à la fois, tu es un vrai exemple pour moi. Ta gentillesse va me manquer !

A Charlotte :

Merci pour ta bonne humeur et de ton soutien pendant toutes ces années et notamment durant le stage d'imagerie digestive où on se retrouvait toutes les deux, un peu en solitaire ! Tu étais toujours là dans l'adversité !!!! Tu es une super amie sur laquelle on peut toujours compter, merci d'être ce que tu es !! Tu es une fille en or !

A Alexandre :

Que dire, à part que tu es un vrai ami fidèle, prévenant, attentionné et tellement gentils! Je suis très contente d'avoir croisé ta route et j'espère qu'on se reverra sur Paris ou ailleurs !!!!

A Charles :

Merci beaucoup pour tout, pour ta sagesse, tes attentions et ta dévotion pour la radio ! Reste comme tu es !!

A Sarah :

Merci pour ton aide, ton soutien et de ta gentillesse envers moi. J'espère vraiment qu'on se reverra par la suite mais je n'en doute pas !

A Sahra et Redoine :

Vous êtes exceptionnels, les amis, merci pour votre soutien et votre bonne humeur à tout niveau ! Vous êtes tellement généreux et attentionnés. J'ai beaucoup de chance de vous avoir.

Restez comme vous êtes !

A Aurélie :

Merci de m'avoir aidé et conseillé. Ta gentillesse et ta générosité font de toi une superbe amie. Merci.

A Clémence, Hélène et Pierre :

Merci d'être rester à mes côtés pendant toutes ses années et de m'avoir soutenue tout ce temps.

Merci d'être ce que vous êtes, des amis en or !!! En plus d'être de superbes témoins, vous êtes des personnes sur lesquelles j'ai toujours pu compter. Je sais que nous resterons toujours liés.

Aux Pitous : à Rosy, Nini, Céline, Romain, Statik, Julien, Clémaille, Kevin, Aurélien et Laurène :

Merci de m'avoir fait partager de très bons moments, de délires, de fêtes et de bonheur !

Même les kilomètres n'ont pu nous séparer, alors rien ne peut le faire ! Vive l'amitié

A Andy : Pour ses performances informatiques en pleine garde ! MERCI !!!

A tous mes co internes :

Un immense merci à Jessy, Jean, Loïc, Emilie, Constance, Emna, Fabrice, Anas, Samir, Jamel, Olivier, Mickael, David, Aurélien, Kevin, Olivier, Quentin, Xavier, Geoffrey, Wesley, Simon, Margaux, Victor, Amélie, Antoine, Naty, JB, JB, Charles, Andy, David, Jean, JF, Mickael, Wilfried, Déesse :

Merci pour ce super internat, savant mélange de rigolades, d'entraide et de bonne humeur ! Merci pour votre bienveillante gentillesse !!!

A mes (jeunes) chefs :

Merci à Jerem, Nico, Raph, Antoine, Kevin, Riyad, Florine, Floriane, Elodie, pour toute votre écoute et votre pédagogie.

A l'équipe de sainte Anne : Joseph, Ghazi, Catherine, Grégoire, Laurence, Laura, Denis, Olivier, Christine, Charles, Eric, Sylvie, Djallel, Jean-François, Jean-Phi, Laurence, Elisa, Around, Jérémy, Kathia, Meline, Christopher, Julien, Maliesse, Mehdi, Nico et Fifi :

Merci pour votre chaleureux accueil et votre joie de vivre. Merci de m'avoir fait passé un interCHU de fou ! Merci pour votre bienveillance et votre gentillesse.

Pour Wagih :

Merci pour ta gentillesse, ton aide, ton humour et ta pédagogie !!

A toute l'équipe d'Amiens : Justine, Taite, Corinne, Francis, Aurore, Manon, Amélie, Mathilde, Julie, Lisa, Valérie, Hélène, Hélène, Gigi, Mathieu, Nico, Arthur, Aubin, Charlotte, Claire, Carole, JP, Camille, Kevin, Mélanie, Perrine, Solène, Lolo, Florie, Elie, Claudia, Clara, Manon, Hélène, Leïla, Pauline, Stéphanie, Quentin, Bene, Elsa, Céline, Marion, Elodie, Sabrina, Elisa, Dodo, Garance, Quentin, Laurent, Clément, Anne-So, Lorie, Nath, Parfait, Serge, Brice, David, Pierrot, Mickael, Mathilde, Bruno, Catherine, Julie, Marie-Anne, Thomas, Brigitte, Mr Yzet : Merci pour tout ! Vous êtes les meilleurs ! Que de merveilleux souvenirs passés à vos côtés !

LISTE DES ABREVIATIONS

2D : Bidimensionnel

3D : Tridimensionnel

3D T1 : Séquence d'IRM tridimensionnelle en pondération T1

ADC : Apparent Diffusion Coefficient : coefficient apparent de diffusion

ADK : Adénocarcinome

ANTs : Advanced Normalization Tools

ARN : Acide RiboNucléique : Acide nucléique essentiel dans le transport du message génétique et la synthèse des protéines

AUC : Area Under Curve : aire sous la courbe

AFP : Alpha-FoetoProtéine : marqueur tumoral dans le carcinome hépatocellulaire, le tératome de l'ovaire et le cancer du testicule

ACE : Antigène Carcino Embryonnaire : retrouvé à la surface des cellules de l'intestin grêle, du colon, du rectum, du pancréas, du poumon et du rein ; produit davantage par les cancers colorectaux

ALK : Anaplastic Lymphoma Kinase : translocation retrouvée dans le cancer du poumon

BHE : Barrière-Hémato-Encéphalique

BTB : Barrière Tumoro-Encéphalique

C : Carcinome non spécifique

CA 125 : Carbohydate Antigen 125 : Marqueur tumoral de cancers ou de maladies digestifs ou génitaux

CA 19-9 : Antigène Carbohydate 19-9 : Marqueur tumoral de cancer pancréatique

CA 15-3 : Antigène Carbohydate 15-3 : Marqueur tumoral de cancer du sein

CBPC : Cancer Bronchique à Petites Cellules

CBNPC : Cancer Bronchique Non à Petites Cellules

CDX2 : protéine : Marqueur de la différenciation digestive et spécialement colorectale

Cho : Choline : Marqueur des membranes, de la myéline, ou de l'inflammation

CK7: CytoKératine 7 : Expression dans les tumeurs du poumon, de l'ovaire, de l'endomètre et du sein

CK20 : CytoKératine 20 : Expression dans les tumeurs gastro-intestinaux et urothéliaux

COX2 : CycloOXYgénase 2 : Enzyme permettant la formation des prostanoides (dont les prostaglandines)

Cr : Créatine : Marqueur du pool énergétique des cellules cérébrales

CUP : Tumeur de Primitif Inconnu

DICOM: Digital imaging and communications in medicine

DOF : Degré de Liberté

DWI : Diffusion Weighted Images : Séquences pondérées en diffusion

EGFR : Epidermal Growth Factor Receptor : Mutation associée à plusieurs cancers dont celui du poumon

ER : Récepteur oestrogène

F1 score : Function measure : mesure de la précision d'un test

FLAIR : T2-Fluid Attenuation Inversion Recovery : Séquence IRM

FDG-PET: Tomographie par émission de positrons sur corps entier 2-désoxy-2-[18F]fluoro-D-glucose

FET : O-(2-[18 F]fluoroethyl)-L-tyrosine : Radiopharmaceutique utilisé en TEP

FN : Faux Négatif

FOV : Field Of View : Champ de vue

FP : Faux Positif

GCDFP-15 : Gross Cystic Disease Fluid Protein 15 : Protéine retrouvée en grand nombre dans les cancers du sein

GLCM : Gray Level Co-occurrence Matrix : Matrice de co-occurrence

GLRLM : Grey Level Run-Length Matrix : Matrice des longueurs de niveaux de gris

GLSZM : Grey Level Size-Zone Matrix : Matrice des tailles de zones homogènes

GLDM : Gray Level Dependence Matrix : Matrice de dépendance des niveaux de gris

hCG : Hormone Chorionique Gonadotrope Humaine : Marqueur tumoral de tératome, cancer testiculaire, choriocarcinome, môle hydatiforme ou de certains cancers neuroendocrines

HD-BET: High-Dimensional - Brain Extraction Tool

HER2 : protéine : Surexpression du Récepteur HER2 retrouvée dans 15 à 20 % de cancers du sein

IBSI : Image Biomarker Standardisation Initiative : Initiative de Normalisation des Biomarqueurs d'Imagerie

ICBM : International Consortium of Brain Mapping

IHC : Immunohistochimie

IsoSVM : algorithme pour distinguer les isoformes et paralogues sur le niveau de protéine

HMB-45 : Mouse Anti-Melanosome Antibody : Anticorps pour l'évaluation qualitative du mélanome malin

IT : Temps d'Inversion

Ki-67 : Marqueur de Prolifération cellulaire tumorale (étude réalisée à l'origine dans la ville de Kiel (Ki) sur un clone de cellules (67))

KPS : Karnofsky Performance Status

KRAS : V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog : Mutation jouant un rôle clé dans la genèse de plusieurs cancers

LCA : Leucocyte Common Antigen : Exprimé dans presque toutes les cellules hématolymphoïdes

LCS : Liquide Cérébro-Spinal

Melan-A : MART-1: Antigène de différenciation des mélanocytes

MET : [11C]-methyl-L-methionine - radiopharmaceutique utilisé en TEP

miARN: micro ARN : Catégorie de petits acides ribonucléiques, simples brins, non codants, propres aux cellules eucaryotes

MITF : Microphthalmia-Associated Transcriptor Factor: Impliqué dans le développement et la survie des mélanocytes

NGDLM : Neighborhood Grey Level Difference Matrix : Matrice de différence des niveaux de gris entre voxels voisins

MLM : Mélanome

MNI : Montreal Neurological Institute

N4ITK : Improved N3 Bias Correction

NAA : N-acétylaspartate : Diminution traduisant une souffrance ou un remplacement neuroaxonal par du tissu tumoral ou de la gliose

Napsine A : Marqueur d'adénocarcinome pulmonaire

NSE : Neuron-specific enolase

OCT 4 : Octamer-binding transcription factor 4

OMS : Organisation Mondiale de la Santé

OR : Odds Ratio : rapport des cotes

PACS : Picture Archiving and Communication System : système d'archivage des images médicales

PgR : Progesterone Receptor : Expression dans les tissus des organes reproducteurs (ovaires++), dans l'os, la prostate, le tissu adipeux et la membrane des vaisseaux

PLAP : Placental Alkaline Phosphatase : Exprimée dans les séminomes et moins fréquemment dans les tumeurs germinales

PSA : Antigène spécifique de la prostate : Protéine fabriquée exclusivement par la prostate

ROI : Region Of Interest : Région d'intérêt

RM : Résonance magnétique

RT : Radiothérapie

RS : Radiothérapie Stéréotaxique

TKI : Inhibiteur de Tyrosine Kinase : Utilisé dans le traitement du cancer pulmonaire

TN : True Negative

TP : True Positive

S : Sarcome

S-100 : Protéine, présente dans les cellules dérivées des crêtes neurales (cellules de Schwann et mélanocytes), chondrocytes, adipocytes, macrophages entre autres et donc un marqueur de tumeurs de mélanome, schwannome et neurofibrome entre autres

SCC : Carcinome Epidermoïde

SE : Spin Echo : Type de Séquence d'IRM

SNC : Système Nerveux Central

SD : Standard deviation ou écart-type

Se : Sensibilité

SMOTE : Synthetic Minority Over-Sampling Technique: technique de ré-échantillonnage en cas de jeux de données non-équilibrés

SOX10 : Sry-Box Transcription Factor 10 : Marqueur très sensible pour les mélanomes

Sp : Spécificité

SPGR : Séquence d'IRM en écho de gradient avec destruction de l'aimantation transversale résiduelle: SPOiled Gradient-Recalled

SRM : Spectroscopie par Résonance Magnétique

STD : Déviation standard

SVM : Support Vector Machine : Type de classifieur utilisé dans les méthodes de machine learning

SWI : Séquence de susceptibilité magnétique, plus sensible que la séquence T2 étoile

T1w-GRE : Séquence T1 par écho de gradient

T1w-SE : Séquence T1 par écho de spin

TE : Temps d'Echo

TEP : Tomographie par Emission de Positons

TDM : TomoDensitoMétrie

TN : Vrai Négatif

TP : Vrai Positif

TTF-1 : Thyroid Transcription Factor-1 : Protéine régulant la transcription de gènes spécifiques à la thyroïde, au poumon et au diencéphale

TR : Temps de Répétition

TSE: Turbo spin écho : Séquence d'IRM rapide

UC : Carcinome Indifférencié

VPN : Valeur Prédicative Négative

VSC : Volume Sanguin Cérébral

INTRODUCTION

I. Contexte général

Les métastases cérébrales sont des localisations secondaires cérébrales d'une tumeur primitive et correspondent aux tumeurs intracrâniennes les plus fréquentes de l'adulte [1]. Elles sont dix fois plus fréquentes que les tumeurs cérébrales primaires [2] et représentent plus de la moitié

de toutes les tumeurs cérébrales [3]. La métastase cérébrale est responsable du premier symptôme chez près de la moitié des patients opérés pour une métastase cérébrale.

70% des métastases cérébrales sont secondaires à cinq cancers primitifs (par ordre décroissant : poumon, sein, mélanome, puis appareil urinaire et tube digestif sensiblement équivalents) [4].

La présentation clinique des symptômes cérébraux peut être due au fait que les symptômes thoraciques apparaissent assez tard dans le cancer du poumon par exemple, alors que les lésions cérébrales donnent des symptômes à un stade beaucoup plus précoce.

L'âge moyen des patients atteints de la maladie varie entre 51 et 55 ans en fonction des études. Le cancer du poumon est la tumeur la plus pourvoyeuse de métastases cérébrales quel que soit le sexe du patient et est la métastase cérébrale la plus fréquente chez les hommes. Chez les femmes, le cancer du sein est la métastase cérébrale la plus fréquente [5].

Comme pour tout syndrome de masse cérébrale, les symptômes sont focaux ou généraux en raison de l'hypertension intra-crânienne secondaire. Les céphalées sont le symptôme le plus courant, suivies par les vomissements, les signes neurologiques déficitaires, les troubles cognitifs et comportementaux, et les crises d'épilepsie (dans environ 15 à 25 % des cas) [6].

L'hémorragie intracrânienne n'est pas une présentation rare de métastases cérébrales (15%), en particulier dans les cas de mélanomes et de choriocarcinomes [7].

Le risque de développer des métastases cérébrales dépend aussi du stade de la tumeur primitive [8];[9] ;([10]. Ainsi, l'incidence des métastases cérébrales au moment du diagnostic est nettement supérieure lorsque la maladie est d'emblée métastatique dans d'autres localisations.

La proportion de patients présentant des métastases cérébrales identifiées au moment du diagnostic varie considérablement selon le type de cancer primaire [8].

26% de toutes les métastases cérébrales diagnostiquées chez les patients présentent un diagnostic synchrone de tumeur primaire [1].

La fréquence de survenue d'une métastase cérébrale dépend non seulement de la neurophilie de la tumeur primitive, mais aussi de la fréquence de cette tumeur. Par exemple, le mélanome malin comporte une ou plusieurs métastases intracrâniennes dans 70% des cas, mais n'est à l'origine que de 5% des métastases cérébrales du fait de sa rareté (environ 1% des tumeurs malignes).

Les leptoméniges, les nerfs crâniens, les vaisseaux sanguins, ou la voûte crânienne peuvent être affectés, mais le site préférentiel des métastases est de loin le parenchyme cérébral.

Le pronostic des patients atteints de métastases cérébrales est connu pour être mauvais, en fonction également de l'origine primaire de la néoplasie, avec une médiane de survie qui varie entre 3 mois et 2 ans, et un pourcentage de survie globale à 2 ans de 8% [11].

Le pronostic varie en fonction du primitif, de l'âge du patient, du degré de handicap fonctionnel défini selon l'index de Karnofsky (KPS), de l'étendue du contrôle de la maladie primaire, de la présence de métastases extracrâniennes ou de maladie leptoméningée [12], et du traitement, y compris les antécédents de résection chirurgicale, de radiothérapie et de chimiothérapie.

Le diagnostic de métastase cérébrale n'est pas évident, et ce notamment lorsqu'aucun cancer primitif n'est connu au moment de la découverte d'une lésion cérébrale. Ainsi, des lésions intracrâniennes solitaires peuvent se révéler métastatiques chez 15% des patients sans cancer systémique connu [13] (figure 1). Le diagnostic est d'autant plus complexe que la lésion secondaire peut être unique.

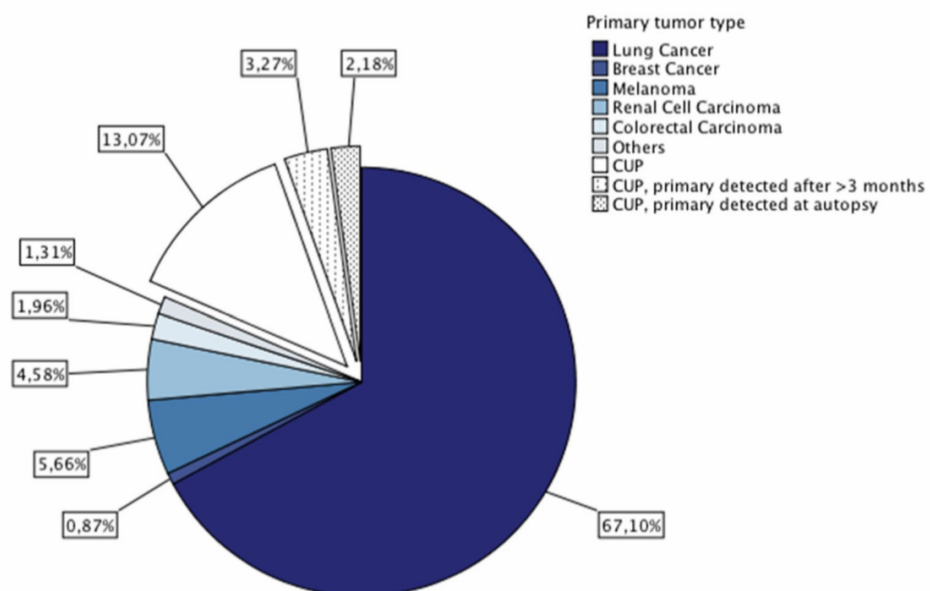


Figure 1: Distribution des types de tumeurs primitives inconnues au moment de la découverte de la métastase [14]

Dans ce contexte, afin de caractériser ces lésions, un bilan IRM est préconisé, comprenant des séquences morphologiques et des séquences avancées afin d'étayer le diagnostic entre lésion secondaire et autre lésion cérébrale. En effet, l'IRM est recommandée dans le but d'identifier

avec précision le nombre, la localisation et la taille des lésions, puis d'avancer sur la caractérisation lésionnelle. Elle sera également un examen clé pour le suivi lors des différentes thérapeutiques entreprises.

II. Que sait-on des voies de dissémination et de la localisation des lésions secondaires cérébrales ?

Les voies de propagation au SNC sont multiples (Figure 2).

La dissémination hémotogène est de loin le mécanisme le plus courant, et le dépôt des lésions a tendance à se produire soit aux endroits où le calibre des vaisseaux change (aux points de flux turbulent), soit aux points terminaux artériels. Dans le cerveau, ces zones constituent l'interface entre la substance grise et la substance blanche, et les régions de limite de différence de vascularisation [15].

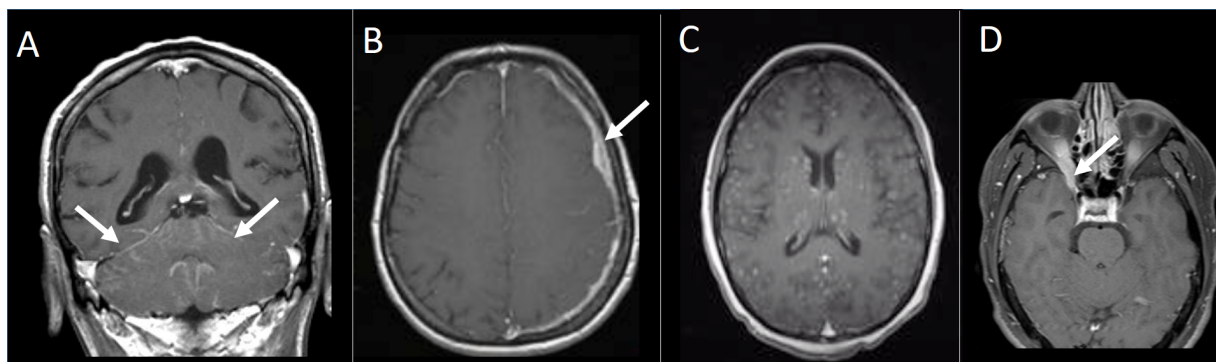


Figure 2 : Différentes voies de dissémination des métastases cérébrales, responsables de différentes localisations métastatiques distinguées en IRM : atteinte leptoméningée (A), pachyméningée (B), miliaire métastatique sur extension hémotogène (C), extension périneurale par un foramen de la base : infiltration de la fosse ptérygopalatine (D)

Dans le cas de la dissémination hémotogène d'une métastase, les cellules tumorales se détachent de la tumeur primitive et envahissent les tissus environnants, les veinules, les capillaires et le système lymphatique [16]. Ce processus s'appelle l'intravasation.

Les cellules tumorales circulantesensemencent ensuite la microvascularisation cérébrale puis commence un processus d'extravasation, qui est facilité par des cellules subissant un arrêt adhésif circulatoire. Une cascade de voies de signalisation intracellulaires impliquant les gènes COX2 et EGFR s'active ensuite pour permettre aux cellules tumorales de traverser la barrière hémato-encéphalique (BHE). Une fois traversée, les reliquats de la barrière hémato-

encéphalique sont appelés barrière tumoro-encéphalique (BTB). La constitution de la barrière hémato-encéphalique est schématisée sur la figure 3 (issue de [16]).

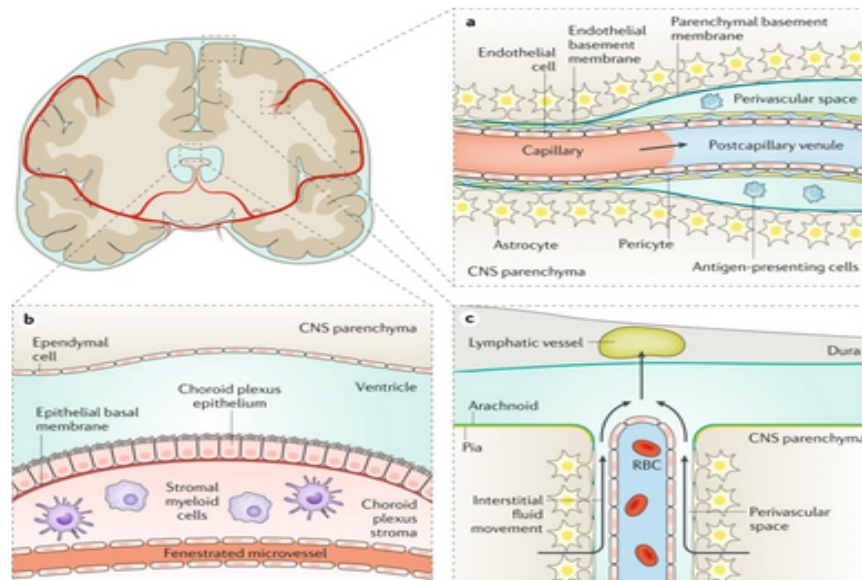


Figure 3: Barrière hémato-encéphalique [14]

- La barrière hémato-encéphalique est un isolant protecteur pour la signalisation neuronale et comprend des cellules endothéliales caractérisées par un faible taux de transcytose et une expression élevée des pompes d'efflux qui sont reliées par des jonctions étroites continues. Deux membranes basales (endothéliale et astrocytaire parenchymateuse), des péricytes intégrés et les pieds d'extrémité des astrocytes contribuent à ses fonctions de barrière. Au niveau des veinules postcapillaires, il existe un espace potentiel entre les membranes basales contenant des cellules clairsemées présentant l'antigène.
- La jonction sang – liquide cérébro-spinal (LCS) est formée de cellules épithéliales du plexus choroïde qui sont connectées via des jonctions serrées ; les capillaires du plexus choroïde ont des fenestrations et des espaces intercellulaires qui permettent le mouvement des molécules. Les leucocytes circulants pénètrent rarement dans le parenchyme cérébral dans des conditions homéostatiques, bien que des cellules immunitaires parenchymateuses endogènes soient présentes.
- Le système lymphatique cérébral permet aux cellules immunitaires et aux protéines du LCS de s'écouler par les canaux lymphatiques des méninges jusqu'aux ganglions lymphatiques cervicaux profonds.

Les cellules tumorales qui quittent le système vasculaire du SNC pour pénétrer dans le cerveau ne se développent pas nécessairement [17]. Si les conditions locales ne sont pas propices, les cellules tumorales peuvent entrer en apoptose ou rester dormantes pendant des mois, voire des années. Des métastases cérébrales solitaires peuvent ainsi se développer chez les patients des années (et parfois des décennies) après la guérison apparente du cancer primitif [18]. Ce phénomène est particulièrement important dans le mélanome et le cancer du sein, mais il se produit également avec d'autres tumeurs. La barrière hémato-encéphalique limitant la pénétration des thérapies systémiques dans le cerveau et ces patients ne présentant pas de récurrence de la maladie dans d'autres régions du corps, la plupart des cellules

cancéreuses doivent avoir été acheminées vers le cerveau et être restées latentes pendant que la tumeur d'origine a été traitée avec succès [19].

Si les conditions locales sont propices, des interactions micro-environnementales complexes, des cascades neuro-inflammatoires et, éventuellement, une néovascularisation (Figure 4 (issue de [16])) sont impliquées dans l'établissement d'une nouvelle métastase cérébrale [16].

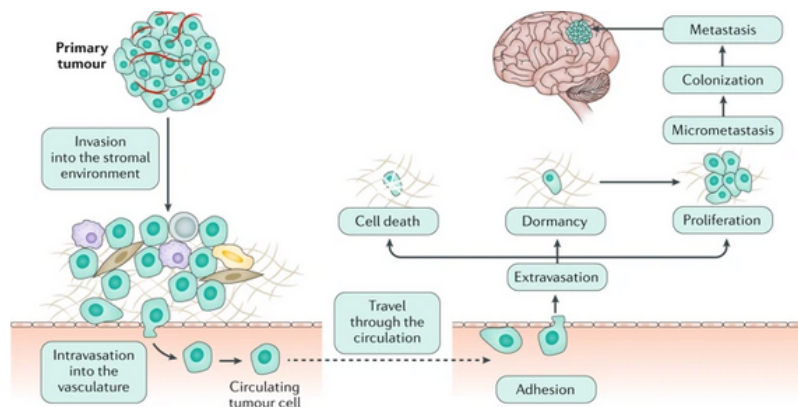


Figure 4: Dissémination des cellules tumorales métastatiques [4]

Environ 85% des métastases cérébrales se trouvent dans les hémisphères cérébraux, 10 à 15 % dans le cervelet et seulement 3 % environ dans le tronc cérébral. Très rarement, les métastases cérébrales peuvent être localisées dans les noyaux gris centraux ou le thalamus.

Certaines tumeurs primaires, comme celles du rein et du côlon, sont plus susceptibles de métastaser dans le cervelet que les cancers du poumon ou ceux qui se développent ailleurs dans le corps [15], ce qui suggère que ces tumeurs primaires accèdent plus souvent au SNC par la circulation postérieure.

La dissémination leptoméningée, qui représente la dissémination sous-arachnoïdienne, se fait par voie artérielle ou bien, est secondaire à une dissémination par le liquide cérébro-spinal (LCS). Elle est surtout observée dans les carcinomes pulmonaires et mammaires [20].

La présence de multiples foyers tumoraux métastatiques autour des vaisseaux sanguins sans invasion parenchymateuse est décrite comme un modèle de métastase cérébrale miliaire, et a été signalée le plus souvent dans les carcinomes pulmonaires non à petites cellules [21].

Les localisations dures résultent d'une dissémination hémotogène ou d'un envahissement à partir des os du crâne. L'atteinte intracrânienne peut résulter d'une extension péri-neurale ou par les foramens de la base du crâne, notamment dans les tumeurs ORL.

Les autres sites rares comprennent le plexus choroïde, la glande pinéale, l'hypophyse et le nerf optique [20].

Par ailleurs, en moyenne 25% des lésions secondaires cérébrales sont uniques et 75 % multiples au moment de leur découverte [22]. Leur distribution au sein du parenchyme cérébral dépend également du primitif. Les cancers du sein et du poumon sont connus comme présentant plus fréquemment des atteintes infratentorielles [23].

III. Démarche diagnostique devant la découverte d'une possible métastase cérébrale ?

Le diagnostic de métastase cérébrale n'est pas évident, et ce notamment lorsqu'aucun cancer primitif n'est connu au moment de la découverte d'une lésion cérébrale. Ainsi, des lésions intracrâniennes solitaires peuvent se révéler métastatiques chez 15% des patients sans cancer systémique connu [13]. Ainsi, le diagnostic est plus complexe lorsque la lésion est unique.

La démarche diagnostique repose d'une part sur l'analyse IRM précise de la lésion cérébrale et d'autre part sur un bilan complet exhaustif à la recherche du primitif lorsque celui-ci n'est pas connu.

1. Bilan IRM devant une possible localisation cérébrale.

La modalité d'imagerie la plus sensible pour détecter les métastases cérébrales est l'IRM. Ses capacités d'imagerie multiplanaire, son contraste tissulaire supérieur, l'élimination des artefacts osseux et ses paramètres polyvalents sont extrêmement efficaces pour l'évaluation de ces patients avec injection de gadolinium pour la détection et la caractérisation des lésions cérébrales.

Le bilan initial en IRM comprend des séquences morphologiques et des séquences avancées afin d'étayer le diagnostic entre lésion secondaire et autre lésion cérébrale, ce d'autant plus, si aucun primitif n'est connu.

Le protocole IRM comprend un ensemble de séquences morphologiques dont le but est de permettre de détecter, localiser et mesurer les lésions secondaires.

De par sa résolution spatiale, son excellent signal sur bruit et sa capacité à détecter une lésion cérébrale présentant un rehaussement, ce qui est le cas des lésions secondaires, la séquence 3DT1 après injection est la séquence clé de la détection des lésions métastatiques intracrâniennes.

a. Séquences T1 sans et avec injection de gadolinium

Les séquences T1 avant et T1 après injection de gadolinium recherchent une prise de contraste, présente dans la vaste majorité des tumeurs cérébrales primitives et secondaires, liée à la non spécifique rupture de barrière hémato encéphalique.

Cette perturbation est responsable d'une fuite de gadolinium dans les espaces extracellulaires de la tumeur. L'élément gadolinium, utilisé comme agent de contraste RM, est fortement paramagnétique à faible concentration, et est responsable d'un effet de raccourcissement du T1, qui augmente le signal T1 à l'intérieur de la tumeur, fournissant un fort contraste par rapport aux tissus normaux. Ce fort contraste permet la détection de petites lésions (moins de 5 mm), en particulier les métastases cérébrales [24].

Plusieurs paramètres peuvent améliorer le taux de détection des petites lésions.

Le taux de détection augmente avec la dose d'agent de contraste [25]. Cependant, l'augmentation de la dose n'est pas recommandée dans la pratique clinique, même avec les agents de contraste modernes, en raison des effets inconnus du dépôt à long terme de gadolinium in vivo chez les patients ayant subi une injection répétée de produit de contraste [26][27].

L'utilisation d'une IRM à haute intensité de champ augmente le rapport contraste/bruit et permet de réduire la dose d'agent de contraste à 3 T par rapport à 1,5 T [28].

Comme la perturbation de la barrière hémato-encéphalique dépend du temps, il a été démontré que le degré de renforcement augmente avec le temps, et il est recommandé d'acquérir la séquence pondérée T1 après le renforcement du contraste 10 à 15 minutes après l'injection de l'agent de contraste [29].

Deux principaux types de séquences peuvent être utilisés pour réaliser une imagerie pondérée T1 : l'imagerie 3D T1 par écho de gradient (T1w-GRE), est la plus répandue [30] (figure 5). Plus récemment, des études ont révélé un taux plus élevé de détection des métastases grâce à

l'imagerie par 3D T1 FSE (écho de spin (T1w-SE) [31]. Ces séquences ont cependant l'inconvénient de ne pas visualiser les veines, ni les reliefs du visage, pouvant gêner leur utilisation dans le cadre d'une éventuelle neuronavigation.

Utilisation des techniques de soustraction :

L'acquisition de séquences pondérées T1 avec et sans injection d'agent de contraste permet de calculer une carte de soustraction, ce qui est très utile pour évaluer les lésions hémorragiques qui peuvent avoir un signal T1 spontanément élevé et pour améliorer le contraste et la délimitation des lésions [32],[31].

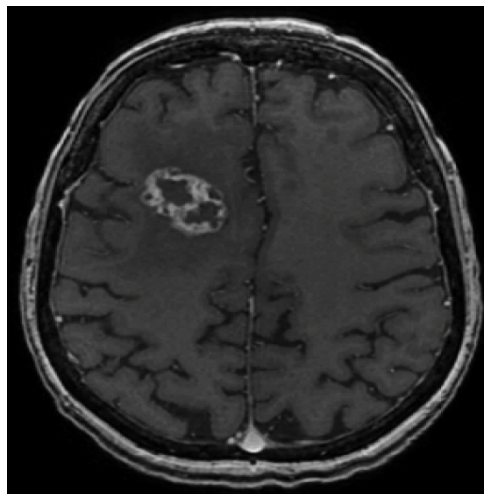


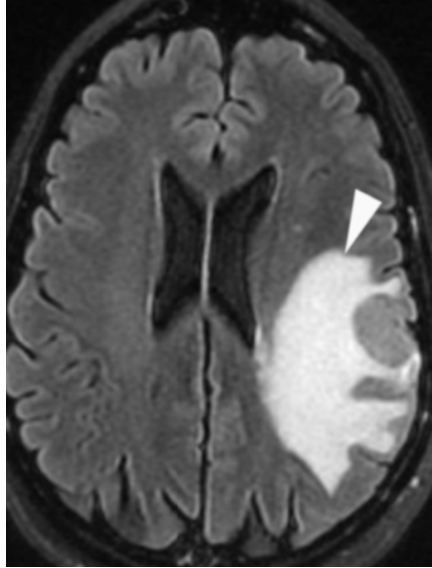
Figure 5: Séquence volumique 3DT1 EG après injection de gadolinium retrouvant un processus expansif occupant l'espace intra parenchymateux, frontal droit, correspondant à une métastase cérébrale. La périphérie se rehausse et le centre paraît de signal liquidien nécrotique.

b. Séquence FLAIR

La séquence FLAIR est une séquence pondérée en T2 avec une impulsion supplémentaire de récupération par inversion qui permet une suppression du signal du liquide céphalorachidien et est donc sensible à l'œdème infiltrant (figure 6).

Dans les métastases cérébrales non traitées, une zone de forte intensité entourant une tumeur accrue est liée à un œdème vasogénique.

Cet œdème réactionnel ne contient pas de cellules tumorales et n'est donc pas considéré comme une cible tumorale [24].



*c. Séquence T2**

La séquence écho de gradient T2* est utilisée au stade du diagnostic afin d'évaluer le contenu des hémorragies des tumeurs cérébrales. Pour les tumeurs secondaires, l'hémorragie est classiquement observée dans les mélanomes, les choriocarcinomes, les carcinomes des cellules rénales, la thyroïde mais aussi dans les cancers du poumon et du sein en raison de leur forte prévalence [34].

d. Séquence de diffusion

La séquence d'imagerie pondérée en fonction de la diffusion se trouve à la jonction de l'imagerie morphologique et fonctionnelle. Dans le contexte de l'exploration des tumeurs cérébrales, une restriction du coefficient de diffusion est la preuve d'une forte cellularité [35] (figure 7) . La diffusion a montré sa pertinence pour le classement des tumeurs.

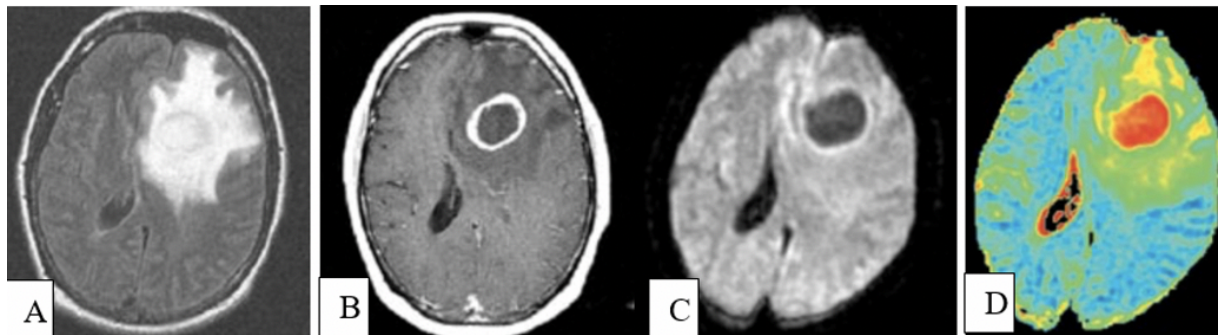


Figure 7: Bilan IRM d'une lésion métastatique en séquences morphologiques.

Métastase nécrotique d'un néoplasme pulmonaire. A) Lésion frontale gauche hyperintense en FLAIR avec oedème périlésionnel. B) SE T1 avec injection : lésion nécrotique hypointense à rehaussement annulaire. C) Diffusion : lésion hypointense. D) ADC nettement augmenté sur la cartographie. Ceci permet d'évoquer le diagnostic de tumeur nécrosée, alors que la séquence SE T1 avec injection, retrouve une lésion nécrotique à rehaussement annulaire, difficile à distinguer d'un abcès.[36]

e. Séquences avancées

En plus des séquences morphologiques dites de base, des **séquences avancées** comme la perfusion cérébrale et la spectroscopie cérébrale nous aident à caractériser les lésions cérébrales, et notamment, devant une lésion d'allure tumorale se rehaussant, nous permettent d'augmenter les éléments pour différencier métastase et lésion gliale de haut grade, fréquent diagnostic différentiel [37][38].

Elles ne nous permettent cependant pas de nous prononcer par rapport à l'origine métastatique.

Certaines études ont essayé de dégager des éléments pouvant orienter vers l'origine du primitif, comme l'étude de Chernov [39], retrouvant, en spectroscopie, une augmentation de la teneur en lipides mobiles comme un signe faible d'une origine colique pour les métastases. L'étude de Grand [38] avait, quant à elle, étudié la perfusion et identifié les métastases hypervasculaires (complicant habituellement un cancer du rein, un mélanome ou un choriocarcinome). Cette information permettrait ainsi, quand la néoplasie primitive n'est pas connue, d'orienter sa recherche.

La spectroscopie combinée à l'IRM de perfusion étudiée dans l'étude de Huang [40], tendait à démontrer que les différences dans les rapports choline-créatine entre les métastases du cancer du poumon et du sein étaient corrélées avec les différences de volume sanguin cérébral relatif, mais les tissus n'étaient pas disponibles pour la comparaison afin de rechercher une base pathologique unificatrice.

Ainsi, les séquences morphologiques aidées des séquences avancées ont pour rôle de détecter et de caractériser les lésions secondaires. Cependant, avancer vers l'origine du primitif en étudiant le signal de la lésion cérébrale en IRM restait un challenge non résolu. En effet, devant ces 4 lésions cérébrales analysées en IRM (figure 8), aucun élément morphologique ou d'imagerie avancée ne permet jusqu'à présent de déterminer leur primitif.

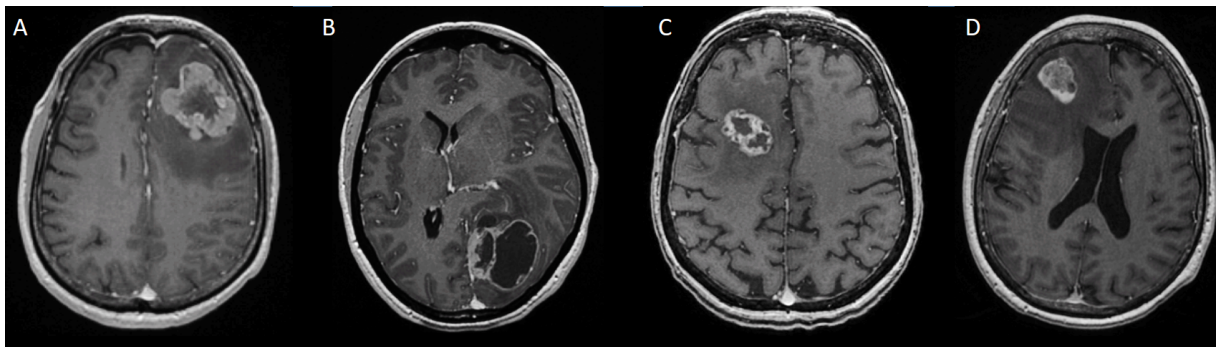


Figure 8: Volume T1 EG après injection de gadolinium retrouvant des lésions métastatiques secondaires d'un colon (A), mélanome (B), poumon (C), sein (D).

2. Bilan à la recherche du primitif

Le bilan devant une lésion cérébrale métastatique probable comprend également la recherche du primitif. Le bilan exhaustif est dans un premier temps clinique. En cas de tumeur primaire inconnue, il est recommandé de procéder à un examen physique complet (y compris les testicules et l'inspection de la peau comprenant un examen de la tête et du cou, un examen rectal, un examen pelvien et un examen mammaire) [41] [42].

L'utilité du scanner thoraco-abdomino-pelvien est très bien documentée, et permet de détecter un site primaire chez 30 à 35 % des patients atteints de métastases cérébrales de primitif inconnu. La tomodensitométrie peut également fournir des indications pour le choix du site optimal pour la biopsie [43].

La mammographie et les autres tests d'imagerie mammaire sont moins fréquemment réalisés chez ces patientes, car il est peu probable que le carcinome du sein se présente sous la forme de métastase cérébrale de primitif inconnu mais doivent être réalisés si le primitif n'est pas retrouvé [44].

La pertinence clinique de la tomographie par émission de positrons (TEP) pour détecter les tumeurs primaires en plus des procédures conventionnelles est démontrée si le bilan suscité est négatif. Le TEP pourrait être une méthode plus sensible pour détecter les tumeurs primaires inconnues qu'une approche diagnostique conventionnelle [45].

La combinaison TEP/TDM pourrait également être un outil diagnostique utile [46].

La tomographie par émission de positrons sur corps entier 2-désoxy-2-[18F]fluoro-D-glucose (CT/FDG-PET) peut contribuer à la prise en charge des patients atteints de tumeurs de primitif inconnu et en particulier de ceux qui présentent des adénopathies cervicales et des métastases uniques [47].

Le bilan biologique comprenant des marqueurs sanguins tumoraux permet également d'orienter le diagnostic étiologique. Ainsi, l'évaluation sérique de l'a-foetoprotéine (AFP), de la gonadotrophine chorionique humaine (hCG), de la chromogranine A plasmatique et du PSA (antigène plasmatique sanguine) est suggérée chez les patients de sexe masculin afin d'exclure les tumeurs extragonadiques des cellules germinales, les tumeurs neuroendocrines et les cancers de la prostate.

Les marqueurs tumoraux sériques CA 125, CA 19-9 et CA 15-3 ne présentent en revanche aucune corrélation avec l'histologie, le site de la maladie primaire ou le nombre de métastases. De plus, ils n'offrent aucune valeur diagnostique ou prédictive chez les patients atteints de métastase cérébrale sans primitif connu. L'ACE n'a une valeur prédictive possible que pour la survie [48].

Enfin, en fonction de la clinique et des résultats du scanner et des marqueurs tumoraux, certains examens spécifiques peuvent être réalisés comme les endoscopies digestives.

3. Diagnostic histologique

Le diagnostic final sera réalisé par l'analyse histologique de la métastase ou du primitif. Il est intéressant d'analyser les différences en fonction du primitif retrouvées en histologie et en immunomarquage, permettant de distinguer les profils de lésion.

a. Caractéristiques macroscopiques

Macroscopiquement, les lésions métastatiques sont généralement des masses sphériques fortement délimitées qui n'infiltrent pas le parenchyme cérébral environnant, mais provoquent un œdème dû à l'effet de masse. L'œdème du parenchyme cérébral environnant est important et parfois disproportionné par rapport à la taille des lésions. Les tumeurs les plus volumineuses ont tendance à avoir des centres nécrotiques. Les hémorragies ne sont pas rares et sont le plus souvent associées à des métastases de mélanome (Figure 9), de choriocarcinome et de CCR à cellules claires [49][50].

Parfois, les métastases des carcinomes du poumon et du sein semblent kystiques. Certaines lésions métastatiques peuvent présenter des caractéristiques de la tumeur primaire, telles que l'aspect mucoïde brillant des adénocarcinomes sécrétant de la mucine ou les lésions pigmentées gris-noir du mélanome (Figures 9 et 10).

Il est toutefois important de noter que l'absence de pigmentation n'exclut pas le mélanome, car certains exemples sont amélaniques.

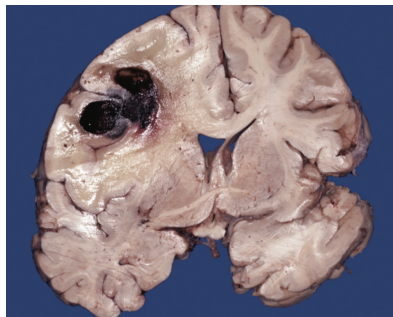


Figure 9 : Foyer de mélanome métastatique impliquant le lobe frontal au niveau de la commissure antérieure avec hémorragie associée et œdème environnant [20]

b. Caractéristiques microscopiques

La majorité des métastases sont nettement délimitées par rapport au cerveau environnant et présentent des caractéristiques histologiques similaires à celles de la tumeur primaire (figure 10). Le cerveau environnant présente une raréfaction des tissus, une néovascularisation et une gliose, qui se manifestent par un œdème vasogénique à l'imagerie.

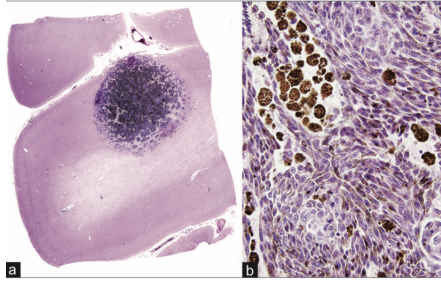


Figure 10 : Mélanome métastatique (a) A faible grossissement, la démarcation est relativement nette, à l'exception d'un étalement périphérique limité le long des espaces périvasculaires. La substance blanche adjacente apparaît pâle et oedémateuse. (b) A plus fort grossissement : un grossissement plus élevé révèle une néoplasie épithélioïde et filiforme à pigmentation variable, avec des nucléoles proéminents, une activité mitotique et des mélanophages dispersés, typiques du mélanome [20]

c. Différences anatomopathologiques entre les métastases

1) Mélanome Malin

Le mélanome est le plus souvent composé de cellules épithélioïdes avec un cytoplasme rose abondant, de gros noyaux et des nucléoles géants rouge cerise (Figure 10b). Il peut également être composé de cellules fusiformes, imitant un sarcome.

Une combinaison de marqueurs à sensibilité variable (protéine S-100, très sensible) et de marqueurs spécifiques (comme HMB-45, Melan-A, tyrosinase) semble être la meilleure approche pour le diagnostic d'un mélanome primaire inconnu du SNC.

2) Adénocarcinome Pulmonaire

Le TTF-1 est exprimé dans les noyaux des cellules thyroïdiennes normales, et l'épithélium respiratoire dans le poumon. Le TTF-1 est positif dans la majorité des carcinomes d'origine pulmonaire.

Une plus grande spécificité est obtenue en combinant les marqueurs, la plupart des adénocarcinomes pulmonaires primaires étant CK7-positifs, CK20-négatifs et TTF-1-positifs [20] (Figure 11).

3) Carcinome Du Sein

Comme les adénocarcinomes pulmonaires, les carcinomes mammaires sont CK7-positifs et CK20-négatifs (figure 11) ; cependant, ils sont négatifs pour le TTF-1 et la napsine A [20]. Les autres marqueurs permettant de vérifier l'origine mammaire d'une métastase comprennent la mammaglobine et le récepteur des œstrogènes, ce dernier étant moins spécifique.

Tous ces marqueurs sont cependant imparfaits et peuvent même être exprimés occasionnellement dans les carcinomes pulmonaires [51].

4) Côlon et Autres Carcinomes Gastro-Intestinaux

Les adénocarcinomes du côlon sont CK7-négatifs et CK20-positifs (figure 11), ce qui permet de les distinguer facilement des adénocarcinomes pulmonaires et mammaires, même s'ils sont mal différenciés. Le CDX2 est un facteur de transcription normalement exprimé dans l'intestin, les voies biliaires, et représente un marqueur très sensible et spécifique des adénocarcinomes intestinaux [52]. Il montre une forte expression nucléaire concernant jusqu'à 90% des adénocarcinomes colorectaux et duodénaux. Les carcinomes mucineux ovariens et certains adénocarcinomes de la vessie peuvent être fortement positifs pour le CDX2 ; cependant, ils ne métastasent que rarement dans le SNC, généralement bien après que le diagnostic primaire ait été établi [20].

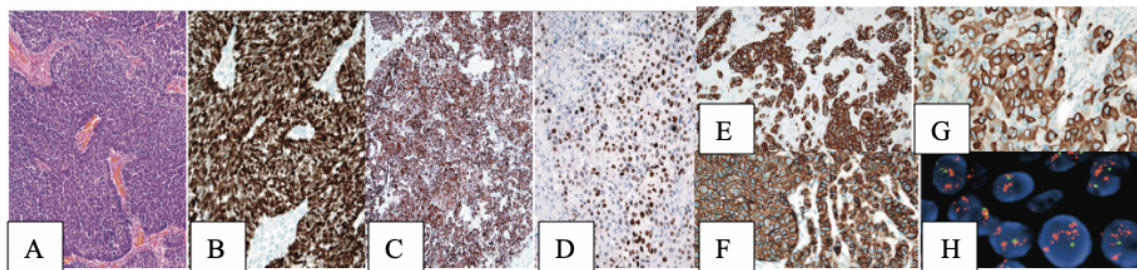


Figure 11 :Carcinome pulmonaire métastatique à petites cellules.

- (a) L'histopathologie typique comprend des nids de petites cellules d'apparence primitive avec des noyaux hyperchromatiques, un rapport nucléaire/cytoplasmique élevé et des moulages nucléaires.

(b) Presque tous les cas sont fortement positifs pour le TTF-1

-Adénocarcinome pulmonaire métastatique peu différencié avec immunoréactivité pour le CK7(c)et le TTF-1 (d)

-Carcinome mammaire métastatique à forte positivité CK7 (e), expression de la mammaglobine (g), immunoréactivité du modèle de membrane pour HER2 (f), et amplification du gène HER2 à haut niveau par hybridation in situ par fluorescence (h ; centromère 17 en vert et HER2 en rouge)[20]

IV. La radiomique, une nouvelle voie d'analyse d'image ?

Au cours des dernières années, le domaine de l'analyse des images médicales a connu une croissance exponentielle, avec un nombre croissant d'outils de segmentation automatique, et une augmentation des jeux de données structurés et annotés disponibles. Ces progrès ont facilité le développement de processus d'extraction à haut débit de caractéristiques quantitatives qui aboutissent à la conversion des images en données exploitables et à l'analyse ultérieure de ces données pour l'aide à la décision ; cette pratique est appelée radiomique.

La radiomique repose sur un constat simple : l'œil humain a des limites, et même entraîné à interpréter des images médicales, il sous-exploite les données contenues dans celles-ci.

Des études récentes ont fait état de liens étroits entre la signature radiomique de diverses tumeurs et leur phénotype et les profils génomiques et protéomiques [53][54][55].

Les images contiendraient des données phénotypiques, morphologiques et génotypiques propres aux tumeurs qui seraient reflétées dans l'intensité des niveaux de gris des voxels, et leurs interactions spatiales.

Ces données sont potentiellement combinées avec d'autres données (données cliniques, biologiques, génomiques...) et exploitées à l'aide d'outils bioinformatiques sophistiqués, par le biais de modèles de classification et de régression, afin de développer des modèles susceptibles d'améliorer la précision du diagnostic, du pronostic et de la prévision de la réponse à la thérapie [56](figure 12).

Cette pratique s'oppose à la pratique traditionnelle qui consiste à traiter les images médicales comme des images destinées uniquement à l'interprétation visuelle.

La radiomique a été initiée dans les études oncologiques, comme par exemple en guidant la classification des tumeurs par stade, par histologie et par profil moléculaire [57] [58] [59], ou encore en guidant le pronostic et la prédiction de la réponse au traitement [60] [61] [62].

Elle peut être réalisée à l'aide d'images tridimensionnelles provenant de la tomodensitométrie, de l'imagerie par résonance magnétique et de la tomographie par émission de positons ou de toute autre modalité d'imagerie.

Les caractéristiques des images sont extraites de volumes d'intérêt, qui peuvent être soit des tumeurs entières, soit des sous-volumes définis à l'intérieur des tumeurs, appelés habitats [56].

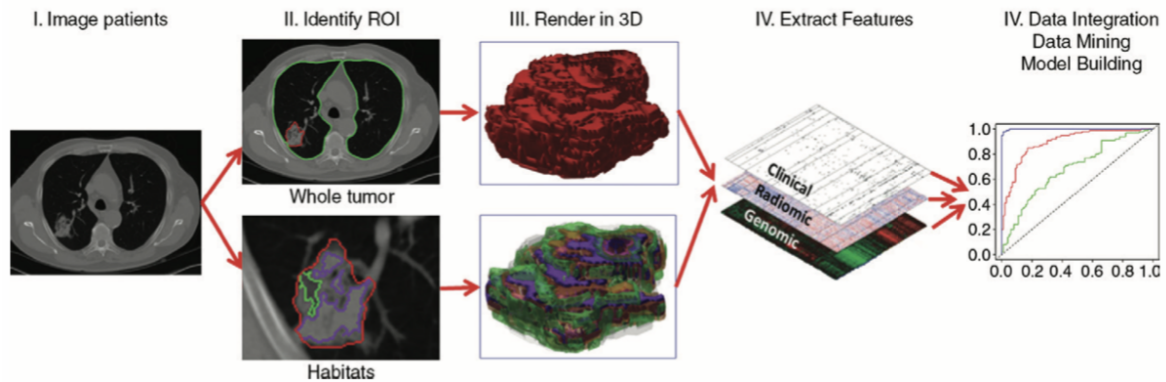


Figure 12: Organigramme de processus de la radiomique et de l'utilisation de la radiomique dans l'aide à la décision.

La radiomique commence par l'acquisition d'images médicales. À partir de ces images, on peut identifier une région d'intérêt (ROI) qui contient soit la tumeur entière, soit des sous-régions (c'est-à-dire des habitats) à l'intérieur de la tumeur. Les caractéristiques quantitatives sont extraites de ces ROI et peuvent être combinées à d'autres données, telles que les données cliniques ou génomiques. Ces données sont ensuite exploitées pour développer des modèles diagnostiques, prédictifs ou pronostiques, figure issue de Gillies RJ et al., *Radiomics :Images are more than pictures, they are data* [56].

Le terme -omique est utilisé dans les domaines de la recherche médicale qui génèrent des données complexes de haute dimension à partir d'objets uniques ou d'échantillons [63].

Le concept -omique s'applique facilement à l'imagerie tridimensionnelle pour plusieurs raisons : une image tridimensionnelle d'un patient peut facilement contenir des millions de voxels, l'unité volumique élémentaire des images radiologiques. En outre, un volume d'intérêt comme une tumeur peut contenir des centaines de caractéristiques mesurables décrivant la taille, la forme et l'hétérogénéité tumorale. La radiomique permettrait ainsi d'analyser la tumeur dans son intégralité, et pourrait être un complément à la biopsie qui n'analyse qu'un échantillon de la région tumorale. De plus, la répétabilité des examens d'imagerie permet une réévaluation dans le temps des critères pronostiques et la surveillance des variations temporelles des caractéristiques de la tumeur.

Plusieurs études ont déjà montré la supériorité de la radiomique, par rapport à la seule interprétation visuelle des images. L'étude de Zhu et al. indique que, dans le cas du cancer du sein invasif, les activités transcriptionnelles de diverses voies génétiques et les expressions des miARN sont associées à la forme et à la texture de la tumeur sur des images dynamiques de résonance magnétique (IRM) avec injection de produit de contraste [54]. De même, Grossmann et al. signalent que l'hétérogénéité intra-tumorale et la dispersion des intensités sur les images TDM acquises de patients atteints d'un cancer du poumon sont associées aux

caractéristiques moléculaires de la tumeur, notamment l'activité de transcription de l'ARN polymérase et la voie d'autodégradation d'une ubiquitine-ligase [55].

L'analyse morphologique des tumeurs en imagerie oncologique repose sur un faisceau d'arguments qualitatifs, basés sur la caractérisation visuelle des tumeurs en termes de distribution des intensités (hétérogène, homogène), de forme (spiculée, non spiculée), ou d'analyse des limites (régulières, floues). Les approches quantitatives peuvent être monodimensionnelles (mesure du plus grand diamètre de la tumeur pour le critère RECIST) ou bidimensionnelles [42] et sont couramment utilisées pour le suivi du traitement et non pour le diagnostic initial.

V. Objectifs de l'étude

Nous avons vu que les métastases cérébrales sont issues de tumeurs primitives très différentes sur les plans physiopathologique, histopathologique et moléculaire, mais que leur aspect morphologique et fonctionnel est très similaire en imagerie.

Le diagnostic de certitude repose donc uniquement sur la biopsie et l'analyse histologique et biomoléculaire de la pièce tumorale.

La radiomique, en tant que domaine de recherche récent en médecine, pourrait constituer un outil d'aide au diagnostic dans la caractérisation et la différenciation des tumeurs primitives de ces lésions secondaires cérébrales.

L'objectif de ce travail a ainsi été de développer un modèle de classification visant à distinguer le site d'origine principal de la métastase cérébrale à partir de données cliniques (âge et sexe), des données spatiales (localisation de la lésion au sein du cerveau) ainsi que des caractéristiques radiomiques provenant de l'IRM 3DT1 en écho de gradient (SPGR) post gadolinium. Pour ce faire, une cohorte de 237 métastases, chez 188 patients, représentant les 4 plus fréquents types de métastases cérébrales a été utilisée.

MATERIEL ET METHODES

I. Différentes étapes d'une analyse radiomique

Le processus conventionnel d'analyse radiomique est constitué de 5 étapes principales :

- Acquisition des images
- Segmentation des volumes d'intérêt
- Pré-traitement des images et des masques
- Extraction des index radiomiques de forme, d'intensité et de texture
- Analyse statistique : Machine Learning

II. Acquisition des images

1. Population

Les données IRM de la cohorte complète ont été collectées rétrospectivement dans deux centres, au Centre Hospitalier de Sainte Anne (GHU de Paris) et au Centre Hospitalier Universitaire d'Amiens.

La base de données comportait des patients pour lesquels les acquisitions ont eu lieu de 2008 à 2019.

Les patients devaient présenter au moins une métastase cérébrale, intra axiale, de diamètre supérieur ou égal à 1 cm, avant traitement ou opération, confirmée par l'analyse histologique après biopsie ou chirurgie du cerveau, provenant soit d'un cancer pulmonaire, soit d'un cancer mammaire, soit d'un mélanome, ou soit d'un cancer colorectal.

Nous avons exclu les petites lésions (diamètre < 1 cm), les lésions extra-axiales et les patients ayant des antécédents de chirurgie cérébrale et/ou de radiothérapie et/ou d'immunothérapie ainsi que les séquences IRM présentant des artéfacts.

La figure 13 montre le flow chart pour l'inclusion des patients.

Par ailleurs, le type de néoplasie primaire, l'âge et le sexe du patient ont été collectés.

Nous avons inclus 188 patients dont 237 métastases.

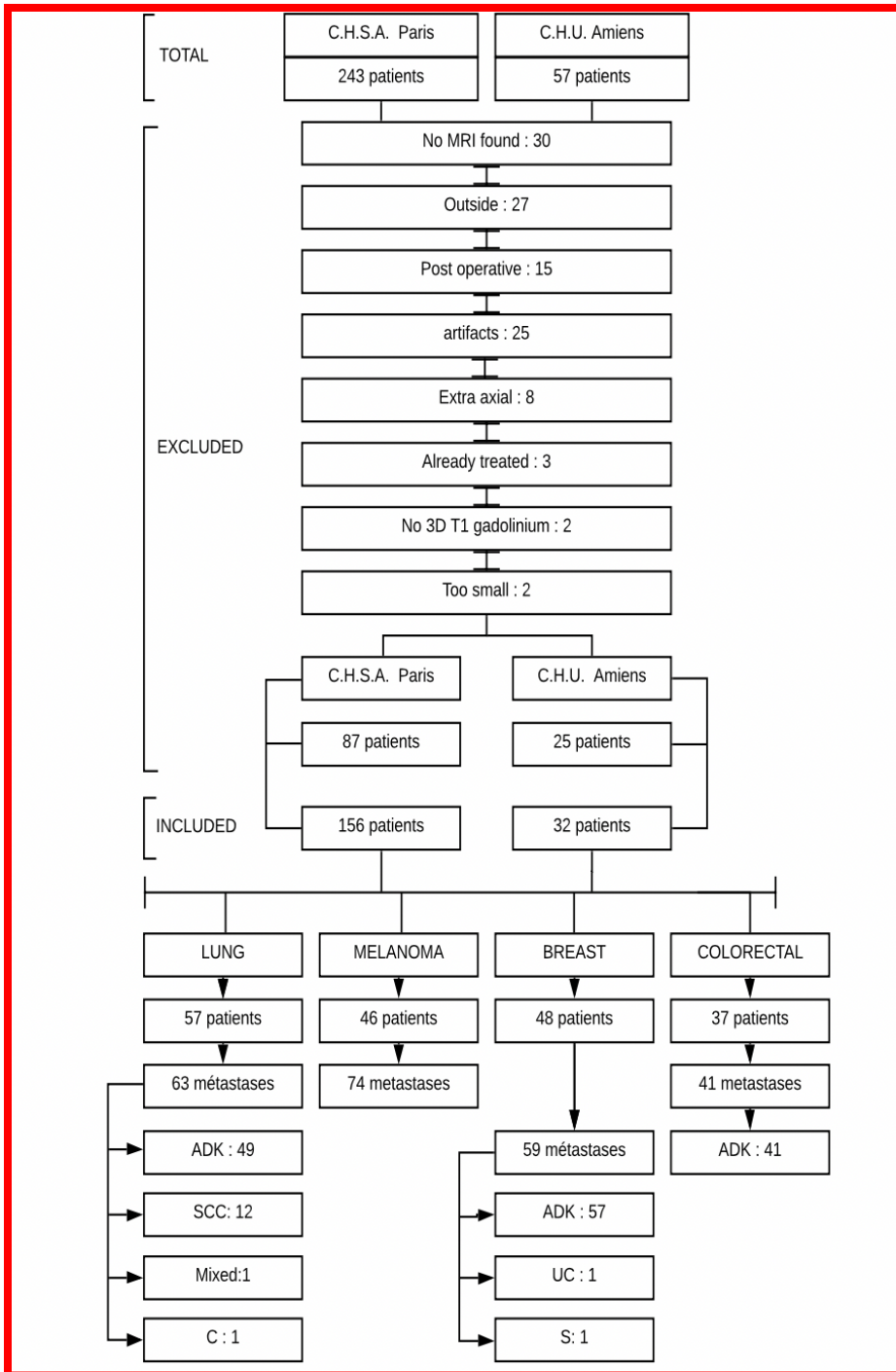


Figure 13 : Critères d'inclusion et d'exclusion des patients.

ADK = adénocarcinome, SCC = cancer épidermoïde (squamous cell carcinoma), Mixed = ADK + SCC, C = carcinoma non spécifique, UC = carcinoma indifférencié, S = sarcome, Outside : examen réalisé avec un autre constructeur que General Electric

2. Données IRM

Pour faire une analyse radiomique, il faut partir d'un choix de séquence.

La séquence 3D T1 post gadolinium a été choisie dans notre étude car elle représente la séquence la plus fréquemment utilisée pour la détection de métastases cérébrales.

Les caractéristiques radiomiques extraites des images IRM sont fortement dépendantes des paramètres d'acquisition (champ magnétique, temps d'écho, temps de répétition, taille des voxels..). Nous avons donc sélectionné deux sites avec des imageurs IRM provenant du même constructeur, General Electric (Milwaukee, WI) (1,5 Tesla et 3 Tesla).

Tous les examens comprenaient une séquence tridimensionnelle pondérée T1 en écho de gradient (SPGR) après injection de contraste gadoliné et ces séquences ont été acquises avec un temps de répétition (TR) = 6,5-15,9 ms, un temps d'écho (TE) = 2,1-7 ms et un temps d'inversion (IT) = 450 ms. Tous les patients ont reçu une dose de produit de contraste de 10 mL de gadolinium (Dotarem, Guerbet, France).

Cette séquence 3DT1 post-gadolinium a été réalisée pour tous les patients et a été la séquence utilisée pour la segmentation des lésions tumorales et comme donnée d'entrée de l'analyse radiomique.

La radiomique étant un domaine de recherche, toutes les imageries implémentées dans le pipeline ont dû être anonymisées puis exportées au format DICOM vers un logiciel de segmentation.

III. Segmentation du volume d'intérêt

La segmentation des volumes d'intérêt est une étape très importante puisqu'elle va délimiter le volume dans lequel est appliquée l'analyse radiomique. Ainsi, tous les voxels contenus dans le volume segmenté seront analysés comme faisant partie de la lésion par le processus de radiomique. Du fait de l'hétérogénéité tumorale, les tumeurs malignes ont généralement des contours irréguliers donc la segmentation est difficile et peut poser une question de reproductibilité inter-opérateurs.

Il existe des méthodes de segmentation totalement automatisées, semi-automatisées, ou manuelles [64][65][66]. Les méthodes automatiques avec correction manuelle ou méthodes semi-automatiques sont le plus souvent utilisées dans les études radiomiques [67].

Nous avons utilisé une méthode de segmentation semi-automatique combinant le seuillage d'intensité et la correction manuelle sur la version 4.0 de Mango (figure 14).

Le seuillage d'intensité repose sur la définition par l'utilisateur d'une plage de niveaux de gris, et la segmentation par le logiciel de tous les voxels compris dans cette plage. L'avantage de cette méthode est sa facilité d'utilisation. La limite de cette technique est qu'elle nécessite une différence de signal importante entre la lésion à segmenter et le tissu avoisinant. Le logiciel de segmentation peut ainsi contourner à défaut des éléments présents dans l'image, de même intensité que le volume d'intérêt, mais ne faisant pas partie de la lésion à segmenter. Dans notre cas des métastases cérébrales, les techniques de seuillage d'intensité peuvent être artéfactées par la segmentation d'éléments vasculaires, ou durs prenant le contraste au même titre que la lésion à segmenter, ce qui suppose une correction manuelle de la segmentation.

Toute la paroi de la lésion et la nécrose centrale ont été incluses dans le volume final.

L'œdème péritumoral vasogénique et les structures adjacentes comme l'os, les sinus durs ou les vaisseaux ont été exclus de la segmentation.

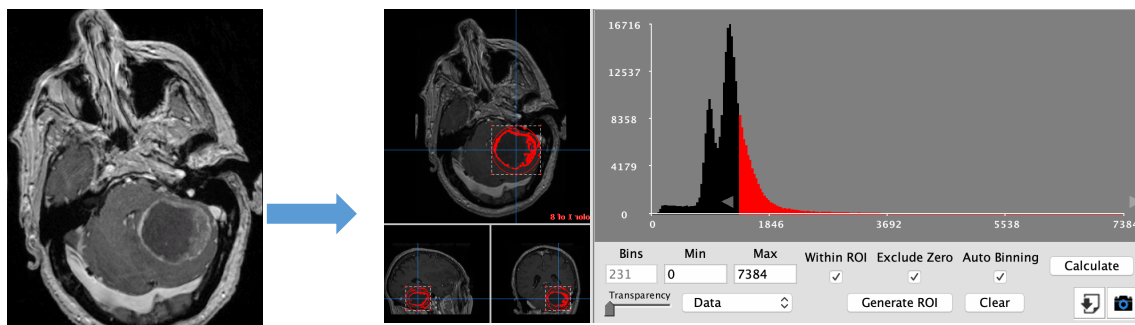


Figure 14 : Interface du logiciel Mango, avec lequel la segmentation a été réalisée. Une première ROI est définie par un cercle coloré rouge sur la séquence 3D T1 injectée, puis un histogramme d'intensités est construit permettant d'isoler la lésion du reste du cerveau. L'opération finale consiste à vérifier manuellement si toute la lésion est bien prise en compte et seulement la lésion, avant de sauvegarder la ROI finale.

IV. Pré-traitement des images

Le but du pré-traitement des images est de diminuer la variabilité inter-acquisitions et l'effet du bruit sur l'étape d'extraction des index radiomiques des régions d'intérêt. En particulier, l'harmonisation des images permet d'homogénéiser à la fois la taille des voxels et les différences d'intensités observables entre différentes machines ou protocoles d'acquisition.

1. Correction de l'inhomogénéité du champ magnétique

Toutes les images ont d'abord été corrigées car artefactées par un bruit basse fréquence présent dans les données IRM, appelé « bias field » (figure 15). Ce bruit, généré par la non-homogénéité du champ magnétique, s'ajoute au vrai signal contenu dans l'image et modifie les valeurs des intensités des voxels de l'image.

Cet artefact dégrade ainsi les performances des algorithmes de traitement d'images tels que les algorithmes de segmentation et de classification ou tout algorithme basé sur l'hypothèse d'une invariance spatiale de l'image traitée [68].

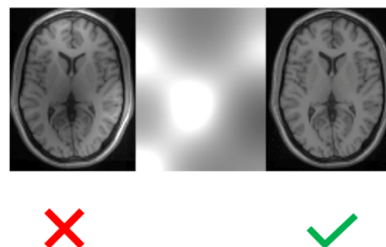


Figure 15: à gauche : image avec bias field et, à droite, image corrigée)[69]

Dans ce travail, cette correction a été effectuée à l'aide de l'algorithme N4ITK [70]. Il permet la correction de la non-uniformité basse fréquence des intensités présentes dans l'image RM. Cette méthode suppose un modèle paramétrique simple et ne nécessite pas de classification des tissus. Cet algorithme est implémenté dans la bibliothèque ANTs (Advanced Normalization Tools) [71] qui représente une bibliothèque populaire en imagerie médicale et qui permet entre autres le recalage d'images ou bien la segmentation.

2. Skull-stripping

Ensuite, afin de mettre en place les méthodes d'harmonisation des intensités, la boîte crânienne a été supprimée des images pour ne garder que le tissu cérébral en utilisant l'outil HD-BET du logiciel FSL (FMRIB's Software Library) [72]. Cet outil est un outil d'extraction du cerveau s'appuyant sur des réseaux de neurones convolutionnels.

3. Z-Score

Enfin, la standardisation des intensités de chacune des IRM a été réalisée. La méthode dite du "Z-Score" a été utilisée à cet effet. Pour cela, l'intensité moyenne à l'intérieur du masque cérébral (μ_{brain}) a été calculée puis soustraite à chaque valeur de voxel $I(x)$. La résultante de chaque voxel a finalement été divisée par l'écart-type des intensités dans le masque (σ_{brain}) (équation 1).

Équation 1: Formule du Z-Score

4. Ré-échantillonnage spatial

Préalablement à l'extraction des index radiomiques, les images ont été ré-échantillonnées dans l'espace pour obtenir des voxels de 1 mm x 1 mm x 1 mm [73][74]. Pour ce faire, une interpolation par spline du 3ème ordre a été utilisée. Les splines sont des fonctions polynomiales par morceaux souvent utilisées pour la modification de l'échantillonnage spatial.

V. Extraction des index radiomiques

Les index radiomiques quantifient différentes informations. Ils comprennent les **index de forme, d'intensité et de texture**.

Le logiciel libre Pyradiomics (version 2.1.2) a été utilisé pour extraire ces caractéristiques [75].

Dans ce logiciel, les caractéristiques radiomiques sont extraites conformément aux définitions consensuelles de l'initiative de normalisation des biomarqueurs d'imagerie (IBSI)[76].

L'IBSI est une collaboration internationale indépendante visant à l'harmonisation de l'extraction de biomarqueurs d'imagerie (radiomics) en fournissant nomenclature, définition et méthodes de calcul entre autres.

1. Index de forme

Les index de forme décrivent de manière avancée et quantitative la géométrie du volume d'intérêt. Ils permettent notamment d'étudier sa surface, son volume, sa divergence par rapport à une sphère ou à une ellipsoïde, sa concavité, sa convexité ou sa sphéricité.

Dans notre étude, **14** caractéristiques basées sur la **forme** ont été extraites (tableau 1).

Order of features	Description	Features name
Shape	Shape-based (3D)	<i>Mesh Volume</i>
		<i>Voxel Volume</i>
		<i>Surface Area</i>
		<i>Surface Area to Volume Ratio</i>
		<i>Sphericity</i>
		<i>Maximum 3D diameter</i>
		<i>Maximum 2D diameter (slice)</i>
		<i>Maximum 2D diameter (column)</i>
		<i>Maximum 2D diameter (row)</i>
		<i>Major Axis Length</i>
		<i>Minor Axis Length</i>
		<i>Least Axis Length</i>
		<i>Elongation</i>
		<i>Flatness</i>

Tableau 1 : Index de Formes

2. Index d'intensité (ou index agnostiques de 1er ordre)

Les index d'intensité sont dérivés de l'histogramme des intensités. Un histogramme décrit le nombre de voxels possédant une intensité donnée, mais sans tenir compte des interactions spatiales entre les voxels.

La moyenne, la médiane, le minimum, le maximum, l'entropie, l'uniformité, l'asymétrie (asymétrie) ou l'aplatissement (kurtosis) de l'histogramme constituent des index d'intensité (figure 17).

C'est pourquoi différents arrangements spatiaux de niveaux de gris correspondant au même histogramme peuvent entraîner des valeurs d'index identiques [77].

Dans notre étude, **18** caractéristiques de **premier ordre** ont été extraites (tableau 2).

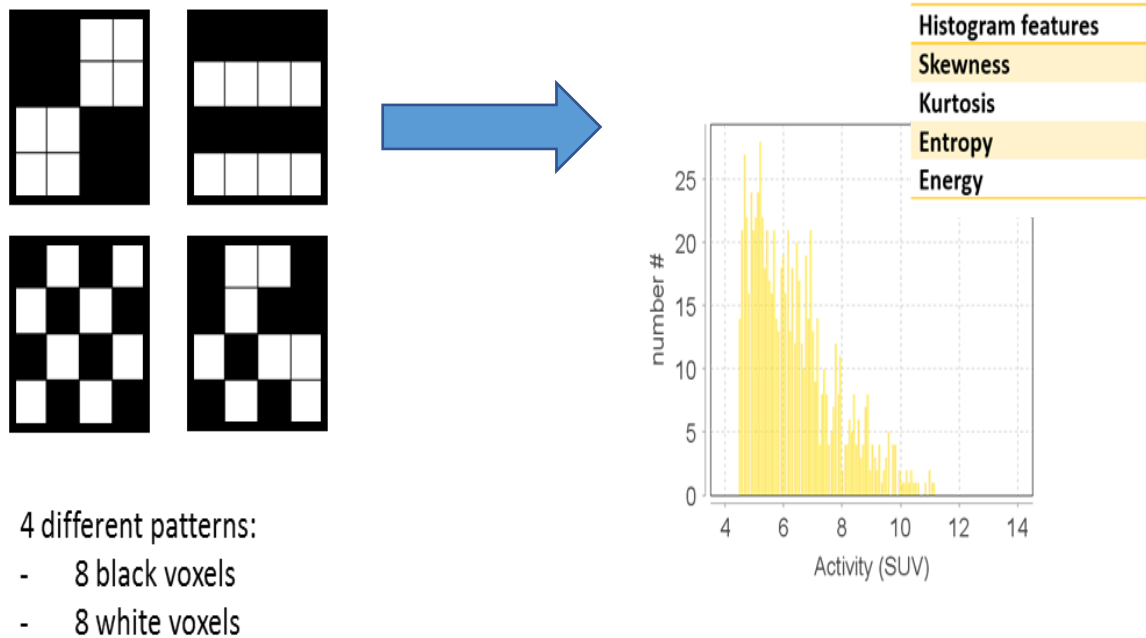


Figure 16: Description d'un histogramme d'intensité

Les index du 1er ordre reflètent uniquement la distribution des intensités, sans tenir compte de la répartition spatiale des voxels. Pour différentes régions possédant le même nombre de pixels noirs et blancs, mais à des positions différentes, l'histogramme est identique. Les valeurs des index dérivés de l'histogramme sont donc égales.[78]

Order of features	Description	Features name
First	First order statistics	<i>Energy</i> <i>Total Energy</i> <i>Entropy</i> <i>Minimum</i> <i>10th percentile</i> <i>90th percentile</i> <i>Maximum</i> <i>Mean</i> <i>Median</i> <i>Interquartile Range</i> <i>Range</i> <i>Mean Absolute Deviation (MAD)</i> <i>Robust Mean Absolute Deviation (rMAD)</i> <i>Root Mean Squared (RMS)</i> <i>Standard Deviation</i> <i>Skewness</i> <i>Kurtosis</i> <i>Variance</i> <i>Uniformity</i>

Tableau 2 : Index de 1er ordre

3. Index de texture (ou index agnostiques de 2ème ordre)

Les index de texture sont dérivés de matrices de niveaux de gris qui décrivent les arrangements spatiaux des voxels dans le volume d'intérêt en fonction de leurs niveaux de gris, et permettent ainsi de décrire l'hétérogénéité tumorale.

Dans notre étude, **73** caractéristiques de **texture** ont été extraites des régions tumorales segmentées à partir de **5 classes** de caractéristiques de texture.

a. La matrice de co-occurrence

- GLCM : Gray Level Co-occurrence Matrix
- **22** caractéristiques extraites
- Décrit la probabilité d'observer un couple de valeurs de niveaux de gris à une distance donnée [78]. En 3D, 13 directions sont nécessaires pour connecter un voxel avec ses 26 voisins et sont donc calculées (figure 18).
- Une matrice de co occurrence est calculée pour chaque direction.
- Sur chaque matrice, 22 index sont extraits. Chaque caractéristique finale correspond à la moyenne des index évaluée à partir de chacune des 13 matrices (figure 19).

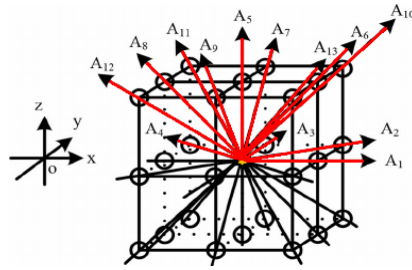


Figure 17: Schéma des 13 directions nécessaires pour connecter un voxel avec ses 26 voisins

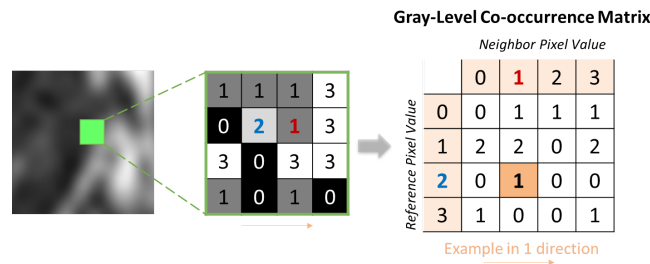


Figure 18: Matrice de Co-occurrence des niveaux de gris (GLCM)[76]

b. La matrice des longueurs de séries homogènes

- GLRLM : Grey Level Run-Length Matrix
- **16** caractéristiques extraites
- Mesure le nombre de voxels consécutifs possédant la même valeur d'intensité, alignés dans une direction donnée pour chaque valeur d'intensité [79].

c. La matrice des tailles de zones homogènes

- GLSZM : Grey Level Size-Zone Matrix
- **16** caractéristiques extraites
- Mesure le nombre de voxels voisins de même intensité, pour chaque valeur d'intensité [80].

d. La matrice de différence des niveaux de gris entre voxels voisins

- NGDLM : Neighborhood Grey Level Difference Matrix
- **5** caractéristiques extraites
- Mesure la différence des intensités entre des voxels voisins [78].

e. La matrice de dépendance des niveaux de gris

- GLDM : Gray Level Dependence Matrix
- **14** caractéristiques extraites (tableau 3)
- Quantifie les dépendances des niveaux de gris dans une image :

Une dépendance de niveaux de gris correspond au nombre de voxels connectés à une distance donnée qui sont dépendants du voxel central.

Order of features	Description	Features name
Second	Grey level co-occurrence matrix (GLCM)	<i>Autocorrelation</i>
		<i>Joint Average</i>
		<i>Cluster Prominence</i>
		<i>Cluster Shade</i>
		<i>Cluster Tendency</i>
		<i>Contrast</i>
		<i>Correlation</i>
		<i>Difference Average</i>
		<i>Difference Entropy</i>
		<i>Difference Variance</i>
		<i>Joint Energy</i>
		<i>Joint Entropy</i>
		<i>Informational Measure of Correlation (IMC) 1</i>
		<i>Informational Measure of Correlation (IMC) 2</i>
		<i>Inverse Difference Moment (IDM)</i>
		<i>Maximal Correlation Coefficient (MCC)</i>
		<i>Inverse Difference Moment Normalized (IDMN)</i>
		<i>Inverse Difference (ID)</i>
		<i>Inverse Difference Normalized (IDN)</i>
		<i>Inverse Variance</i>
<i>Maximum Probability</i>		
<i>Sum Entropy</i>		
<i>Sum of Squares</i>		
Third	Grey level size zone matrix (GLSZM)	<i>Small Area Emphasis (SAE)</i>
		<i>Large Area Emphasis (LAE)</i>
		<i>Grey Level Non-Uniformity (GLN)</i>
		<i>Grey Level Non-Uniformity Normalized (GLNN)</i>
		<i>Size-Zone Non-Uniformity (SZN)</i>
		<i>Size-Zone Non-Uniformity Normalized (SZNN)</i>
		<i>Zone Percentage (ZP)</i>
		<i>Grey Level Variance (GLV)</i>
		<i>Zone Variance (ZV)</i>
		<i>Zone Entropy (ZE)</i>
		<i>Low Grey Level Zone Emphasis (LGLZE)</i>
		<i>High Grey Level Zone Emphasis (HGLZE)</i>
		<i>Small Area Low Grey Level Emphasis (SALGLE)</i>
		<i>Small Area High Grey Level Emphasis (SAHGLE)</i>
		<i>Large Area Low Grey Level Emphasis (LALGLE)</i>
<i>Large Area High Grey Level Emphasis (LAHGLE)</i>		
Third	Grey level run length matrix (GLRLM)	<i>Short Run Emphasis (SRE)</i>
		<i>Long Run Emphasis (LRE)</i>
		<i>Grey Level Non-Uniformity (GLN)</i>
		<i>Grey Level Non-Uniformity Normalized (GLNN)</i>
		<i>Run Length Non-Uniformity (RLN)</i>
		<i>Run Length Non-Uniformity Normalized (RLNN)</i>
		<i>Run Percentage (RP)</i>
		<i>Grey Level Variance (GLV)</i>
		<i>Run Variance (RV)</i>
		<i>Run Entropy (RE)</i>
		<i>Low Grey Level Run Emphasis (LGLRE)</i>
		<i>High Grey Level Run Emphasis (HGLRE)</i>
		<i>Short Run Low Grey Level Emphasis (SRLGLE)</i>
		<i>Short Run High Grey Level Emphasis (SRHGLE)</i>
		<i>Long Run Low Grey Level Emphasis (LRHGLE)</i>
<i>Long Run High Grey Level Emphasis (LRHGLE)</i>		
Third	Neighbouring grey tone difference matrix (NGTDM)	<i>Coarseness</i>
		<i>Contrast</i>
		<i>Busyness</i>
		<i>Complexity</i>
		<i>Strength</i>
Third	Grey level dependence matrix (GLDM)	<i>Small Dependence Emphasis (SDE)</i>
		<i>Large Dependence Emphasis (LDE)</i>
		<i>Grey Level Non-Uniformity (GLN)</i>
		<i>Dependence Non-Uniformity (DN)</i>
		<i>Dependence Non-Uniformity Normalized (DNN)</i>
		<i>Grey Level Variance (GLV)</i>
		<i>Dependence Variance (DV)</i>
		<i>Dependence Entropy (DV)</i>
		<i>Low Grey Level Emphasis (LGLE)</i>
		<i>High Grey Level Emphasis (HGLE)</i>
<i>Small Dependence Low Grey Level Emphasis (SDLGLE)</i>		
<i>Small Dependence High Grey Level Emphasis (SDHGLE)</i>		
<i>Large Dependence Low Grey Level Emphasis (LDLGLE)</i>		
<i>Large Dependence High Grey Level Emphasis (LDHGLE)</i>		

Tableau 3: Index de texture

La figure 20 résume les matrices de texture usuelles et les index de texture dérivés.

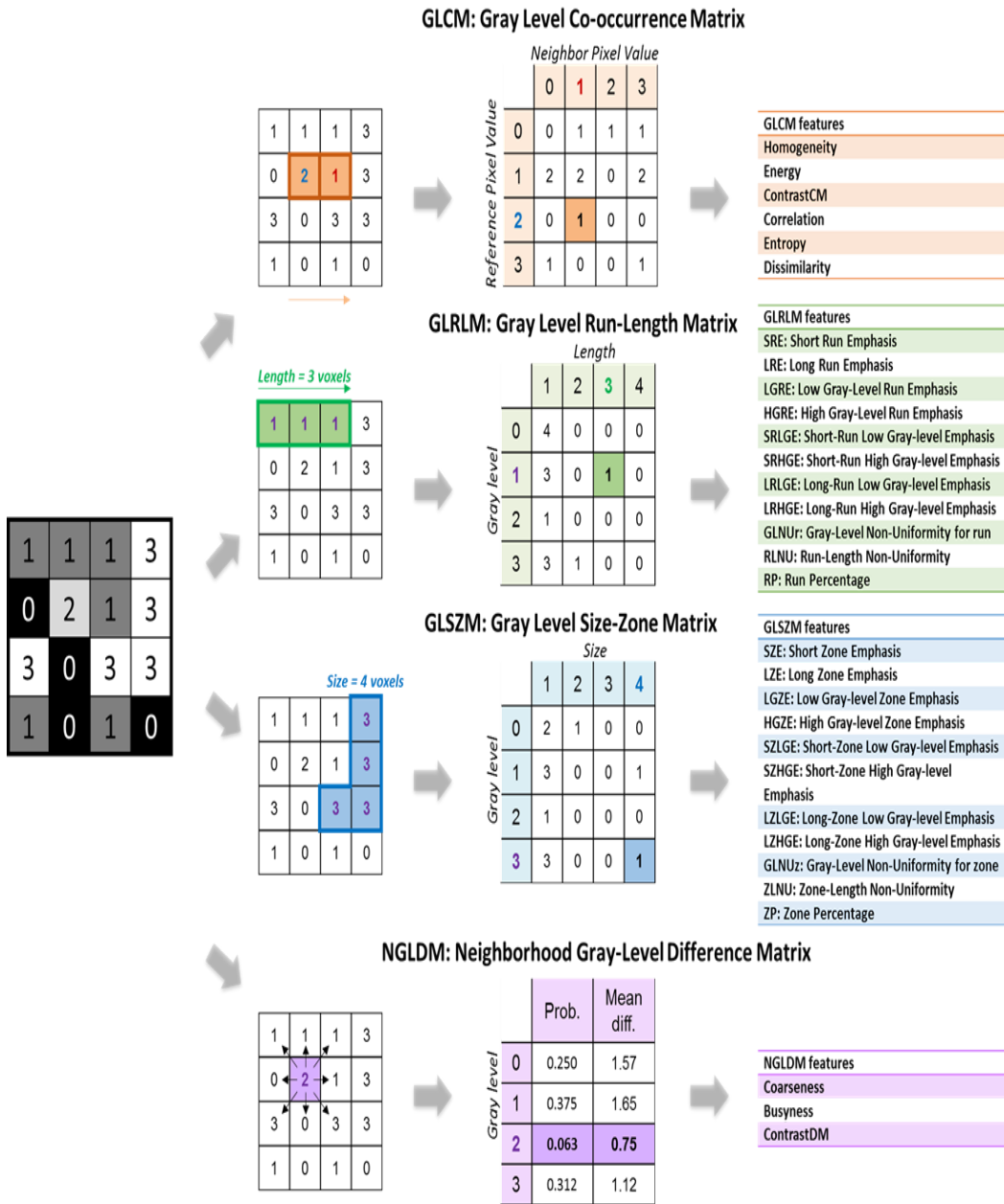


Figure 19 : Matrices de texture usuelles et index de texture dérivés [78]

VI. Information spatiale

Afin d'obtenir l'information spatiale sur la localisation de la tumeur, toutes les images ont été recalées sur l'atlas de MNI (Montreal Neurological Institute), en utilisant la bibliothèque ANTs. Pour ce faire, des transformations affines incluant 12 degrés de liberté ont été considérées.

L'atlas MNI représente une version du cerveau moyen proposée par l'ICBM (International Consortium of Brain Mapping). Dans ce projet, une moyenne de 152 IRM pondérées T1 a été considérée [81][82].

Les coordonnées 3D du centre de masse de la tumeur dans le référentiel de l'atlas du MNI ont été incluses en tant que caractéristiques dans le modèle de classification tumorale.

VII. Analyse statistique

1. Généralités

Une fois qu'un sous-ensemble de caractéristiques importantes est identifié, il est possible de générer un modèle mathématique qui prédit une classe (tâche de classification) ou une valeur continue (tâche de régression) à partir de ces grandeurs. C'est le principe du machine-learning.

Les algorithmes d'apprentissage machine couramment utilisés pour la génération de modèles en radiomique sont les arbres de décision (par exemple, les random forest), la régression linéaire ou logistique, les machines à vecteurs de support (SVM pour Support Vector Machines) et la méthode des k plus proches voisins (k-nearest neighbors).

Les paramètres de ces algorithmes sont optimisés pour répondre à une tâche donnée dans un sous-ensemble de données (ensemble de données d'apprentissage). Les méthodes d'apprentissage supervisées apprennent notamment à prédire une variable à partir de données annotées sur une cohorte d'apprentissage. Dans notre étude, le logiciel doit apprendre à différencier les sous types de tumeurs cérébrales métastatiques, à partir d'une cohorte constituée des différentes métastases cérébrales de cancer pulmonaire, de cancer mammaire, de mélanome et de cancer colorectal.

Ensuite, afin d'évaluer que le modèle ne souffre pas de sur-apprentissage, le modèle est appliqué à un autre sous-ensemble de données qui n'a pas été utilisé pendant le processus de génération du modèle (ensemble de données de validation).

Idéalement, le modèle identifié comme le plus performant est finalement appliqué à un troisième ensemble de données (ensemble de données de test, données totalement indépendantes, en aveugle) comprenant des données d'imagerie acquises auprès de différentes institutions à l'aide de différents IRM et de différents protocoles d'acquisition afin d'évaluer son caractère généralisable.

Conventionnellement, environ 60 % des données servent pour l'apprentissage, 20 % pour la validation et 20% pour le test quand un jeu de données indépendant n'est pas accessible.

Dans les cas où le nombre de patients est faible et où il n'est pas possible d'effectuer une division raisonnable et équilibrée des données en une cohorte d'apprentissage/validation et une cohorte de test avant la génération du modèle, des méthodes statistiques telles que la validation croisée peuvent être appliquées pour estimer la performance du modèle sans disposer d'un ensemble de données de test indépendant. Les ensembles de données disponibles sont ainsi divisés en k sous-ensembles de taille égale, et un sous-ensemble est conservé comme données de validation, tandis que les $k-1$ ensembles de données restants sont utilisés comme données d'apprentissage. Ensuite, le processus est répété k fois, chaque sous-ensemble étant utilisé une fois comme données de validation. Les estimateurs des performances du modèle de chaque k itération peuvent alors être moyennés pour produire une seule estimation de la performance du modèle [83]. Nous avons réalisé cette démarche dans notre étude (figure 21).



Figure 20 : Illustration de la méthodologie de la validation croisée

2. Analyse statistique

Quatre classes de métastases cérébrales ont été considérées selon le site d'origine primaire.

Un total de **74, 63, 59, 41** lésions ont été incluses respectivement pour le mélanome, le cancer primaire du poumon, du sein et du colorectal.

80% des lésions ont été **utilisées** pour l'entraînement et la validation du modèle par méthode de validation croisée comme détaillé préalablement.

20% du nombre total de lésions par classe ont été mises de côté pour constituer le jeu de test indépendant, qui ne sera pas étudié ici.

La répartition était stratifiée selon un ratio égalitaire de pourcentages de métastases de sein, de mélanome, de poumon et de cancer colorectal entre les différents sets de la validation croisée (80% des lésions) et le jeu de test indépendant.

Ensuite, deux stratégies de classification ont été testées :

- **One vs One** : opposition d'un type de néoplasie primitive par rapport à une autre (i.e. sein versus poumon)
- **One vs Others** : opposition d'un type de néoplasie primitive par rapport à toutes les autres (i.e. sein versus toutes les autres étiologies)

4 types de modèles ont été finalement développés. Ces modèles intégraient :

- Une combinaison de toutes les caractéristiques,
- Uniquement les coordonnées spatiales de la localisation de la tumeur,
- Uniquement les caractéristiques cliniques (âge, sexe),
- Uniquement les caractéristiques radiomiques.

3. Ré-échantillonnage des populations

Nous avons affaire à un problème de classes déséquilibrées : nous avons par exemple presque 2 fois plus de plus de lésions étudiées secondaires à un mélanome (74 lésions) que de lésions colorectales (41 lésions).

Ce déséquilibre de classe augmente nettement la difficulté de l'apprentissage par l'algorithme de classification et rend complexe l'évaluation de ses performances. En effet, l'algorithme a moins d'exemples de la classe minoritaire. Il est donc biaisé sur cette population minoritaire et produit des prédictions potentiellement moins robustes qu'en l'absence de déséquilibre. Les performances peuvent être par ailleurs faussement très bonnes si l'algorithme classe l'ensemble des lésions comme faisant partie de la classe majoritaire.

Pour pallier à ce déséquilibre, nous avons utilisé 3 techniques de ré-échantillonnage :

Sous-échantillonnage aléatoire :

Sous-échantillonner la ou les classes majoritaires en supprimant au hasard des échantillons (ici des lésions)

Sur-échantillonnage aléatoire :

Sur-échantillonnage de la ou des classes minoritaires en dupliquant au hasard des échantillons

SMOTE (Technique de sur-échantillonnage synthétique des minorités) :

SMOTE est une technique de sur-échantillonnage qui consiste à synthétiser de nouveaux échantillons dans la classe minoritaire, c'est-à-dire en la création de profils cliniques, spatiaux

et/ou radiomiques statistiquement probables pour la ou les classes minoritaires mais ne correspondant pas à de vraies lésions.

4. Principales mesures de performance

Plusieurs outils sont à notre disposition pour évaluer les performances du modèle généré. Nous allons dans un premier temps redéfinir les termes de *precision*, *recall* et *accuracy* dont nous aurons besoin pour expliciter ce qu'est le F1 score et l'aire sous la courbe ROC (roc-AUC pour Area Under the Receiver Operating Characteristic Curve).

La *precision* compte la proportion d'items pertinents parmi les items sélectionnés, représentant donc la valeur prédictive positive (équation 2).

Le *recall* compte la proportion d'items pertinents sélectionnés parmi tous les items pertinents sélectionnables, représentant donc la sensibilité (équation 3).

Équation 2: Formule de la precision

Équation 3: Formule du recall

L'*accuracy* désigne le rapport entre le nombre de prévisions correctes et le nombre total d'échantillons d'entrée (équation 4).

Équation 4: Formule de l'accuracy

avec TP : True Positive : Vrai Positif, TN : True Negative : Vrai Négatif, FP : False Positive : Faux Positif, FN : False Negative : Faux Négatif

Le score F1 est une mesure globale de la précision d'un modèle qui combine précision et rappel (équation 5). Il s'agit d'une métrique bien adaptée aux configurations déséquilibrées comme c'est le cas dans notre étude.

Un bon score F1 signifie que l'on a peu de faux positifs et peu de faux négatifs. Un score F1 est considéré comme parfait lorsqu'il est de 1, alors que le modèle est un échec total lorsqu'il est de 0.

Équation 5: Formule du score F1

Une courbe ROC est un graphique en 2 dimensions qui représente le taux de vrais positifs (sensibilité) en fonction du taux de faux positifs ($1 - \text{spécificité}$) pour différents seuils de classification d'un modèle analysé. Le **score roc-AUC** mesure l'aire sous la courbe de ce graphique. Il s'agit d'une mesure de la performance d'un classificateur binaire, c'est-à-dire d'un système qui a pour objectif de catégoriser des éléments en deux groupes distincts sur la base d'une ou plusieurs des caractéristiques de chacun de ces éléments. Le **score roc-AUC** est biaisé dans le cas de classes déséquilibrées, d'où l'intérêt du score F1 dans notre étude.

Un **score roc-AUC** de 0,5 suggère l'absence de discrimination (performances aléatoires) alors qu'un roc-AUC de 0,7 à 0,8 est considéré comme acceptable, de 0,8 à 0,9 comme excellent et supérieur à 0,9 comme exceptionnel [84].

5. Modèles testés

L'objectif du Machine Learning est d'apprendre au modèle à prédire des données de sorties (ici primitif associé à chaque métastase cérébrale) à partir de données d'entrée (ici les informations radiomiques, spatiaux, cliniques extraits), et à sélectionner le modèle le plus performant.

Dans ce travail, **648 combinaisons de modèles** ont été testées à partir d'une méthodologie de validation croisée. Chaque combinaison incluait un classifieur (au nombre de **18**; tableau 4), une méthode de mise à l'échelle des index radiomiques (au nombre **9**), et une méthode de rééquilibrage des classes (au nombre de $3 + 1 = 4$).

L'étape de « mise à l'échelle » des index radiomiques consiste à ramener les différentes caractéristiques radiomiques, initialement dans des gammes de valeurs très différentes, dans une plage de valeurs similaires. Dans ce travail, 9 méthodes de mises à l'échelle ont été testées (StandardScaler, MinMaxScaler, MaxAbsScaler, RobustScaler, Quantile Transformer-Normal, Quantile Transformer-Uniform, PowerTransformer-Yeo-Johnson, Normalizer, pas de mise à l'échelle).

Concernant, les méthodes de rééquilibrage des classes, le nombre 4 provient du fait qu'une solution est de ne pas chercher à ré-équilibrer les classes dans l'étape d'apprentissage. Dans ce cas, les performances sur le jeu de données d'apprentissage doivent être analysées avec méfiance.

QuadraticDiscriminantAnalysis
AdaBoostClassifier
BaggingClassifier
ExtraTreesClassifier
GradientBoostingClassifier
RandomForestClassifier
DecisionTreeClassifier
ExtraTreeClassifier
RidgeClassifier
SGDClassifier
BernoulliNB
GaussianNB
KNeighborsClassifier
MLPClassifier
LinearSVC
NuSVC
SVC
LinearDiscriminantAnalysis

Tableau 4 : Classifieurs testés pour la construction des modèles.

6. Stratégie One versus One

La stratégie One vs One consiste à opposer une classe par rapport à une autre, c'est-à-dire, sur les 80 % de métastases de chaque groupe de primitifs : sein vs poumon (47 vs 50), colique vs sein (33 vs 47), sein vs mélanome (47 vs 59), colique vs poumon (33 vs 50), poumon vs mélanome (50 vs 59), colique vs mélanome (33 vs 59), pour juger de la présence de signal ou non dans les caractéristiques d'entrée du modèle.

Pour évaluer les performances du modèle, la *precision*, le *recall*, l'*accuracy*, le score roc-AUC et le score F1 ont été évalués en réalisant une moyenne des performances obtenues sur les jeux d'apprentissage et sur les 5 jeux de données restants de la validation croisée.

7. Stratégie One versus Others

La stratégie One vs. Others consiste à opposer une classe par rapport à toutes les autres, c'est-à-dire, sur les 80 % de métastases de chaque groupe de primitifs : sein contre autres classes (47 contre 142), poumon contre autres classes (50 contre 139), mélanome contre autres classes (59 contre 130), coliques contre autres classes (33 contre 156), et est plus représentative de la question clinique à laquelle on souhaite répondre.

La même méthodologie d'évaluation a été adoptée que le cadre de la stratégie One vs One.

RESULTATS

I. Population

La population étudiée comprenait 188 patients (âge moyen \pm écart type, 58,75 \pm 12,07 ans), dont 75 hommes (39,9%) et 113 femmes (60,1%).

Un total de **237** métastases cérébrales ont été incluses.

157 patients avaient une métastase cérébrale unique, 31 présentaient des lésions métastatiques multiples de plus de 1 cm (19 patients présentaient 2 lésions, 11 patients moins de 5, 1 patient avec 5 lésions).

Les métastases cérébrales provenant d'un cancer du sein représentaient 59/237 (24,9 %) de la population, la néoplasie pulmonaire 63/237 (26,6 %), le mélanome 74/237 (31,2 %) et la néoplasie colorectale 41/237 (17,3 %).

La population de patients présentant des métastases de mélanome provenait de **46** patients dont 27 femmes et 19 hommes, soit **58,7 %** de femmes, avec un âge moyen de **55,37** ans avec un écart type de 13,20 ans.

La population de patients présentant des métastases pulmonaires provenait de **57** patients dont 27 femmes et de 30 hommes soit **47,37 %** de femmes, avec un âge moyen de **61,56** ans avec un écart type de 9,67 ans.

La population de patients présentant des métastases mammaires provenait de **48** patients dont 44 femmes et de 4 hommes, soit **91,67 %** de femmes, avec un âge moyen de **52,21** ans avec un écart type de 12,77 ans.

La population de patients présentant des métastases de cancer colorectal provenait de **37** patients dont 15 femmes et de 22 hommes, soit **40,54 %** de femmes, avec un âge moyen de **63,32** ans avec un écart type de 7,15 ans.

II. Résultats de la stratégie One versus One

Le tableau 5 synthétise les performances de classification pour la stratégie One vs One (sein vs poumon, sein vs colique, sein vs mélanome, poumon vs colique, poumon vs mélanome, mélanome vs colique) selon les 4 approches :

- Une combinaison de toutes les caractéristiques,
- Uniquement les coordonnées spatiales de la localisation de la tumeur,
- Uniquement les caractéristiques cliniques (âge, sexe),
- Uniquement les caractéristiques radiomiques.

Les résultats indiqués correspondent à la moyenne des scores roc-AUC et F1, évalués sur les jeux de données restants de la validation croisée (validation) (avec en couleur rouge, les scores roc-AUC et F1 aux alentours de 0,6, en jaune, les scores roc-AUC et F1 aux alentours de 0,7 et en vert, le score roc-AUC et F1 aux alentours de 0,8 et plus).

Pour rappel, seules 80 % des métastases de chaque groupe de primitif (poumon, sein, mélanome, colorectal) ont été utilisées.

L'analyse complète des résultats, détaillant les classifieurs, est résumée en annexe 1.

L'analyse complète des résultats portant sur la *precision*, le *recall*, l'*accuracy* et l'apprentissage des scores roc-AUC et F1 est détaillée aux annexes 2, 3, 4 et 5.

Classification One versus One	Informations	Moyenne Validation roc-AUC	Moyenne Validation F1
Mammaire Vs Pulmonaire 47 VS 50	Modèle complet	0,824245443	0,735315072
	Informations Spatiales	0,633452758	0,628090385
	Informations Cliniques	0,781035625	0,689524048
	Informations Radiomiques	0,664732556	0,68506242
Mammaire Vs Colique 47 VS 33	Modèle complet	0,770347791	0,84626807
	Informations Spatiales	0,551681291	0,717147023
	Informations Cliniques	0,897229056	0,846683528
	Informations Radiomiques	0,639110736	0,731212755
Mammaire Vs Mélanome 47 VS 59	Modèle complet	0,787002446	0,748116042
	Informations Spatiales	0,665078048	0,598875185
	Informations Cliniques	0,680674367	0,706490867
	Informations Radiomiques	0,764059956	0,732116515
Pulmonaire Vs Colique 50 VS 33	Modèle complet	0,664585123	0,747341182
	Informations Spatiales	0,514792277	0,718220976
	Informations Cliniques	0,59902501	0,747341182
	Informations Radiomiques	0,64946115	0,747341182
Pulmonaire Vs Mélanome 50 VS 59	Modèle complet	0,787774118	0,711373673
	Informations Spatiales	0,570676171	0,538526411
	Informations Cliniques	0,749529401	0,707526556
	Informations Radiomiques	0,780560591	0,698796843
Mélanome Vs Colique 59 VS 33	Modèle complet	0,820278832	0,850763093
	Informations Spatiales	0,553189366	0,778074301
	Informations Cliniques	0,741416833	0,784038953
	Informations Radiomiques	0,790530969	0,842251797

Tableau 5: Performances de Classification pour la stratégie One vs One

1. Métastases d'origine mammaire versus pulmonaire

Comme illustré dans le tableau 5, les données de radiomique ne permettent pas de différencier clairement ces types lésionnels. En effet, les valeurs de roc-AUC et de F1 sont faibles. La valeur du score roc-AUC vaut $0,66 \pm 0,10$ sur les jeux de validation et le score F1, $0,68 \pm 0,06$.

Les données spatiales recueillies n'apportent pas non plus d'information supplémentaire avec respectivement un score roc-AUC de $0,63 \pm 0,11$ et un score F1 de $0,63 \pm 0,10$.

En revanche, ainsi qu'attendu, il existe une forte pondération du genre qui permet aux données cliniques d'être significatives, la classe du genre féminin étant plus représentée au sein des lésions métastatiques mammaires que pulmonaires. Ainsi le roc-AUC atteint $0,78 \pm 0,12$ et le score F1 $0,69 \pm 0,11$.

Le modèle global généré obtient des performances assez bonnes, mais cela par la forte pondération du genre sur ces deux populations.

2. Métastases d'origine mammaire versus colique

Ici à nouveau, ni les caractéristiques radiomiques [score roc-AUC égal à $0,64 \pm 0,14$ et score F1 égal à $0,73 \pm 0,11$] ni les données spatiales [score roc-AUC égal à $0,55 \pm 0,12$ et score F1 égal à $0,72 \pm 0,09$] ne permettent de différencier significativement les deux types lésionnels.

La clinique, de par le sexe des patients différent au sein des deux groupes, est le seul facteur qui l'emporte sur le modèle général.

3. Métastases d'origine mammaire versus mélanome

Cette fois, un signal franc issu de l'analyse radiomique permet de séparer les deux types néoplasiques. La classification des métastases d'origine mammaire par rapport aux mélanomes est guidée par la radiomique avec un score roc-AUC de $0,76 \pm 0,12$ et un score F1 de $0,73 \pm 0,11$.

Les performances sont un peu améliorées en combinant l'ensemble des données avec un roc-AUC de $0,79 \pm 0,09$ et un score F1 de $0,75 \pm 0,11$.

L'information clinique guide de façon modérée le modèle avec un score roc-AUC de $0,68 \pm 0,07$ et un score F1 de $0,71 \pm 0,10$.

Les données spatiales orientent peu la classification avec respectivement, des scores roc-AUC et F1 de $0,67 \pm 0,1$ et $0,60 \pm 0,12$.

4. Métastase d'origine pulmonaire versus colique

Aucun modèle n'est particulièrement discriminant que ce soit l'analyse radiomique, clinique, spatiale ou le modèle global.

La classification des métastases d'origine pulmonaire par rapport aux coliques est ainsi faiblement guidée par la radiomique avec un score roc-AUC de $0,65 \pm 0,15$ et un score F1 de $0,75 \pm 0,08$.

Les données spatiales et cliniques orientent peu la classification avec respectivement, des scores roc-AUC et F1 de $0,52 \pm 0,21$ et $0,72 \pm 0,14$ pour les données spatiales et $0,60 \pm 0,10$ et $0,75 \pm 0,08$ pour les données cliniques.

Le score est peu modifié en combinant l'ensemble des données avec un score roc-AUC de $0,66 \pm 0,16$ et un score F1 de $0,75 \pm 0,08$.

5. Métastase d'origine mélanome versus colique

La classification des métastases d'origine de mélanome par rapport aux coliques est ici clairement guidée par la radiomique avec des performances élevées : roc-AUC de $0,79 \pm 0,13$ et score F1 de $0,84 \pm 0,06$.

Les données spatiales orientent peu la classification avec respectivement, des scores roc-AUC et F1 de $0,55 \pm 0,16$ et $0,78 \pm 0,07$. Les données cliniques guident également peu la classification avec un score roc-AUC de $0,74 \pm 0,10$ et un score F1 de $0,78 \pm 0,07$.

Le score global est très pondéré par les données radiomiques et varie peu en termes de performances entre radiomique seule et ensemble des données combinées. Pour le modèle intégrant l'ensemble des données, le score roc-AUC vaut $0,82 \pm 0,11$ et le score F1 de $0,85 \pm 0,07$.

III. Résultats de la stratégie One versus Others

Le tableau 6 synthétise les performances de classification pour la stratégie One vs Others selon les 4 approches :

- Une combinaison de toutes les caractéristiques,
- Uniquement les coordonnées spatiales de la localisation de la tumeur,
- Uniquement les caractéristiques cliniques (âge, sexe),
- Uniquement les caractéristiques radiomiques.

Les résultats correspondent à la moyenne des scores roc-AUC et F1, évalués sur les jeux de données restants de la validation croisée (avec en couleur rouge, les scores roc-AUC et F1 aux alentours de 0,6, en jaune, les scores roc-AUC et F1 aux alentours de 0,7 et en vert, le score roc-AUC et F1 aux alentours de 0,8 et plus).

Pour rappel, seules 80 % des métastases de chaque groupe de primitif (poumon, sein, mélanome, colorectal) ont été utilisées.

L'analyse complète des résultats, détaillant les classifieurs, est résumée en annexe 6.

L'analyse complète des résultats portant sur la *precision*, le *recall*, l'*accuracy* et l'apprentissage des scores roc-AUC et F1 est détaillée aux annexes 7,8,9 et 10.

Classification One versus Others	Informations	Moyenne Validation roc-AUC	Moyenne Validation F1
Pulmonaire Vs les Autres 50 VS 139	Modèle complet	0,689429542	0,462603193
	Informations Spatiales	0,587463085	0,43207895
	Informations Cliniques	0,684498376	0,476902167
	Informations Radiomiques	0,639681588	0,452049545
Mélanome Vs les Autres 59 VS 130	Modèle complet	0,754685477	0,847266188
	Informations Spatiales	0,515126747	0,816350366
	Informations Cliniques	0,693568255	0,841336254
	Informations Radiomiques	0,758482778	0,845824702
Mammaire Vs les Autres 47 VS 142	Modèle complet	0,630686287	0,855666614
	Informations Spatiales	0,6424833	0,855666614
	Informations Cliniques	0,783786445	0,862030669
	Informations Radiomiques	0,630686287	0,855666614
Colique Vs les Autres 33 VS 156	Modèle complet	0,799757507	0,910192137
	Informations Spatiales	0,575940274	0,745126917
	Informations Cliniques	0,769453018	0,770028147
	Informations Radiomiques	0,762449449	0,905325445

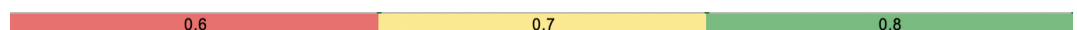


Tableau 6: Performances de Classification pour la stratégie One vs Others

Lors de cette analyse, nous retrouvons que les métastases de mélanome se détachent clairement des autres types métastatiques par leur signature radiomique basée sur l'analyse du volume T1 écho de gradient après injection de gadolinium.

Ainsi, la classification des mélanomes par rapport aux autres tumeurs primitives est guidée par la radiomique avec un roc-AUC de $0,76 \pm 0,08$ et un score F1 de $0,85 \pm 0,04$.

Les données cliniques guident elles aussi modérément la classification avec un score roc-AUC de $0,69 \pm 0,07$ et un score F1 de $0,84 \pm 0,04$.

Les données spatiales sont moins informatives avec, respectivement, des scores roc-AUC et F1 de $0,52 \pm 0,03$ et $0,82 \pm 0,03$.

Les performances du modèle combinant l'ensemble des données sont malgré tout peu améliorées en comparaison au modèle incluant les caractéristiques radiomiques seules puisque pour ce modèle combiné possède un score roc-AUC de $0,75 \pm 0,08$ et un score F1 de $0,85 \pm 0,04$.

Pour les métastases mammaires, le sexe guide la classification avec un score roc-AUC de $0,78 \pm 0,07$ et un score F1 de $0,86 \pm 0,05$.

Pour les métastases coliques vs les autres, nous avons observé de bons scores quand le modèle inclut les index radiomiques, avec des performances roc-AUC de $0,76 \pm 0,12$ et un score F1 de $0,90 \pm 0,03$.

Cependant, ce dernier score est certainement lié au non-équilibre des classes (33 coliques vs. 156 autres) car nous avons noté une absence de signal préalable dans la stratégie one versus one.

DISCUSSION

I. Rappel des principaux résultats et implication clinique

Notre étude a permis de développer et d'évaluer les performances de plusieurs modèles d'analyse radiomique basés sur l'imagerie IRM cérébrale en 3DT1 post injection de gadolinium, dans le but de différencier les métastases cérébrales en fonction de leur primitif. Notre modèle permet de faire ressortir une signature claire des métastases de mélanome par rapport aux autres primitifs testés, à savoir le sein, le poumon et le colon. Ainsi, la classification des mélanomes par rapport aux autres tumeurs primitives est guidée par la radiomique avec un roc-AUC de $0,76 \pm 0,08$ et un score F1 de $0,85 \pm 0,04$ sur les 5 cohortes de validation de la validation croisée.

Jusqu'à 15 % de tous les patients présentant des métastases cérébrales n'ont pas de site primitif clairement détecté malgré une évaluation intensive [13]. Pouvoir s'orienter vers le site primaire, en analysant les métastases cérébrales, aurait un impact important sur les soins aux patients, en aidant à concentrer et à poursuivre les investigations sur une origine. Si nous prenons l'exemple du mélanome, certains endroits cachés sont difficiles à explorer, tels que les ongles, le cuir chevelu, la paume des mains, la plante des pieds, mais peuvent aussi être complètement inaccessibles à l'inspection visuelle, comme la choroïde intra-oculaire ou le mélanome de l'œsophage. Dans ces circonstances, le fait d'avoir une signature qui s'oriente vers un mélanome primaire, conduira à poursuivre des investigations spécifiques sur des endroits cachés afin de trouver le site primaire, et d'adapter la stratégie de traitement.

Pour les autres primitifs en revanche, notre modèle radiomique ou complet comprenant également données cliniques et spatiales, ne permet pas de discriminer une signature orientant vers le primitif.

Concernant les métastases d'origine mammaire, le score global de notre modèle peut paraître assez bon (score F1 de $0,86 \pm 0,05$). Cependant, ce résultat est fortement guidé par le genre féminin surexprimé pour ce cancer. Il n'est donc pas envisageable d'utiliser le modèle global pondéré quasi uniquement par le genre pour distinguer des métastases d'origine mammaire des autres métastases.

II. Comparaison à la littérature

Le tableau 7 fait un bilan de l'ensemble des études récentes ayant proposé un modèle intégrant des variables radiomiques pour la classification de lésions cérébrales métastatiques en fonction du cancer primitif (cancers pulmonaire, mammaire et mélanome au minimum).

ETUDE	Notre étude	Ortiz et al. /85/	Knip et al. /86/
Année	2020	2018	2019
Nombre total de Patients/ Métastases	188 /237	38/67	189/658
Nombre de cas par classe	63 Poumon/ 74 Mélanome/ 59 Sein /41 Colorectal	27 Poumon / 23 Mélanome/ 17 Sein	151 Poumon <u>petites cellules</u> / 225 Poumon <u>non petites cellules</u> / 89 Mélanome/143 Sein/50 Colorectal
Séquence(s) utilisée(s)	3D T1 SPGR post gadolinium	3D T1 IR SPGR post gadolinium	3D T1 GE post gadolinium ou Axial T1 SE post gadolinium / FLAIR/ Axial T1 SE
Segmentation	3D	3D et 2D	3D
Logiciel d'extraction des features	Pyradiomics	MATLAB	Pyradiomics
Features	- 18 caractéristiques de premier ordre - 14 caractéristiques basées sur la forme - 73 caractéristiques de texture GLDM, NGTDM, GLSZM, GLRLM, GLCM	- 3 caractéristiques de premier ordre - 40 caractéristiques de texture : NGTDM, GLSZM, GLRLM, GLCM	- 18 caractéristiques de premier ordre - 17 caractéristiques basées sur la forme - 56 caractéristiques de texture - caractéristiques de premier ordre et caractéristiques de texture calculées en plus après 8 décompositions par ondelette = 1423 caractéristiques d'images
Données cliniques	Age et sexe	∅	Age et sexe
Données spatiales	Coordonnées 3D du centre de la lésion	∅	∅
Combinaison de l'ensemble des données	Radiomiques + Cliniques + Spatiales	∅	Radiomiques + Cliniques
Modèle	Validation croisée: -One vs one -One vs others	Validation croisée : - Modèle multiclassé (pour comparer 3D vs 2D) - One vs one pour différencier les primitifs des métastases	Validation croisée + comparaison à la lecture de deux radiologues: - Modèle à 3 classes (Poumon à petites cellules / Sein /Mélanome) - Modèle à 5 classes (Poumon à petites et non petites cellules / Sein / Mélanome / Colorectal)
Metric(s) utilisé(s)	F1 et Roc AUC	Roc AUC	Roc AUC
Résultats Présentés sur	One vs others	One versus one	Modèle à 5 classes
Roc Auc Radiomics seul Poumon	<u>0,64</u> (vs others) / 0,66 (sein) / 0,78 (Mélanome)	0,96 (sein) / 0,93 (mélanome)	<u>0,74</u> (Petites cellules) <u>0,63</u> (non à petites cellules)
Roc Auc Radiomics seul Sein	0,63 / clinique : sexe : 0,78	0,60 (vs mélanome)	0,61 / clinique: sexe :0,73
Roc Auc Radiomics seul Mélanome	0,75	0,60 (vs sein)	0,80
Roc Auc Radiomics seul Colon	0,76	Non réalisé	0,68
Conclusions supplémentaires de		Caractéristiques 3D > caractéristiques 2D	Machine learning pour tous les types de primitif > Lecture radiologues

Concernant la méthodologie :

Les premières grandes différences de **l'étude de Kniep** [86] avec notre étude concernent le nombre important de métastases étudiées (au nombre de 658 métastases pour 189 patients contre 237 métastases pour 188 patients dans notre étude) et la séparation en sous-groupes du cancer du poumon, en petites cellules et non à petites cellules.

Les séquences FLAIR et T1 non injectées ont été ajoutées à l'analyse et ont servi également à l'étude sur les critères morphologiques connus de métastases cérébrales à l'IRM (annexe 12), pour la lecture des deux radiologues (de 8 ans et 7 ans d'expérience en neuroradiologie), associées aux données cliniques (âge et sexe).

Une autre différence à noter est la réalisation de la segmentation de manière semi-automatique par deux neuroradiologues de 11 ans et de 3 ans d'expérience contre celle réalisée par une seule interne, de fin d'internat, dans notre étude.

La validation de leur modèle a été effectuée sur la base de validation croisée avec des sets d'apprentissage et de validation indépendants.

Pour les patients présentant plusieurs métastases, les métastases respectives ont été affectées soit au set de validation, soit au set d'apprentissage.

Les caractéristiques radiomiques ont aussi été définies par le même logiciel que le nôtre (PyRadiomics Python).

La deuxième grande différence est le nombre important de caractéristiques extraites par l'équipe de Kniep [86] au nombre de 1423 caractéristiques contre 105 dans notre étude.

Cependant, de nombreuses caractéristiques de leur étude, étaient corrélées du fait de la transformation en ondelettes (n'étant que des filtres successifs appliqués à l'image et donc l'évaluation ne concernait que les mêmes types de caractéristiques).

Au final, ils ont obtenu similairement le même nombre de caractéristiques extraites que dans notre étude (18 caractéristiques de premier ordre, 17 caractéristiques basées sur la forme, et 56 caractéristiques de texture dans leur étude et respectivement 18/14/73, dans notre étude).

La création de leur modèle est basée sur un schéma en deux étapes :

-D'abord, un modèle optimiste à trois classes intégrant les trois types de tumeurs les plus fréquentes (cancer du poumon à petites cellules, sein et mélanome) comprenant seulement des images de haute résolution.

-Puis, un modèle à cinq classes intégrant les cinq types de tumeurs primaires (cancer du poumon à petites cellules et non à petites cellules, cancer du sein, mélanome, cancer colorectal) comprenant des images de résolution inférieure.

Le modèle à 5 classes a pour but, par conséquent, de saisir de manière exhaustive, le potentiel du diagnostic basé sur la radiomique dans les environnements cliniques habituels. Les informations ont été analysées en trois combinaisons, informations cliniques à savoir le sexe et l'âge isolément, les caractéristiques radiomiques seules, et l'association de ces biomarqueurs d'imagerie aux informations cliniques. De plus, l'algorithme Random Forest a été entraîné avec des classes de poids égaux et avec des classes de poids en fonction des probabilités préalables des types tumoraux métastatiques, rapportées dans la littérature.

Les scores roc-AUC pour les problèmes multiclassés ont été calculés en utilisant les performances de prédiction de la classe positive (c'est-à-dire du type de tumeur métastatique respectif) par rapport à toutes les autres classes. Les scores roc-AUC ont été basés sur les moyennes calculées à partir de tous les sets de la validation croisée.

Par ailleurs, **l'équipe d'Ortiz** et al. [85] a seulement étudié les métastases de cancers pulmonaire, mammaire et de mélanome (et non de cancer colorectal comme dans notre étude), et ce, sur une petite cohorte de 67 métastases (de 38 patients) (versus 237 métastases de 188 patients dans notre étude).

La segmentation a de plus, été réalisée, sur les séquences 3D T1 après injection de gadolinium mais en 2D (sur une seule coupe axiale représentant la composante solide lésionnelle la plus importante) par un neuroradiologue de 20 ans d'expérience et en 3D, sur toutes les coupes axiales représentant la tumeur, de manière semi automatique, selon un logiciel développé spécifiquement pour cette étude.

Avant l'extraction des caractéristiques par le logiciel MATLAB, une quantification des niveaux de gris a été appliquée pour réduire le temps de calcul et améliorer le rapport signal sur bruit du résultat de texture. Différents nombres de niveaux de gris ont été testés (8,16,32,64 et 128) afin d'étudier l'influence du processus de quantification sur le pouvoir discriminant des caractéristiques. Les caractéristiques, au nombre de 43, dont 3 caractéristiques de premier ordre (contre 18 dans notre étude), et 40 autres, issues des caractéristiques de texture (versus 73 dans notre étude), ont été extraites par la suite.

Enfin, 10 sets de données de caractéristiques de texture différents ont été obtenus : cinq jeux de données, un par type de discrétisation des images natives (quantification des niveaux de

gris en cinq niveaux : 8, 16, 32, 64 et 128), provenant de l'extraction 2D et cinq autres jeux, provenant de l'extraction 3D.

Ces données ont été analysées séparément en utilisant une stratégie multiclasse. Le Random Forest a été choisi comme modèle prédictif afin d'évaluer le pouvoir discriminant des caractéristiques. Les métriques statistiques résultantes dérivées de la performance du modèle de chaque jeu de données ont été comparées afin d'identifier le jeu de données offrant les meilleurs résultats de classification. Puis, l'ensemble des données a été analysé à l'aide de la stratégie one versus one afin d'examiner la capacité de ces fonctionnalités à différencier les types individuels de métastases cérébrales (poumon, sein, mélanome). Compte tenu de la petite taille de l'échantillon, les performances du classifieur ont été évaluées à l'aide d'une validation croisée imbriquée. Dans chaque groupe, 25 % des échantillons ont été sélectionnés de manière aléatoire en tant que jeux de validation et 75 % restants ont été utilisés en tant qu'apprentissage.

Les métastases cérébrales provenant du même patient ont été traitées de manière non distinctive.

Le roc-Auc a été calculé en calculant la moyenne de l'une par rapport à l'ensemble des statistiques.

Concernant les résultats :

Si l'on fait un bilan des résultats de l'ensemble des études, quelle que soit la stratégie mise en œuvre, les performances des modèles intégrant des caractéristiques radiomiques sont très limitées pour discriminer les métastases du sein (avec un roc-AUC aux alentours de 0,60) mais les données cliniques ressortent dans notre étude et celle de Kniep comme guidant le modèle (roc-AUC supérieure à 0,7). Il est intéressant de noter que la classification des métastases cérébrales du cancer du sein et du mélanome était insatisfaisante dans l'étude d'Ortiz avec un score roc-AUC de seulement 0,60.

Les données radiomiques seules et l'ensemble des données (cliniques, spatiales et radiomiques) ont guidé la classification pour le mélanome dans notre étude comme dans celle de Kniep avec respectivement des scores roc-AUC de 0,75 et 0,75 dans notre étude et de 0,80 et de 0,82 dans celle de Kniep.

Dans cette même étude, les performances en termes de score roc-AUC du modèle final pour la classification des cinq types de cancer primaire étaient égales à 0,63 pour les métastases cérébrales du cancer du poumon non à petites cellules, 0,74 pour les métastases du cancer du poumon à petites cellules, 0,61 les pour métastases mammaires et de 0,80 pour les métastases cérébrales du mélanome, et 0,68 pour le cancer colique, reflétant des résultats similaires aux nôtres. Cela signifie que les signatures radiomiques seraient moins discriminantes pour le cancer du poumon non à petites cellules et le cancer du sein. Ce résultat peut s'expliquer par la plus grande hétérogénéité lésionnelle de ces deux types de cancers : les cancers pulmonaire et mammaire ont davantage de sous-types anatomopathologiques (ex : pour le cancer du poumon : adénocarcinome (avec plusieurs variantes), carcinome épidermoïde, carcinome à grandes cellules, carcinome à petites cellules, tumeur carcinoïde, carcinome adénoquameux, carcinome sarcomatoïde, ... [87] ou pour le cancer mammaire : carcinome intracanaux, carcinome lobulaire in situ, carcinome canalaire infiltrant, lobulaire infiltrant, mucineux, médullaire, papillaire, tubuleux, adénoïde kystique, sécrétant [88]...) que le mélanome (mélanome superficiel extensif, nodulaire, acrolentigineux, de Dubreuilh [89]), et le cancer colorectal (adénocarcinome lieberkuhnien dans 95% des cas [90]).

Par ailleurs, les ajustements des poids de classe pendant l'apprentissage en fonction des probabilités antérieures connues de types de tumeurs métastatiques n'ont pas modifié les résultats dans l'étude de Kniep [86].

Concernant l'importance des fonctionnalités, 59 caractéristiques ont été classées parmi les 100 caractéristiques les plus importantes pour les sets d'apprentissage (29 de 1^{er} ordre sur le T1 injecté, 20 caractéristiques de texture sur les séquences T1 injectées, la position longitudinale centrale du ROI et 9 de 1^{er} ordre sur les séquences T1 non injectées) d'après l'étude de Kniep [86]. Les métastases de mélanome ont présenté par ailleurs les niveaux d'intensité les plus élevés alors que des intensités faibles ont été observées pour les métastases de cancer pulmonaire à petites cellules et de cancer colorectal. Les signatures du cancer non à petites cellules ont montré un modèle d'expression moins spécifique.

Contrairement à l'étude d'Ortiz [85], ce dernier n'a retrouvé de manière significative aucune des caractéristiques de radiomique lors de la classification des métastases de cancer mammaire et du mélanome ; à l'inverse de celles du poumon (au nombre de 10 : par ordre décroissant de rang moyen: variance (GLCM), low gray-level zone emphasis (GLSZM), sum

average (GLCM), small zone low gray-level emphasis (GLSZM), short run low gray-level emphasis (GLRLM), low gray-level run emphasis (GLRLM), high gray-level run emphasis (GLRLM), high gray-level zone emphasis (GLSZM), auto correlation (GLCM), gray-level non iniformity (GLSZM), avec $10^{-8} < p < 10^{-2}$), avec un score roc-AUC de 0,96 avec 5 caractéristiques pour les métastases pulmonaires versus mammaires et un score roc-AUC de 0,93 avec 10 caractéristiques pour les métastases pulmonaires versus de mélanome contre un score roc-AUC de 0,60 avec 42 caractéristiques pour les métastases de mélanome versus mammaires.

De plus, les auteurs de l'étude d'Ortiz [85] ont conclu que l'évaluation volumétrique (3D) des caractéristiques de texture serait plus informative et donc de plus grande valeur pour l'identification du cancer primaire que les caractéristiques en 2D.

Cependant, le nombre de niveaux de gris utilisés pour la quantification a affecté les performances du modèle. Les caractéristiques 3D des lésions quantifiées avec des niveaux de gris à 8, 16 et 32 ont fourni un meilleur score roc-AUC que les caractéristiques 2D équivalentes (roc-AUC de 0,865 (3D) versus 0,79 (2D) pour 8 niveaux de gris, roc-AUC de 0,86 (3D) versus 0,84 (2D) pour 16 niveaux de gris, 0,873 (3D) versus 0,78 (2D) pour 32 niveaux de gris). Cependant pour les niveaux de gris à 64, le roc-AUC était similaire (roc-AUC de 0,76 (3D) versus 0,765 (2D)) et pour les niveaux de gris à 128, les caractéristiques 2D étaient plus discriminantes que les 3D (roc-AUC de 0,71 (3D) versus 0,76 (2D)). Par conséquent, les caractéristiques 3D étaient davantage influencées par la quantification des régions IRM que les 2D, perdant du pouvoir discriminant lorsqu'elles augmentaient.

Le roc-AUC le plus élevé a donc été obtenu à l'aide de caractéristiques 3D de lésions quantifiées avec 32 niveaux de gris (roc-Auc de 0,873 +/- 0,064).

A ce propos, plusieurs études ont déjà comparé les performances de classification de différents types de tumeurs cérébrales à partir de caractéristiques de texture 3D ou 2D issues de l'IRM [91][92][93]. Ces dernières avaient déjà démontré une augmentation des performances de classification lors de l'utilisation de caractéristiques 3D.

Enfin, l'étude de Kniep et al. [86] a démontré la supériorité des performances de classification du machine learning comparativement à la lecture seule de deux radiologues (de 8 ans et 7 ans d'expérience en neuroradiologie). Le gain de sensibilité le plus élevé concernait les métastases de mélanome avec une sensibilité de 43 % et 36% pour les deux lecteurs contre

61% et 53% pour le classifieur radiomique. Ces résultats démontrent l'utilité de la machine learning dans la pratique quotidienne, comme aide complémentaire aux radiologues.

III. Forces de l'étude

D'un point de vue méthodologique, le logiciel Pyradiomics [75] avec lequel nous avons travaillé est un logiciel disponible gratuitement, qui permet l'extraction d'index radiomiques à partir d'images médicales à l'aide d'un large panel d'algorithmes codés en Python, un langage qui s'est imposé comme langage open-source populaire pour le calcul scientifique. Ce logiciel bénéficie d'une certification IBSI pour la plupart des index radiomiques extraits [76]. Cette certification est essentielle pour la validation par d'autres équipes des modèles radiomiques publiés. Par ailleurs, malgré l'inclusion d'images IRM réalisées avec un champ magnétique différent (1,5 T et 3 T), une méthodologie complexe du prétraitement des images (correction des inhomogénéités de champ, suppression du crâne, harmonisation des intensités, rééchantillonnage spatial) a permis de rendre les caractéristiques plus robustes au changement de machine. De même, la phase de construction des modèles de Machine Learning a été réalisée en utilisant une validation croisée. Cette méthodologie a permis d'évaluer la généralisabilité des résultats sur 5 cohortes de validation indépendantes. Les performances devront cependant être de nouveau évaluées sur les 20% des métastases du jeu de test indépendant, laissés de côté jusqu'à aujourd'hui, pour fortifier les conclusions de robustesse des modèles proposés. Finalement, des stratégies d'équilibrage des classes ont été mises en œuvre pour éviter des prédictions faussées vers la classe majoritaire et l'analyse de performances faussement élevées. Cette méthodologie était essentielle pour la stratégie One vs Others qui considère dans tous les cas des sous-ensembles très déséquilibrés.

IV. Limites de l'étude

Plusieurs limites à notre étude peuvent être notées. Tout d'abord, un modèle radiomique se base sur les données recueillies pour entraîner le modèle. Dans notre étude, nous avons fait le choix de ne travailler que sur la séquence 3D T1 post gadolinium car il s'agit d'une séquence classique présente dans tous les protocoles standards d'IRM de tumeurs et notamment de la séquence utilisée pour la réalisation des biopsies. Le 3DT1 post gadolinium permet d'autre part de réaliser les analyses radiomiques sur l'intégralité du volume de la masse tumorale définie par la prise de contraste et le centre nécrotique, sans interposition de tissu cérébral sain. Dans la suite, il pourrait être intéressant de rajouter à notre analyse d'autres pondérations, notamment la séquence FLAIR pouvant potentiellement permettre une amélioration des performances du modèle. Deuxièmement, étant donné le peu de données disponibles, la validation croisée a été utilisée pour évaluer les performances de la classification, ce qui nous a permis de tirer des conclusions de manière efficace grâce aux différentes partitions d'apprentissage et de validation. L'étape finale d'évaluation des performances de généralisabilité des signatures doit se faire sur le jeu de test, une fois les modèles optimaux définis pour chacune des tâches. Pour généraliser parfaitement le modèle, il faudrait idéalement l'entraîner par ailleurs sur des données enrichies, c'est-à-dire qu'il y ait suffisamment de patients inclus dans la cohorte d'apprentissage pour que le modèle apprenne les différentes présentations physiologiques, y compris les formes rares d'une même maladie et que les patients proviennent de plusieurs centres en tenant compte de la diversité technique dans la réalisation des protocoles d'acquisition et l'utilisation de machines de constructeurs différents. Par ailleurs, une limitation importante à notre étude est due au fait que les lésions d'un même patient ont été considérées comme indépendantes. De ce fait, dans le cadre de la validation croisée, il est possible d'avoir dans le set d'apprentissage et dans le jeu de validation des lésions appartenant au même patient. Finalement, nous n'avons étudié que les métastases provenant des sites d'origine les plus fréquents. D'autres types de métastases, comme celles du cancer du rein, devraient être pris en compte dans de futures analyses, car nous devrions considérer tous les sites possibles afin de construire un modèle prédictif final efficace. Il serait également intéressant d'analyser les sous-types anatomopathologiques des cancers pour étudier si une signature domine en fonction du sous-type, par exemple les adénocarcinomes par rapport aux carcinomes épidermoïdes. Nous n'avons pas assez d'échantillons pour étudier ce critère. Li et al. [94] avaient utilisé un réseau de neurones pour différencier les métastases cérébrales des différents sous-types de cancer du poumon et obtenu un roc AUC $\geq 0,9$ lors de la différenciation du carcinome du poumon à petites cellules des

autres types de cancer du poumon. De même, l'étude de Kniep et al. [86] a séparé dans son analyse les petites cellules du cancer du poumon des autres types de cancer du poumon. Enfin, il serait intéressant de combiner la radiomique aux séquences avancées telles que la perfusion et la spectroscopie.

V. Perspectives

Concernant les perspectives, le travail devait montrer si une signature apparaissait concernant la classification des sites primaires de métastases cérébrales. Pour ce faire, une première étape de comparaison de chacune des classes deux à deux a été réalisée afin d'évaluer la présence de signal différentiant. Dans un deuxième temps, une tâche correspondant plus à une tâche clinique, la tâche « One Vs. Others » a été mise en œuvre. L'objectif final serait d'implémenter un problème de classification multi-classes. Par ailleurs, dans ce travail, nous avons souhaité voir si la localisation des lésions était informative pour la tâche de classification. Pour ce faire, chaque cerveau a été recalé sur un atlas et les coordonnées spatiales du barycentre de la lésion dans cet atlas ont été extraites. Pour résumer cette information continue en variables discrètes, une idée pourrait être de subdiviser l'atlas en zones anatomiques de faible nombre. Finalement, la segmentation 3D des lésions est chronophage et expert-dépendante. La technique de segmentation utilisée dans ce travail était semi-automatique et incluait une phase de correction manuelle. Cette technique a permis de segmenter l'ensemble du volume de la masse tumorale définie par la prise de contraste et son centre nécrotique ou non, mais est difficilement implémentable en routine clinique. Il serait là aussi intéressant d'utiliser l'intelligence artificielle et notamment des méthodes de Deep Learning et de réseaux neuronaux convolutionnels pour accélérer cette étape. Cette technique de segmentation a déjà fait ses preuves pour la segmentation de métastases cérébrales sur des séquences d'IRM 3D T1 post-gadolinium [95][96][97].

CONCLUSION

Cette étude montre que la radiomique a le potentiel de constituer un outil d'aide au diagnostic dans la différenciation de métastases cérébrales de mélanome.

Les performances diagnostiques obtenues par le modèle n'étaient pas significativement robustes pour différencier les autres sous types lésionnels (pulmonaires, coliques et mammaires).

L'étape finale consistera à réaliser un problème de classification multi-classes pour consolider cette méthodologie et à construire des bases de données standardisées mettant en relation les phénotypes de tumeurs avec des signatures de caractéristiques d'images spécifiques.

BIBLIOGRAPHIE

- [1] Barnholtz-Sloan JS, Sloan AE, Davis FG. Incidence proportions of brain metastases in patients diagnosed (1973 to 2001) in the Metropolitan Detroit Cancer Surveillance System. *J Clin Oncol Off J Am Soc Clin Oncol* 2004; 22: 2865–2872.
- [2] Petrovich Z, Yu C, Giannotta SL. Survival and pattern of failure in brain metastasis treated with stereotactic gamma knife radiosurgery. *J Neurosurg* 2002; 97: 499–506.
- [3] Nayak L, Lee EQ, Wen PY. Epidemiology of brain metastases. *Curr Oncol Rep* 2012; 14: 48–54.
- [4] Taillibert S, Le Rhun É. Épidémiologie des lésions métastatiques cérébrales. *Cancer/Radiothérapie* 2015; 19: 3–9.
- [5] Smedby KE, Brandt L, Bäcklund ML, et al. Brain metastases admissions in Sweden between 1987 and 2006. *Br J Cancer* 2009; 101: 1919–1924.
- [6] T LC, Fp S, P C, et al. Sites of primary malignancies in patients presenting with cerebral metastases. A review of 120 cases. *Cancer* 1985; 56: 880–882.
- [7] Mandybur TI. Intracranial hemorrhage caused by metastatic tumors. *Neurology* 1977; 27: 650–655.
- [8] Cagney DN, Martin AM, Catalano PJ, et al. Incidence and prognosis of patients with brain metastases at diagnosis of systemic malignancy: a population-based study. *Neuro-Oncol* 2017; 19: 1511–1521.
- [9] Steeg PS, Camphausen KA, Smith QR. Brain metastases as preventive and therapeutic targets. *Nat Rev Cancer* 2011; 11: 352–363.
- [10] Yamanaka R. Medical Management of Brain Metastases from Lung Cancer. *Brain Tumors - Curr Emerg Ther Strateg*.
- [11] Hall W, Djalilian H, Nussbaum E, et al. Long-term survival with metastatic cancer to the brain. *Med Oncol* 2000; 17: 279–286.
- [12] Valiente M, Ahluwalia MS, Boire A, et al. The Evolving Landscape of Brain Metastasis. *Trends Cancer* 2018; 4: 176–196.
- [13] Voorhies RM, Sundaresan N, Thaler HT. The single supratentorial lesion. An

- evaluation of preoperative diagnostic tests. *J Neurosurg* 1980; 53: 364–368.
- [14] Füreder LM, Widhalm G, Gatterbauer B, et al. Brain metastases as first manifestation of advanced cancer: exploratory analysis of 459 patients at a tertiary care center. *Clin Exp Metastasis Off J Metastasis Res Soc* 2018; 35: 727.
- [15] Polyzoidis KS, Miliaras G, Pavlidis N. Brain metastasis of unknown primary: A diagnostic and therapeutic dilemma. *Cancer Treat Rev* 2005; 247.
- [16] Achrol AS, Rennert RC, Anders C, et al. Brain metastases. *Nat Rev Dis Primer* 2019; 5: 5.
- [17] Gavrilovic IT, Posner JB. Brain metastases: epidemiology and pathophysiology. *J Neurooncol* 2005; 75: 5–14.
- [18] Chambers AF, Groom AC, MacDonald IC. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer* 2002; 2: 563–572.
- [19] Puhalla S, Elmquist W, Freyer D, et al. Unsanctifying the sanctuary: challenges and opportunities with brain metastases. *Neuro-Oncol* 2015; 17: 639–651.
- [20] Pekmezci M, Perry A. Neuropathology of brain metastases. *Surg Neurol Int* 2013; 4: S245-255.
- [21] Ogawa M, Kurahashi K, Ebina A, et al. Miliary brain metastasis presenting with dementia: progression pattern of cancer metastases in the cerebral cortex. *Neuropathol Off J Jpn Soc Neuropathol* 2007; 27: 390–395.
- [22] Dietemann J-L, p.284-290. *Neuro-imagerie diagnostique*, Elsevier Masson, 2012,.
- [23] Bender ET, Tomé WA. Distribution of brain metastases: implications for non-uniform dose prescriptions. *Br J Radiol* 2011; 84: 649–658.
- [24] Benzakoun J, Robert C, Legrand L, et al. Anatomical and functional MR imaging to define tumoral boundaries and characterize lesions in neuro-oncology. *Cancer/Radiothérapie* 2020; 24: 453–462.
- [25] Yuh WT, Engelken JD, Muhonen MG. Experience with high-dose gadolinium MR imaging in the evaluation of brain metastases. *AJNR Am J Neuroradiol* 1992; 13: 335–345.
- [26] Ramalho J, Semelka RC, Ramalho M, et al. Gadolinium-Based Contrast Agent Accumulation and Toxicity: An Update. *AJNR Am J Neuroradiol* 2016; 37: 1192–1198.
- [27] Bjørnerud A, Vatnehol SAS, Larsson C, et al. Signal Enhancement of the Dentate Nucleus at Unenhanced MR Imaging after Very High Cumulative Doses of the Macrocyclic Gadolinium-based Contrast Agent Gadobutrol: An Observational Study. *Radiology* 2017; 285: 434–444.
- [28] Brain tumors: full- and half-dose contrast-enhanced MR imaging at 3.0 T compared with 1.5 T--Initial Experience. - Abstract - Europe PMC.
- [29] Cohen-Inbar O, Xu Z, Dodson B, et al. Time-delayed contrast-enhanced MRI improves detection of brain metastases: a prospective validation of diagnostic yield. *J Neurooncol* 2016; 130: 485–494.
- [30] Brant-Zawadzki MN, Gillan GD, Atkinson DJ, et al. Three-dimensional MR imaging and display of intracranial disease: improvements with the MP-RAGE sequence and gadolinium. *J Magn Reson Imaging JMRI* 1993; 3: 656–662.
- [31] Suh CH, Jung SC, Kim KW, et al. The detectability of brain metastases using contrast-enhanced spin-echo or gradient-echo images: a systematic review and meta-analysis. *J Neurooncol* 2016; 129: 363–371.
- [32] Hanna SL, Langston JW, Gronemeyer SA. Value of subtraction images in the detection of hemorrhagic brain lesions on contrast-enhanced MR images. *AJNR Am J Neuroradiol* 1991; 12: 681–685.
- [33] Ellingson BM, Kim HJ, Woodworth DC, et al. Recurrent glioblastoma treated with bevacizumab: contrast-enhanced T1-weighted subtraction maps improve tumor delineation

- and aid prediction of survival in a multicenter clinical trial. *Radiology* 2014; 271: 200–210.
- [34] Fink KR, Fink JR. Imaging of brain metastases. *Surg Neurol Int* 2013; 4: S209-219.
- [35] Hayashida Y, Hirai T, Morishita S, et al. Diffusion-weighted imaging of metastatic brain tumors: comparison with histologic type and tumor cellularity. *AJNR Am J Neuroradiol* 2006; 27: 1419–1425.
- [36] Hamon M, Coskun O, Courthéoux P, et al. IRM de diffusion du système nerveux central : applications cliniques. *J Radiol* 2005; 86: 369–385.
- [37] Galanaud D, Nicoli F, Roche P et al. Spectroscopie par résonance magnétique des tumeurs cérébrales. *EM-Consulte*, <https://www.em-consulte.com/article/122045> (accessed 15 August 2020).
- [38] Grand S, Lefournier V, Bessou P et al. le point sur... - Imagerie de perfusion : principes et applications cliniques, *Journal de radiologie*, <https://www.em-consulte.com/article/122157> (accessed 15 August 2020).
- [39] Chernov MF, Ono Y, Kubo O, et al. Comparison of 1H-MRS-detected metabolic characteristics in single metastatic brain tumors of different origin. *Brain Tumor Pathol* 2006; 23: 35–40.
- [40] Huang BY, Kwok L, Castillo M, et al. Association of choline levels and tumor perfusion in brain metastases assessed with proton MR spectroscopy and dynamic susceptibility contrast-enhanced perfusion weighted MRI. *Technol Cancer Res Treat* 2010; 9: 327–337.
- [41] Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol (Berl)* 2016; 131: 803–820.
- [42] Jaffe CC. Measures of response: RECIST, WHO, and new alternatives. *J Clin Oncol Off J Am Soc Clin Oncol* 2006; 24: 3245–3251.
- [43] Karsell PR, Sheedy PF, O’Connell MJ. Computed tomography in search of cancer of unknown origin. *JAMA* 1982; 248: 340–343.
- [44] van de Pol M, van Aalst VC, Wilmink JT, et al. Brain metastases from an unknown primary tumour: which diagnostic procedures are indicated? *J Neurol Neurosurg Psychiatry* 1996; 61: 321–323.
- [45] Kole AC, Nieweg OE, Pruim J, et al. Detection of unknown occult primary tumors using positron emission tomography. *Cancer* 1998; 82: 1160–1166.
- [46] Gutzeit A, Antoch G, Kühl H, et al. Unknown Primary Tumors: Detection with Dual-Modality PET/CT—Initial Experience | *Radiology*, <https://pubs.rsna.org/doi/abs/10.1148/radiol.2341031554> (accessed 15 August 2020).
- [47] Pavlidis N, Fizazi K. Cancer of unknown primary (CUP). *Crit Rev Oncol Hematol* 2005; 54: 243–250.
- [48] Milović M, Popov I, Jelić S. Tumor markers in metastatic disease from cancer of unknown primary origin. *Med Sci Monit Int Med J Exp Clin Res* 2002; 8: MT25-30.
- [49] Wronski M, Arbit E. Surgical treatment of brain metastases from melanoma: A retrospective study of 91 patients | Request PDF, https://www.researchgate.net/publication/12435014_Surgical_treatment_of_brain_metastases_from_melanoma_A_retrospective_study_of_91_patients (accessed 15 August 2020).
- [50] Wronski M, Arbit E, Russo P, et al. Surgical resection of brain metastases from renal cell carcinoma in 50 patients. *Urology* 1996; 47: 187–193.
- [51] Yang M, Nonaka D. A study of immunohistochemical differential expression in pulmonary and mammary carcinomas. *Mod Pathol Off J U S Can Acad Pathol Inc* 2010; 23: 654–661.
- [52] Werling RW, Yaziji H, Bacchi CE, et al. CDX2, a highly sensitive and specific marker

of adenocarcinomas of intestinal origin: an immunohistochemical survey of 476 primary and metastatic carcinomas. *Am J Surg Pathol* 2003; 27: 303–310.

[53] Aerts HJWL, Velazquez ER, Leijenaar RTH, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014; 5: 4006.

[54] Zhu Y, Li H, Guo W, et al. Deciphering Genomic Underpinnings of Quantitative MRI-based Radiomic Phenotypes of Invasive Breast Carcinoma. *Sci Rep* 2015; 5: 17787.

[55] Grossmann P, Stringfield O, El-Hachem N, et al. Defining the biological basis of radiomic phenotypes in lung cancer. *eLife*; 6. Epub ahead of print 21 2017. DOI: 10.7554/eLife.23421.

[56] Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* 2016; 278: 563–577.

[57] Wu W, Parmar C, Grossmann P, et al. Exploratory Study to Identify Radiomics Classifiers for Lung Cancer Histology. *Front Oncol* 2016; 6: 71.

[58] Mu W, Chen Z, Liang Y, et al. Staging of cervical cancer based on tumor heterogeneity characterized by texture features on (18)F-FDG PET images. *Phys Med Biol* 2015; 60: 5123–5139.

[59] Liang C, Huang Y, He L, et al. The development and validation of a CT-based radiomics signature for the preoperative discrimination of stage I-II and stage III-IV colorectal cancer. *Oncotarget* 2016; 7: 31401–31412.

[60] Cozzi L, Dinapoli N, Fogliata A, et al. Radiomics based analysis to predict local control and survival in hepatocellular carcinoma patients treated with volumetric modulated arc therapy. *BMC Cancer* 2017; 17: 829.

[61] Zhang B, Tian J, Dong D, et al. Radiomics Features of Multiparametric MRI as Novel Prognostic Factors in Advanced Nasopharyngeal Carcinoma. *Clin Cancer Res Off J Am Assoc Cancer Res* 2017; 23: 4259–4269.

[62] Huang Y, Liu Z, He L, et al. Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer. *Radiology* 2016; 281: 947–957.

[63] Micheel CM, Nass SJ, Omen GS et al. Committee on the Review of Omics-Based Tests for Predicting Patient Outcomes in Clinical Trials | The National Academies Press, <https://www.nap.edu/initiative/committee-on-the-review-of-omics-based-tests-for-predicting-patient-outcomes-in-clinical-trials> (accessed 15 August 2020).

[64] Raphael Meier, Urspeter Knecht, Stefan Bauer. Clinical Evaluation of a Fully-automatic Segmentation Method for Longitudinal Brain Tumor Volumetry, https://www.researchgate.net/publication/299382182_Clinical_Evaluation_of_a_Fully-automatic_Segmentation_Method_for_Longitudinal_Brain_Tumor_Volumetry (accessed 15 August 2020).

[65] Dzung L. Pham, Chenyang Xu, and Jerry L. Prince. Current Methods in Medical Image Segmentation | Annual Review of Biomedical Engineering, <https://www.annualreviews.org/doi/abs/10.1146/annurev.bioeng.2.1.315> (accessed 15 August 2020).

[66] Michael Kass, Andrew Witkin & Demetri Terzopoulos. Snakes: Active contour models | SpringerLink, https://link.springer.com/article/10.1007%2F978-1-4020-0741-2_13 (accessed 15 August 2020).

[67] Limkin EJ, Sun R, Dercle L, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol Off J Eur Soc Med Oncol* 2017; 28: 1191–1206.

[68] Jaber Juntu, Jan Sijbers, Dirk Van Dyck et al. Bias Field Correction for MRI Images | SpringerLink, https://link.springer.com/chapter/10.1007/3-540-32390-2_64 (accessed 15 August 2020).

August 2020).

- [69] Delingette H. Segmentation d'images médicales, Université de Strasbourg, <https://igg.icube.unistra.fr/index.php/Publi:ThesesHDR> (2008, accessed 15 August 2020).
- [70] Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* 2010; 29: 1310–1320.
- [71] Avants BB, Tustison NJ, Song G, et al. A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration. *NeuroImage* 2011; 54: 2033–2044.
- [72] Isensee F, Schell M, Pflueger I, et al. Automated brain extraction of multisequence MRI using artificial neural networks. *Hum Brain Mapp* 2019; 40: 4952–4964.
- [73] Shafiq-Ul-Hassan M, Zhang GG, Latifi K, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med Phys* 2017; 44: 1050–1062.
- [74] Mackin D, Fave X, Zhang L, et al. Harmonizing the pixel size in retrospective computed tomography radiomics studies. *PloS One* 2017; 12: e0178524.
- [75] van Griethuysen JJM, Fedorov A, Parmar C. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* 2017; 77: e104–e107.
- [76] Zwanenburg A, Leger S, Vallières M. Image biomarker standardisation initiative. *ArXiv161207003 Cs Eess*.
- [77] Reuzé S, Schernberg A, Orhac F, et al. Radiomics in Nuclear Medicine Applied to Radiation Therapy: Methods, Pitfalls, and Challenges. *Int J Radiat Oncol Biol Phys* 2018; 102: 1117–1142.
- [78] Thibault G, FERTIL B, Navarro C, et al. Texture Indexes and Gray Level Size Zone Matrix Application to Cell Nuclei Classification. 2009.
- [79] Amadasun M, King R. Textural features corresponding to textural properties. *IEEE Trans Syst Man Cybern* 1989; 19: 1264–1274.
- [80] Dong-hui Xu, Arati S. Kurani, D. Raicu et al. RUN-LENGTH ENCODING FOR VOLUMETRIC TEXTURE, <https://www.semanticscholar.org/paper/RUN-LENGTH-ENCODING-FOR-VOLUMETRIC-TEXTURE-Xu-Kurani/dddc1a5e7d45077aea2689a79b8609e79238c338> (accessed 15 August 2020).
- [81] Collins DL, Holmes CJ, Peters TM, et al. Automatic 3-D model-based neuroanatomical segmentation. *Hum Brain Mapp*. Epub ahead of print 1 January 1995. DOI: 10.1002/hbm.460030304.
- [82] Mazziotta J, Toga A, Evans A, et al. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos Trans R Soc Lond Ser B* 2001; 356: 1293–1322.
- [83] Lohmann P, Kocher M, Ruge MI, et al. PET/MRI Radiomics in Patients With Brain Metastases. *Front Neurol*; 11. Epub ahead of print 7 February 2020. DOI: 10.3389/fneur.2020.00001.
- [84] Mandrekar JN. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *J Thorac Oncol* 2010; 5: 1315–1316.
- [85] Ortiz-Ramón R, Larroza A, Ruiz-España S. Classifying brain metastases by their primary site of origin using a radiomics approach based on texture analysis: a feasibility study. *Eur Radiol* 2018; 28: 4514–4523.
- [86] Kniep HC, Madesta F, Schneider T. Radiomics of Brain MRI: Utility in Prediction of Metastatic Tumor Type. *Radiology* 2019; 290: 479–487.
- [87] [guide_k_bronchopulmonaires_finalweb__091013.pdf](https://www.has-sante.fr/upload/docs/application/pdf/2013-10/guide_k_bronchopulmonaires_finalweb__091013.pdf), https://www.has-sante.fr/upload/docs/application/pdf/2013-10/guide_k_bronchopulmonaires_finalweb__091013.pdf (accessed 22 August 2020).
- [88] Collège Français des Pathologistes (CoPath. Item 309 (ex item 159) – Tumeurs du

sein. 11.

[89] Collège Français des Pathologistes (CoPath). Annales de Dermatologie et de Vénérologie - Item 299 (ex item 149) – Tumeurs cutanées, épithéliales et mélaniques, <https://www.em-consulte.com/article/187849> (accessed 22 August 2020).

[90] Collège Français des Pathologistes (CoPath). Campus d'Anatomie Pathologique - Item 298 (ex item 148) – Tumeurs du côlon et du rectum, http://campus.cerimes.fr/anatomie-pathologique/enseignement/anapath_18/site/html/2.html (accessed 22 August 2020).

[91] Georgiadis P, Cavouras D, Kalatzis I, et al. Enhancing the discrimination accuracy between metastases, gliomas and meningiomas on brain MRI by volumetric textural features and ensemble pattern recognition methods. *Magn Reson Imaging* 2009; 27: 120–130.

[92] Fetit AE, Novak J, Peet AC. Three-dimensional textural features of conventional MRI improve diagnostic classification of childhood brain tumours. *NMR Biomed* 2015; 28: 1174–1184.

[93] Mahmoud-Ghoneim D, Toussaint G, Constans JM. Three dimensional texture analysis in MRI: a preliminary evaluation in gliomas. *Magn Reson Imaging* 2003; 21: 983–987.

[94] Li Z, Mao Y, Li H. Differentiating brain metastases from different pathological types of lung cancers using texture analysis of T1 postcontrast MR. *Magn Reson Med* 2016; 76: 1410–1419.

[95] Grøvik E, Yi D, Iv M. Deep learning enables automatic detection and segmentation of brain metastases on multisequence MRI. *J Magn Reson Imaging JMRI* 2020; 51: 175–182.

[96] Xue J, Wang B, Ming Y. Deep learning-based detection and segmentation-assisted management of brain metastases. *Neuro-Oncol* 2020; 22: 505–514.

[97] Nalepa J, Marcinkiewicz M, Kawulok M. Data Augmentation for Brain-Tumor Segmentation: A Review. *Front Comput Neurosci* 2019; 13: 83.

ANNEXES

ANNEXE 1 : Modèles définis comme optimaux pour répondre à la stratégie « One vs. One »

Classification	Informations	Classifieur(s)
Mammaire Vs Pulmonaire 47 VS 50	Modèle complet	RandomOverSampler_MaxAbsScaler_LogisticRegression
	Informations Spatiales	Normalizer_KNeighborsClassifier
	Informations Cliniques	RandomOverSampler_RobustScaler_KNeighborsClassifier
	Informations Radiomiques	SMOTE_Normalizer_LinearSVC
Mammaire Vs Colique 47 VS 33	Modèle complet	RandomOverSampler_BernoulliNB
	Informations Spatiales	MinMaxScaler_SVC
	Informations Cliniques	RandomOverSampler_QuantileTransformer-normal_QuadraticDiscriminantAnalysis
	Informations Radiomiques	MinMaxScaler_SVC
Mammaire Vs Mélanome 47 VS 59	Modèle complet	MinMaxScaler_LogisticRegression
	Informations Spatiales	QuantileTransformer-normal_SVC
	Informations Cliniques	Normalizer_LogisticRegression
	Informations Radiomiques	SMOTE_QuantileTransformer-uniform_LogisticRegression
Pulmonaire Vs Colique 50 VS 33	Modèle complet	MinMaxScaler_SVC
	Informations Spatiales	Normalizer_MLPClassifier
	Informations Cliniques	Normalizer_LogisticRegression
	Informations Radiomiques	MinMaxScaler_SVC
Pulmonaire Vs Mélanome 50 VS 59	Modèle complet	RandomUnderSampler_RidgeClassifier
	Informations Spatiales	RandomUnderSampler_MinMaxScaler_RandomForestClassifier
	Informations Cliniques	RandomOverSampler_QuantileTransformer-normal_GradientBoostingClassifier
	Informations Radiomiques	SMOTE_RidgeClassifier
Mélanome Vs Colique 59 VS 33	Modèle complet	StandardScaler_SVC
	Informations Spatiales	MaxAbsScaler_LogisticRegression
	Informations Cliniques	StandardScaler_BernoulliNB
	Informations Radiomiques	RobustScaler_SVC

ANNEXE 2 : Résultats portant sur la *precision* de la classification One versus One

Classification One Vs One :	Informations	Validation_moyenne_precision	Validation_déviati on standard	Apprentissage_moyenne_precision	Apprentissage_déviati on standard
Mammaire Vs Pulmonaire 47 VS 50	Modèle complet	0,799693307	0,140499787	0,868356057	0,02894179
	Informations Spatiales	0,643288822	0,157687601	0,742117433	0,033184682
	Informations Cliniques	0,845629926	0,14315827	0,89331249	0,036404273
	Informations Radiomiques	0,607393273	0,086753034	0,604646461	0,024192629
Mammaire Vs Colique 47 VS 33	Modèle complet	0,792979354	0,117324885	0,795084633	0,032635135
	Informations Spatiales	0,584925075	0,128216742	0,597634764	0,033201367
	Informations Cliniques	0,803433566	0,122317397	0,803390978	0,035862134
	Informations Radiomiques	0,5875	0,133463478	0,5875	0,03336587
Mammaire Vs Mélanome 47 VS 59	Modèle complet	0,741544456	0,147637619	0,821096428	0,026557911
	Informations Spatiales	0,633323676	0,128347433	0,6990071	0,033058548
	Informations Cliniques	0,556017316	0,122551842	0,556565826	0,030466749
	Informations Radiomiques	0,769252969	0,158147233	0,831672559	0,035700262
Pulmonaire Vs Colique 50 VS 33	Modèle complet	0,603676471	0,108225703	0,602487562	0,026542715
	Informations Spatiales	0,592100302	0,115670292	0,608602077	0,031596884
	Informations Cliniques	0,603676471	0,108225703	0,602487562	0,026542715
	Informations Radiomiques	0,603676471	0,108225703	0,602487562	0,026542715
Pulmonaire Vs Mélanome 50 VS 59	Modèle complet	0,729989788	0,129703189	0,776608562	0,044357071
	Informations Spatiales	0,484907535	0,141208208	0,721089318	0,042199514
	Informations Cliniques	0,713014097	0,092751338	0,864140104	0,034247423
	Informations Radiomiques	0,72638456	0,125783228	0,802193878	0,05274022
Mélanome Vs Colique 59 VS 33	Modèle complet	0,769902761	0,098254055	0,866659598	0,02498367
	Informations Spatiales	0,642105263	0,092884939	0,641132914	0,023947939
	Informations Cliniques	0,764841991	0,099894243	0,75619426	0,030787516
	Informations Radiomiques	0,764970193	0,095107122	0,855073016	0,023247176
		0,6		0,7	0,8

ANNEXE 3 : Résultats portant sur le *recall* de la classification One versus One

Classification One Vs One :					
Classification	Informations	Validation_moyenne_recall	Validation_déviati on standard	Apprentissage_moyenne_recall	Apprentissage_déviati on standard
Mammaire Vs Pulmonaire 47 VS 50	Modèle complet	0,699313575	0,119489914	0,803805416	0,049093691
	Informations Spatiales	0,64518759	0,121119719	0,744316327	0,056680721
	Informations Cliniques	0,764526316	0,113318417	0,827352647	0,026273871
	Informations Radiomiques	0,808026307	0,108388927	0,795713976	0,029095741
Mammaire Vs Colique 47 VS 33	Modèle complet	0,922536353	0,068096547	0,918162654	0,017092268
	Informations Spatiales	0,973150183	0,082395895	0,97513352	0,070554069
	Informations Cliniques	0,908536353	0,071917714	0,906914407	0,030362555
	Informations Radiomiques	1	0	1	0
Mammaire Vs Mélanome 47 VS 59	Modèle complet	0,773724997	0,108449131	0,839663107	0,030336906
	Informations Spatiales	0,594739816	0,164425151	0,651324938	0,056358398
	Informations Cliniques	1	0	1	0
	Informations Radiomiques	0,717539572	0,114682806	0,788459112	0,031850576
Pulmonaire Vs Colique 50 VS 33	Modèle complet	1	0	1	0
	Informations Spatiales	0,938846154	0,120086808	0,97112653	0,057230274
	Informations Cliniques	1	0	1	0
	Informations Radiomiques	1	0	1	0
Pulmonaire Vs Mélanome 50 VS 59	Modèle complet	0,712562993	0,147074569	0,798819681	0,041681808
	Informations Spatiales	0,621431457	0,216609085	0,851996088	0,050329894
	Informations Cliniques	0,717330669	0,138525396	0,894021475	0,019731329
	Informations Radiomiques	0,695849151	0,145370332	0,800792097	0,055436751
Mélanome Vs Colique 59 VS 33	Modèle complet	0,959606616	0,062238372	0,981359413	0,009350084
	Informations Spatiales	1	0	0,999090909	0,004453618
	Informations Cliniques	0,816225108	0,098583079	0,813683981	0,024405506
	Informations Radiomiques	0,945371961	0,053208476	0,980470524	0,01233939
			0,6	0,7	0,8

ANNEXE 4 : Résultats portant sur l'accuracy de la classification One versus One

Classification One Vs One :					
Classification	Informations	Validation_moyenne_accuracy	Validation_déviati on standard	Apprentissage_moyenne_accuracy	Apprentissage_déviati on standard
Mammaire Vs Pulmonaire 47 VS 50	Modèle complet	0,748421053	0,082673472	0,838603936	0,025517344
	Informations Spatiales	0,619789474	0,089804713	0,734572095	0,024861042
	Informations Cliniques	0,810654532	0,115131318	0,905426	0,01841316
	Informations Radiomiques	0,626526316	0,069285711	0,626273726	0,020462701
Mammaire Vs Colique 47 VS 33	Modèle complet	0,815	0,073908727	0,813125	0,02186607
	Informations Spatiales	0,5675	0,110283725	0,5975	0,033154468
	Informations Cliniques	0,815	0,085659208	0,814375	0,019425338
	Informations Radiomiques	0,5875	0,133463478	0,5875	0,03336587
Mammaire Vs Mélanome 47 VS 59	Modèle complet	0,726233766	0,09222531	0,809417367	0,022677947
	Informations Spatiales	0,581731602	0,101158522	0,648140056	0,030864065
	Informations Cliniques	0,556017316	0,122551842	0,556565826	0,030466749
	Informations Radiomiques	0,722251082	0,092445723	0,793865546	0,02564098
Pulmonaire Vs Colique 50 VS 33	Modèle complet	0,603676471	0,108225703	0,602487562	0,026542715
	Informations Spatiales	0,572941176	0,12208398	0,605526911	0,033079059
	Informations Cliniques	0,603676471	0,108225703	0,602487562	0,026542715
	Informations Radiomiques	0,603676471	0,108225703	0,602487562	0,026542715
Pulmonaire Vs Mélanome 50 VS 59	Modèle complet	0,747099567	0,078894082	0,801818182	0,033309423
	Informations Spatiales	0,540952381	0,12175346	0,778980846	0,023177118
	Informations Cliniques	0,739047619	0,067120734	0,886227795	0,017824497
	Informations Radiomiques	0,733852814	0,080159041	0,816901776	0,047185727
Mélanome Vs Colique 59 VS 33	Modèle complet	0,788538012	0,093192671	0,890751573	0,020537615
	Informations Spatiales	0,642105263	0,092884939	0,640814513	0,024680134
	Informations Cliniques	0,716374269	0,093208376	0,711877083	0,028950634
	Informations Radiomiques	0,777894737	0,083960317	0,880422066	0,018478273
			0,6	0,7	0,8

ANNEXE 5 : Résultats portant sur l'apprentissage des scores roc-AUC et F1 de la classification One versus One

Classification One Vs One :					
Classification	Informations	Apprentissage_moyenne_roc-AUC	Apprentissage_déviati on standard	Apprentissage_moyenne_F1	Apprentissage_déviati on standard_F1
Mammaire Vs Pulmonaire 47 VS 50	Modèle complet	0,910803583	0,017071966	0,833898704	0,03065629
	Informations Spatiales	0,79534095	0,030427591	0,741782187	0,032108784
	Informations Cliniques	0,841476902	0,027945998	0,753033109	0,03360285
	Informations Radiomiques	0,670592308	0,028032703	0,686647896	0,018614983
Mammaire Vs Colique 47 VS 33	Modèle complet	0,804824401	0,023515008	0,851808584	0,02088363
	Informations Spatiales	0,665346426	0,031189867	0,738792639	0,028310105
	Informations Cliniques	0,89965196	0,013537892	0,851068497	0,019001925
	Informations Radiomiques	0,8675614	0,031285799	0,739598748	0,02659158
Mammaire Vs Mélanome 47 VS 59	Modèle complet	0,905121571	0,015894208	0,82999029	0,023824501
	Informations Spatiales	0,709387657	0,026589067	0,672211434	0,030369977
	Informations Cliniques	0,68064721	0,01760904	0,714628811	0,02511078
	Informations Radiomiques	0,881976866	0,017670377	0,80906298	0,028081745
Pulmonaire Vs Colique 50 VS 33	Modèle complet	0,842205434	0,024880711	0,751594757	0,020871858
	Informations Spatiales	0,571602909	0,050503279	0,746997253	0,028622724
	Informations Cliniques	0,594510129	0,03435569	0,751594757	0,020871858
	Informations Radiomiques	0,826416082	0,025614816	0,751594757	0,020871858
Pulmonaire Vs Mélanome 50 VS 59	Modèle complet	0,872183201	0,018283765	0,786895703	0,036388923
	Informations Spatiales	0,872328569	0,024912845	0,779127751	0,023730146
	Informations Cliniques	0,961506198	0,009713684	0,878255351	0,017552957
	Informations Radiomiques	0,887988305	0,034569852	0,800943787	0,050265922
Mélanome Vs Colique 59 VS 33	Modèle complet	0,95541803	0,014863826	0,920224061	0,014112729
	Informations Spatiales	0,646762442	0,036045823	0,780817286	0,018346532
	Informations Cliniques	0,747360561	0,029299119	0,783504681	0,022307086
	Informations Radiomiques	0,927232904	0,015017858	0,913240174	0,012732477
		0,6		0,7	
				0,8	

ANNEXE 6 : Modèles définis comme optimaux pour répondre à la stratégie « One vs. Others »

Classification	Informations	Classifieur(s)
Pulmonaire Vs les Autres 50 VS 139	Modèle complet	SMOTE_MaxAbsScaler_LogisticRegression
	Informations Spatiales	RandomUnderSampler_StandardScaler_LinearSVC
	Informations Cliniques	RandomOverSampler_LogisticRegression
	Informations Radiomiques	SMOTE_MaxAbsScaler_SVC
Mélanome Vs les Autres 59 VS 130	Modèle complet	StandardScaler_SVC
	Informations Spatiales	MinMaxScaler_BernoulliNB
	Informations Cliniques	Normalizer_GradientBoostingClassifier
	Informations Radiomiques	StandardScaler_SVC
Mammaire Vs les Autres 47 VS 142	Modèle complet	Normalizer_LogisticRegression
	Informations Spatiales	Normalizer_LogisticRegression
	Informations Cliniques	QuantileTransformer-normal_MLPClassifier
	Informations Radiomiques	Normalizer_LogisticRegression
Colique Vs les Autres 33 VS 156	Modèle complet	QuantileTransformer-uniform_LogisticRegression
	Informations Spatiales	SMOTE_QuantileTransformer-uniform_RandomForestClassifier
	Informations Cliniques	RandomUnderSampler_MinMaxScaler_GaussianNB
	Informations Radiomiques	QuantileTransformer-uniform_LogisticRegression

ANNEXE 7 : Résultats portant sur la *precision* de la classification One versus Others

Classification One Vs Others :		Validation_moyenne_precision	Validation_déviati on standard	Apprentissage_moyenne_precision	Apprentissage_déviati on standard
Classification					
Informations					
Pulmonaire Vs les Autres	Modèle complet	0,412073139	0,115676057	0,525911928	0,031052187
50 VS 139	Informations Spatiales	0,332312939	0,063367424	0,345694191	0,032166486
	Informations Cliniques	0,392839293	0,123290522	0,389625131	0,033592191
	Informations Radiomiques	0,410315244	0,130241987	0,425606856	0,040317667
Mélanome Vs les Autres					
59 VS 130	Modèle complet	0,767712026	0,059806549	0,80976739	0,014867067
	Informations Spatiales	0,691976404	0,05043315	0,696114541	0,012607835
	Informations Cliniques	0,753021107	0,05917202	0,764081866	0,014172266
	Informations Radiomiques	0,766163921	0,060000982	0,808096811	0,016106619
MAMMAIRE Vs les Autres					
47 VS 142	Modèle complet	0,751379801	0,079494619	0,751326246	0,019821022
	Informations Spatiales	0,751379801	0,079494619	0,751326246	0,019821022
	Informations Cliniques	0,787649424	0,08791455	0,798420765	0,016527664
	Informations Radiomiques	0,751379801	0,079494619	0,751326246	0,019821022
COLIQUE Vs les Autres					
33 VS 156	Modèle complet	0,859968499	0,066245765	0,890206992	0,013123017
	Informations Spatiales	0,857705169	0,057017108	0,937706369	0,019080969
	Informations Cliniques	0,910768152	0,073672219	0,925430227	0,014505058
	Informations Radiomiques	0,845058325	0,061441549	0,861496466	0,014347688
			0,6	0,7	0,8

ANNEXE 8 : Résultats portant sur le *recall* de la classification One versus Others

Classification One Vs Others :		Validation_moyenne_recall	Validation_déviati on standard	Apprentissage_moyenne_recall	Apprentissage_déviati on standard
Classification					
Informations					
Pulmonaire Vs les Autres	Modèle complet	0,54614208	0,146434391	0,705655917	0,042246718
50 VS 139	Informations Spatiales	0,63940959	0,154102224	0,641676772	0,059357452
	Informations Cliniques	0,627972694	0,190617415	0,628425611	0,05417722
	Informations Radiomiques	0,517082473	0,168847292	0,543398729	0,070589034
Mélanome Vs les Autres					
59 VS 130	Modèle complet	0,949689529	0,045289506	0,988371233	0,007869003
	Informations Spatiales	0,9984	0,007838367	1	0
	Informations Cliniques	0,956850121	0,041482207	0,972640808	0,007762386
	Informations Radiomiques	0,947995922	0,044228872	0,987235427	0,007730458
MAMMAIRE Vs les Autres					
47 VS 142	Modèle complet	1	0	1	0
	Informations Spatiales	1	0	1	0
	Informations Cliniques	0,961836361	0,042818197	0,969438065	0,016101091
	Informations Radiomiques	1	0	1	0
COLIQUE Vs les Autres					
33 VS 156	Modèle complet	0,969972957	0,032582923	0,987139671	0,005624423
	Informations Spatiales	0,668501564	0,109051441	0,735313716	0,066679415
	Informations Cliniques	0,678205955	0,112037055	0,696482369	0,094121038
	Informations Radiomiques	0,978852217	0,027165411	0,990317299	0,006897778
			0,6	0,7	0,8

ANNEXE 9 : Résultats portant sur l'*accuracy* de la classification One versus Others

Classification One Vs Others :					
Classification	Informations	Validation_moyenne_accuracy	Validation_déviati on standard	Apprentissage_moyenne_accuracy	Apprentissage_déviati on standard
Pulmonaire Vs les Autres 50 VS 139	Modèle complet	0,672062589	0,068267223	0,753706867	0,017636264
	Informations Spatiales	0,569388336	0,062054514	0,583379226	0,033275312
	Informations Cliniques	0,647709815	0,089855788	0,640487975	0,029055016
	Informations Radiomiques	0,675874822	0,079123001	0,683351342	0,03019305
Mélanome Vs les Autres 59 VS 130	Modèle complet	0,766201991	0,058255008	0,832270826	0,015236346
	Informations Spatiales	0,693086771	0,050757898	0,699729871	0,012527317
	Informations Cliniques	0,753257468	0,064261537	0,77459045	0,014552996
	Informations Radiomiques	0,764096728	0,060247059	0,829884977	0,015924824
MAMMAIRE Vs les Autres 47 VS 142	Modèle complet	0,751379801	0,079494619	0,751326246	0,019821022
	Informations Spatiales	0,751379801	0,079494619	0,751326246	0,019821022
	Informations Cliniques	0,772432432	0,078763124	0,793112583	0,018670715
	Informations Radiomiques	0,751379801	0,079494619	0,751326246	0,019821022
COLIQUE Vs les Autres 33 VS 156	Modèle complet	0,844495021	0,06831995	0,888866372	0,01324077
	Informations Spatiales	0,632944523	0,077810018	0,740501917	0,047962987
	Informations Cliniques	0,675078236	0,086970058	0,702917393	0,065806082
	Informations Radiomiques	0,832802276	0,056342916	0,860575113	0,01558575
			0,6	0,7	0,8

ANNEXE 10 : Résultats portant sur l'apprentissage des scores roc-AUC et F1 de la classification One versus Others

Classification One Vs Others :					
Classification	Informations	Apprentissage_moyenne_roc-AUC	Apprentissage_déviati on standard	Apprentissage_moyenne_F1	Apprentissage_déviati on standard_F1
Pulmonaire Vs les Autres 50 VS 139	Modèle complet	0,83513989	0,01752545	0,602026904	0,029660244
	Informations Spatiales	0,605319366	0,023337563	0,44873836	0,03832445
	Informations Cliniques	0,693825546	0,029728698	0,480158879	0,036327216
	Informations Radiomiques	0,68645132	0,025325446	0,474829207	0,038412263
Mélanome Vs les Autres 59 VS 130	Modèle complet	0,93556812	0,011164092	0,890131677	0,010487958
	Informations Spatiales	0,528533849	0,006888523	0,820768902	0,008805489
	Informations Cliniques	0,760907798	0,013070413	0,855767996	0,010022069
	Informations Radiomiques	0,920699039	0,013169653	0,888639819	0,010780597
MAMMAIRE Vs les Autres 47 VS 142	Modèle complet	0,634892734	0,03452636	0,857862196	0,012912248
	Informations Spatiales	0,65049473	0,026054248	0,857862196	0,012912248
	Informations Cliniques	0,802379123	0,018740261	0,875511548	0,012182647
	Informations Radiomiques	0,634892734	0,03452636	0,857862196	0,012912248
COLIQUE Vs les Autres 33 VS 156	Modèle complet	0,906313762	0,018980631	0,936116804	0,007958999
	Informations Spatiales	0,830936726	0,025256233	0,822130093	0,040616258
	Informations Cliniques	0,775977831	0,027009621	0,790565917	0,063524347
	Informations Radiomiques	0,874145588	0,016495911	0,921361423	0,00921228
			0,6	0,7	0,8

ANNEXE 11 : Caractéristiques Morphologiques des Métastases Cérébrales à l'IRM [86]

Tumor type	Leptomeningeal spread	Solitary metastases	Cystic appearance	Hemorrhage	Spread to bones	Central necrosis	T1w CE intensity	Homogeneity of CE accumulation	T1w nonenhanced intensity	Flair Intensity	Perifocal edema
Small cell lung cancer	+	-	+	+			+		iso	cystic: + solid: (-)	+
Breast cancer		-	+	+			+		iso	cystic: + solid: (-)	+
Malignant melanoma	++	+	+	+	+		+		iso	+	
Gastrointestinal cancer	+	-	-	+		+	+	+	+	+	++
Non-small cell lung cancer		+				+	+	+	iso/-	-	++

RESUME et MOTS CLES

Titre de la thèse en français : Caractérisation du sous type histologique par analyse radiomique des IRM cérébrales de métastases
Titre de la thèse en anglais : Characterization of the histological subtype by radiomic analysis of brain metastases MRI.

Résumé en français : Objectif : Examiner la capacité à discriminer le site primitif des métastases cérébrales, sur la base d'un modèle combinant des données cliniques, des données spatiales grâce à un atlas de localisation des tumeurs basé sur l'IRM et des caractéristiques radiomiques provenant de l'IRM 3D pondérée T1 après injection de gadolinium. **Méthodes :** Cette étude rétrospective bicentrique a inclus les caractéristiques radiomiques de 4 types de métastases cérébrales les plus fréquentes parmi 237 métastases cérébrales provenant de la séquence 3DT1 post gadolinium, chez 188 patients non traités, avec 63 métastases provenant de cancer du poumon, 59 de cancer du sein, 74 de mélanome et 41 de cancer colorectal. L'analyse incluait la segmentation de la zone tumorale, le prétraitement des images IRM et l'extraction de caractéristiques radiomiques dont 18 caractéristiques du premier ordre, 14 caractéristiques basées sur la forme de la région tumorale et 73 caractéristiques basées sur la texture des niveaux de gris dans cette même région. Des données cliniques et les coordonnées spatiales 3D du barycentre de la tumeur dans un atlas cérébral ont également été analysées. Les modèles de classification ont été examinés selon quatre approches différentes : uniquement les caractéristiques cliniques (âge, sexe), uniquement les coordonnées spatiales de la localisation de la tumeur, uniquement les caractéristiques radiomiques et une combinaison de toutes les caractéristiques. Deux stratégies ont été adoptées pour examiner les performances du pouvoir discriminant : la stratégie "One-on-one" et la stratégie "One-on-Other". Une validation croisée répétée a été effectuée et les moyennes des scores roc-AUC, F1 (mesure combinant *precision* et *recall* avec leur moyenne harmonique), *recall*, *precision*, *accuracy* ont été rapportés sur l'ensemble des apprentissages et validations de la validation croisée. **Résultats :** Les caractéristiques radiomiques ont montré les meilleurs résultats pour les mélanomes par rapport aux autres (F1 de 0,84). Pour les autres types lésionnels, cependant, ni la radiomique, ni les données spatiales ni les modèles finaux combinés ne permettent d'être assez discriminants. Pour les métastases mammaires, le sexe guide la classification avec score roc-AUC de $0,78 \pm 0,07$ et un score F1 de $0,86 \pm 0,05$. Pour les métastases coliques vs les autres, nous avons observé de bons scores obtenus pour la radiomique, avec un score roc-AUC de $0,76 \pm 0,12$ et un score F1 de $0,90 \pm 0,03$. Cependant ces scores étaient certainement liés au non-équilibre des classes (33 métastases coliques vs. 156 autres) car nous avons noté une absence de signal préalable dans la stratégie one versus one. Enfin, la classification des tumeurs pulmonaires par rapport aux autres catégories n'a pas été satisfaisante d'après les données radiomiques, cliniques et spatiales. **Conclusion :** Cette étude montre que la radiomique a le potentiel de constituer un outil d'aide au diagnostic dans la différenciation de métastases cérébrales de mélanome. Les performances diagnostiques obtenues par le modèle n'étaient pas significativement robustes pour différencier les métastases pulmonaires, coliques et mammaires. L'étape finale consistera à réaliser un problème de classification multi-classes pour consolider cette méthodologie et à construire des bases de données standardisées mettant en relation les phénotypes de tumeurs avec des signatures de caractéristiques d'images plus spécifiques.

Résumé en anglais : Objective : Examine the ability to discriminate the primary site of brain metastases, based on a model combining clinical data, spatial data from an MRI-based tumor localization atlas and radiomic features from T1-weighted 3D MRI after gadolinium injection. **Methods:** This retrospective bicenter study included radiomic features of 4 of the most common brain metastases among 237 brain metastases from the 3DT1 post-gadolinium sequence in 188 untreated patients with 63 metastases from lung cancer, 59 from breast cancer, 74 from melanoma and 41 from colorectal cancer. The analysis included segmentation of the tumor region, pre-processing of MRI images and extraction of radiomic features including 18 first-order features, 14 features based on the shape of the tumor region and 73 features based on the texture of gray levels in the same region. Clinical data and 3D spatial coordinates of the tumor's barycenter in a brain atlas were also analyzed. Classification models were examined using four different approaches: only clinical features (age, sex), only the spatial coordinates of the tumor location, only radiomic features and a combination of all features. Two strategies were adopted to examine the performance of discriminant power: the "One-on-one" strategy and the "One-on-Other" strategy. Repeated cross-validation was performed and the mean roc-AUC score, f1 score (a measure combining *precision* and *recall* with their harmonic mean), *recall*, *precision*, *accuracy* were reported on the whole test, and the cross-validation training set. **Results:** The radiomic characteristics showed the best results for melanomas compared to the others (F1 of 0.84). For the other lesion types, however, neither radiomics, nor spatial data nor combined final models allowed to be discriminating. For breast metastases, gender guides classification with a roc-AUC score of 0.78 ± 0.07 and an F1 score of 0.86 ± 0.05 . For colic metastases vs. others, we observed good scores for radiomics, with a roc-AUC score of 0.76 ± 0.12 and an F1 score of 0.90 ± 0.03 . However, these scores were certainly related to the unbalanced classes (33 colic metastases vs. 156 others) because we noted an absence of signal in the one versus one strategy. Finally, the classification of lung tumors to other primitives was not satisfactory based on radiomic, clinical and spatial data. **Conclusion:** This study shows that radiomics has the potential to be a diagnostic help in the differentiation of brain metastases of melanoma. The diagnostic performance obtained by the model was not significantly robust to differentiate lung, colic and breast metastases. The final step will be to perform a multi-class classification problem to consolidate this methodology and to build standardized databases relating tumor phenotypes to more specific image feature signatures.

Mots clés en français : Métastase cérébrale, Primitif inconnu, IRM 3D T1 post-injection, Radiomique, Apprentissage-machine.

Mots clés en anglais :Brain Metastase, Unknown primary, post-contrast 3D T1-weighted MR, Radiomics, Machine-learning.