كلية الطب والصيدلة وطب الأسنان
FACULTÉ DE MÉDECINE, DE PHARMACIE ET DE MÉDECINE DENTAIRE

جامعة سيدي محمد بن عبد الله - فاس
UNIVERSITÉ SIDI MOHAMED BEN ABDELLAH DE FES

# INTÉRÊT DE CT – RADIOMICS DANS LA PRÉDICTION DES MUTATIONS EGFR POUR LES CANCERS PULMONAIRES NON À PETITES CELLULES

MÉMOIRE PRESENTÉ PAR :

Docteur Mourabiti Abdelaali Yahya

Née le 21/01/1992

POUR L'OBTENTION DU DIPLÔME DE SPECIALITE EN MEDECINE

OPTION : Radiologie

Sous la direction du Professeur Badreeddine Alami

Session Septembre 2023

Dr. ALAMI BADREEDDINE
Professeur Agrégé
Spécialiste en Radiologie
CHU Hassan II - Fès

كلية الطب والصيدلة وطب الأسنان
FACULTÉ DE MÉDECINE, DE PHARMACIE ET DE MÉDECINE DENTAIRE
جامعة سيدي محمد بن عبد الله - فاس
UNIVERSITÉ SIDI MOHAMED BEN ABDELLAH DE FES

# INTÉRÊT DE CT – RADIOMICS DANS LA PRÉDICTION DES MUTATIONS EGFR POUR LES CANCERS PULMONAIRES NON À PETITES CELLULES

MÉMOIRE PRESENTÉ PAR :

Docteur Mourabiti Abdelaali Yahya

Née le 21/01/1992

POUR L'OBTENTION DU DIPLÔME DE SPECIALITE EN MEDECINE

OPTION : Radiologie

Sous la direction du Professeur Badreeddine Alami

Session Septembre 2023

# *Remerciements*

***A notre maître, Monsieur le Professeur***
***et Chef du service de Radiologie centrale***
***Professeur Mustapha MAAROUFI***

*Votre bonté, votre modestie, ainsi que vos qualités professionnelles ne peuvent que susciter mon estime et mon profond respect.*

*Veuillez trouver ici cher Maître, l'assurance de ma reconnaissance, de mon admiration et de mon plus profond respect.*

***A notre maitre, Madame le professeur***

***et Chef du service de Radiologie Mère-Enfant***
***Meryem BOUBBOU,***

*Votre bonté, votre écoute ainsi que vos qualités professionnelles ne peuvent que susciter mon estime et mon profond respect.*

*Veuillez, Chère Professeur, trouver dans ce modeste travail l'expression de ma haute considération et mon profond respect pour avoir guidé mespremiers pas en radiologie pédiatrique.*

***A notre maitre, Monsieur le professeur***
***Moulay Youssef ALAOUI LAMRANI***

*Votre dévouement, votre écoute ainsi que vos qualités professionnelles ne peuvent que susciter mon estime et mon profond respect.*

*Veuillez, Cher Maître, trouver dans ce modeste travail l'expression de ma haute considération et mon profond respect.*

## *A notre maitre, Monsieur le professeur*
## *Badreddine ALAMI*

*Votre dévouement, votre écoute ainsi que vos qualités professionnelles ne peuvent que susciter mon estime et mon profond respect.*

*Veuillez, Cher Maître, trouver dans ce modeste travail l'expression de ma haute considération et mon profond respect*

## *A notre maitre, Madame le professeur*
## *Meryem HALOUA*

*Votre bonté, votre écoute ainsi que vos qualités professionnelles ne peuvent que susciter mon estime et mon profond respect.*

*Veuillez, Cher Maître, trouver dans ce modeste travail l'expression de ma haute considération et mon profond respect*

## *A notre maitre, Monsieur le professeur*
## *Nizar EL BOUARDI*

*Votre modestie, votre bonté, votre écoute ainsi que vos qualités professionnelles ne peuvent que susciter mon estime et mon profond respect.*

*Veuillez, Cher Maître, trouver dans ce modeste travail l'expression de ma haute considération et mon profond respect*

***A notre maitre, Madame le professeur***

***Amal AKAMMAR***

*Votre modestie, votre bonté, votre écoute ainsi que vos qualités professionnelles ne peuvent que susciter mon estime et mon profond respect.*

*Veuillez, Cher Maître, trouver dans ce modeste travail l'expression de ma haute considération et mon profond respect*

***A notre maitre, Monsieur le professeur***

***Hassan QJIDAA***

*Votre bonté, votre modestie, ainsi que vos qualités professionnelles ne peuvent que susciter mon estime et mon profond respect.*

*Veuillez trouver ici cher Maître, l'assurance de ma reconnaissance, de mon admiration et de mon plus profond respect*

## *A mes collègues radiologues :*

*Plus que des ami(e)s, vous avez été pour moi des frères et des sœurs.*

*Je vous remercie pour tous les bons moments passés avec vous au service de radiologie. Et je vous souhaite beaucoup de bonheur et de réussite dans vos projets professionnels et familiaux.*

## *A mes parents, ma femme et ma famille :*

*En particulier ma défunte mère qui n'a pu assister encore à la fin d'une œuvre de ma vie.*

*Merci pour vos sacrifices et votre présence rassurante. Que ce travail soit l'expression de ma profonde reconnaissance*

# PLAN

# Table des matières

*PLAN* ..... *7*

*LISTE DES ABREVIATION* ..... *11*

*LISTE DES FIGURES* ..... *12*

*LISTE DES TABLEAUX* ..... *15*

*INTRODUCTION* ..... *16*

*OBJECTIFS* ..... *19*

*MATERIELS ET METHODES* ..... *21*

*I. Étude épidémiologique :* ..... *22*

**A. Le type d'étude** ..... **22**

**B. Critères d'inclusion et d'exclusion** ..... **22**

1. Critères d'inclusion ..... 22
2. Critères d'exclusion ..... 22

**C. Méthodes de recueil des données** ..... **22**

**D. Analyse statistique :** ..... **23**

*II. Segmentation d'image :* ..... *24*

**A. Protocole de tomodensitométrie :** ..... **24**

**B. Base de données : dataset** ..... **24**

**C. Prétraitement d'images :** ..... **24**

**D. Segmentation :** ..... **25**

1. Réseau neuronal convolutif (Convolutional neural network:CNN) ..... 26
2. U-Net: ..... 28
3. VGG U-Net: ..... 29
4. ResU-Net : ..... 30
5. MobileNetV2U-Net : ..... 31
6. Mesures d'évaluation ..... 32

*III. Radiomics :* ..... *33*

**A. Segmentation et extraction des caractéristiques radiomiques :** ..... **33**

**B. Sélection des caractéristiques ; FeatureWiz** ..... **33**

**C. Machine learning :** ..... **34**

*IV. Ethique :* ..... *34*

*RESULTATS* ..... *35*

*I. Étude épidémiologique* ..... *36*

**A. Étude descriptive : ..... 36**
1. La description des paramètres sociodémographiques : ..... 36
a) Âge : ..... 36
b) Sexe : ..... 36
c) Consommation du tabac : ..... 37
2. La Mutation EGFR: ..... 37
3. La description des paramètres scanographiques : ..... 38
a) Localisation : ..... 38
b) Lobe : ..... 38
c) Contours : ..... 38
d) Spiculation : ..... 39
e) Forme : ..... 39
f) Limites : ..... 40
g) Densité : ..... 40
h) Bronchogramme aérique : ..... 40
i) Attachement scissurale : ..... 40
j) Attachement pleural : ..... 40
k) Rétraction scissurale : ..... 41
l) Rétraction pleurale : ..... 41
m) Rehaussement : ..... 42
n) Cavitation : ..... 44
o) Emphysème : ..... 45
p) Calcifications : ..... 45
q) Fibrose : ..... 46
r) Taille: ..... 47
s) Nodule pulmonaire dans le même lobe : ..... 47
t) Nodule pulmonaire dans un autre lobe : ..... 47
u) ADP médiastinale : ..... 47
v) Epanchement pleural homolatéral : ..... 47
w) Épanchement pleural controlatérale : ..... 47
x) Métastases à distance : ..... 47

**B. Étude analytique : ..... 49**
1. Les facteurs associés à la mutation EGFR par analyse univariée ..... 49
a) Paramètres sociodémographiques: ..... 49
b) Paramètres scanographiques : ..... 49
2. Les facteurs associés à la mutation EGFR par analyse multivariée ..... 53

***II. Segmentation des tumeurs pulmonaires : ..... 54***

***III. Radiomics : ..... 65***

**Classes des caractéristiques ..... 65**

**Classes de filtres ..... 65**

**Modèle1 : modèle développé à partir de toutes les caractéristiques cliniques : ..... 68**

**Modèle 2 : modèle développé à partir des caractéristiques cliniques pertinentes : ..... 69**

**Modèle 3 : modèle développé à partir de toutes les caractéristiques radiomiques : ..... 70**

**Modèle 4 : modèle développé à partir des caractéristiques radiomiques pertinentes : ..... 71**

**Modèle 5 : modèle développé à partir des caractéristiques cliniques et radiomiques pertinentes : 72**

***DISCUSSION ..... 75***

*I. Généralité : ..... 76*

*II. Caractéristiques clinico-radiologiques :..... 79*

**A. Caractérisations cliniques.....79**

Mutation EGFR et sexe féminin ..... 79

Mutation EGFR et tabagisme..... 80

**B. Caractérisations scanographiques .....81**

Mutation EGFR et contours de la tumeur ..... 82

Mutation EGFR et rehaussement hétérogène..... 83

Mutation EGFR et absence d'emphysème ..... 83

Mutation EGFR et attachement pleural : ..... 84

Mutation EGFR et métastase à distance ..... 84

Mutation EGFR et taille tumorale..... 84

Mutation EGFR et localisation tumorale ..... 85

Mutation EGFR et verre dépoli..... 85

Mutation EGFR et bronchogramme aérique ..... 86

Mutation EGFR et rétraction pleurale ..... 86

Mutation EGFR et cavitation ..... 87

*III. Segmentation d'images : ..... 88*

*IV. Radiomics : ..... 92*

*CONCLUSION..... 96*

*RESUMÉ ..... 99*

*BIBLIOGRAPHIE .....105*

# LISTE DES ABREVIATION

**NSCLC** : cancer du poumon non à petites cellules.

**EGFR** : récepteur du facteur de croissance épidermique.

**KRAS** : Kirsten Rat Sarcoma Viral Oncogene Homologue

**ALK** : anaplastic lymphoma kinase

**ITK** : inhibiteurs de la tyrosine kinase

**CT** : scanner

**TDM** : Tomodensitométrie

**IRM** : Imagerie par résonance magnétique.

**HU** : unité de hounsfield

**ADK** : adénocarcinome

**PDC** : produit de contraste

**ADP** : adénopathie.

**CNN** : Convolutional neural network

**MSCNN** : modèle d'empilement de réseaux CNN

**ROC** : Receiver Operating Characteristic

**AUC** : area under curve

**CHU** : centre hospitalier hassan II

**HOSIX** : système d'information Hospitalier

**PACS** : Pictures Archieving and Communicating System

# LISTE DES FIGURES

**Figure 1** : (a) L'image originale et (b) l'histogramme de l'unité de Hounsfield (UH) correspondante, (c) Résultat de l'image après application de l'unité de Hounsfield de [-2000,0] et (d) l'histogramme de l'unité de Hounsfield (UH) correspondante, (e) Résultat de l'image après redimensionnement.
**Figure 2** : Modèle d'empilement utilisant cinq architectures différentes avec une image de sortie plus précise.
**Figure 3** Architecture U-Net. L'architecture est symétrique. L'image se contracte par sous-échantillonnage sur le côté gauche ; l'image s'étend par suréchantillonnage sur le côté droit.
**Figure 4** : Architecture VGG U-Net. L'architecture est symétrique. L'image se contracte par sous-échantillonnage sur le côté gauche ; l'image s'étend par suréchantillonnage sur le côté droit.
**Figure 5** Architecture du ResU-Net. L'architecture est symétrique. L'image se contracte par sous-échantillonnage sur le côté gauche ; l'image s'étend par suréchantillonnage sur le côté droit.
**Figure 6** Architecture de MobileNetV2U-Net. L'architecture est symétrique. L'image se contracte par sous-échantillonnage sur le côté gauche ; l'image s'étend par suréchantillonnage sur le côté droit
**Figure 7** : Répartition des patients en fonction du sexe
**Figure 8** : Répartition des patients en fonction de la consommation du tabac
**Figure 9** : Répartition des patients en fonction de la mutation EGFR
**Figure 10** : Répartition des lésions en fonction du lobe atteint
**Figure 11 :** Cas clinique montant une tumeur pulmonaire à **contours spiculés**.
**Figure 12** : Répartition des lésions en fonction de la forme
**Figure 13** : Répartition des lésions en fonction de la présence ou non d'attachement pleural
**Figure 14 :** Cas clinique montrant une tumeur pulmonaire avec **une rétraction scissurale et pleurale**.
**Figure 15** : Répartition des lésions en fonction du type de rehaussement
**Figure 16** : Cas clinique montrant une tumeur pulmonaire **rehaussée de façon hétérogène après contraste**.
F**igure 17** : Cas clinique montrant une tumeur pulmonaire **excavée**.

**Figure 18** : Cas clinique montrant une tumeur pulmonaire **contenant une calcification**.
**Figure 19** : Cas clinique montrant une tumeur pulmonaire associée à une **fibrose pulmonaire**.
**Figure 20** Cas clinique montrant une tumeur pulmonaire avec **miliaire carcinomateuse**, des **localisations secondaires surrénalienne gauche** et **osseuse**.
**Figure 21** : comparaison entre la mesure manuelle du volume et du diamètre tumoral par le radiologue (bleu) et la mesure automatique.
**Figure 22** : Comparaisons de segmentation. De gauche à droite, les colonnes représentent l'image originale, le masque original, les prédictions U-Net, les prédictions VGG16U-Net, les prédictions VGG19U-Net, les prédictions ResU-Net, les prédictions MobileNetV2U-Net et les prédictions du modèle d'empilement.
**Figure 23** : (a) segmentation manuelle et (b) segmentation automatique
**Figure 24** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique apicale droite
**Figure 25** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique lobaire moyenne
**Figure 26** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique apicale gauche
**Figure 27** (a) segmentation manuelle et (b) segmentation automatique d'une petite tumeur pulmonaire périphérique lobaire moyenne.
**Figure 28** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique lobaire inférieure droite
**Figure 29** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique lobaire inférieure droite
**Figure 30** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique apicale droite
**Figure 31** : Courbes de perte de formation (a) et de validation (b) de la validation croisée (cinq fois).
**Figure 32** : les courbes de Dice de formation (a) et de validation (b) de la validation croisée à cinq facteurs.
**Figure 33** : comparaison des résultats des moyennes des trois critères d'évaluation.
**Figure 34** : algorithme du radiomics dans la détection d'image.
**Figure 35** : modèle récapitulatif de la métanalyse montrant la corrélation entre le sexe féminin et la mutation EGFR.

**Figure 36** : modèle récapitulatif des 27 études montrant la corrélation entre le statut non-fumeur et la mutation EGFR.
Figure 37 : modèle récapitulatif des 21 études montrant la corrélation entre la présence de contours spiculés et la mutation EGFR.
**Figure 38** : modèle récapitulatif des 23 études montrant la corrélation entre la présence de verre dépoli et la mutation EGFR.
**Figure 39** : modèle récapitulatif des 27 études montrant la corrélation entre la présence de bronchogramme aérique et la mutation EGFR.
**Figure 40** : modèle récapitulatif des 11 études montrant la corrélation entre la présence de rétraction pleurale et la mutation EGFR.
**Figure 41** : modèle récapitulatif des 15 études montrant la corrélation entre la présence de cavitation et la mutation EGFR.
**Figure 42** Schéma de l'étude

# LISTE DES TABLEAUX

**Tableau 1** détaille les caractéristiques et les données démographiques des patients pour cinq validations croisées.

**Tableau 2:** Les facteurs socio-démographiques associés à la mutation EGFR

**Tableau 3** : Les paramètres scanographiques associés à la mutation EGFR par analyse univariée

**Tableau 4** : Les facteurs associés à la mutation EGFR par analyse multivariée

**Tableau 5** : comparaison des résultats des trois coefficients avec la validation croisée des cinq modèles

**Tableau 6** : comparaison des résultats des moyennes des trois critères d'évaluation

**Tableau 7** : comparaison des diamètres et des volumes tumoraux mesurés automatiquement, mesurés par le radiologue et avec la moyenne du pourcentage d'erreur MAPE chez 10 patients choisis au hasard

**Tableau 8** : les caractéristique radiomiques les plus pertinentes (n=124

**Tableau 9** : Matrice de confusion pour les 5 modèles étudiés

**Tableau 10** : comparaison de précision, recall et F1 score entre les 5 modèles étudié

**Tableau 11** : comparaison des différents résultats de segmentation des NSCLC.

# INTRODUCTION

Le cancer du poumon est l'un des cancers les plus mortels dans le monde entier, avec des taux de mortalité élevés en raison d'un diagnostic souvent tardif. Le cancer du poumon non à petites cellules (NSCLC) représente environ 85 % de tous les cancers du poumon(1), et il est souvent diagnostiqué à un stade avancé, réduisant considérablement le pronostic. Le récepteur du facteur de croissance épidermique (EGFR) est un marqueur moléculaire important pour le cancer du poumon, car il est sur-exprimé dans de nombreux cas de cancers NSCLC et joue un rôle clé dans la progression de la maladie. La détection précoce de la mutation EGFR chez les patients atteints de cancers pulmonaires non à petite cellule leurs permets de bénéficier d'un traitement ciblé.

La détection radiomique est une méthode d'imagerie qui utilise des techniques d'apprentissage automatique pour extraire des caractéristiques radiologiques à partir d'images médicales. En particulier, les caractéristiques radiomiques sont des mesures quantitatives de l'hétérogénéité et de la texture des tissus dans les images médicales(2), et peuvent fournir des informations précieuses sur la nature et la progression des tumeurs. L'utilisation de la détection radiomique pour identifier des marqueurs moléculaires dans les images médicales peut aider à améliorer la précision du diagnostic, en fournissant des informations supplémentaires sur la biologie tumorale (2).

L'identification de biomarqueurs prédictifs de réponse aux traitements est un enjeu majeur en cancérologie. Dans le cas du cancer du poumon non à petites cellules (NSCLC), une des cibles thérapeutiques clés est le récepteur du facteur de croissance épidermique (EGFR). Cependant, la présence de mutations dans le gène EGFR peut influencer la réponse au traitement, ainsi que la survie des patients atteints de NSCLC. Dans ce contexte, la détection radiomique de l'expression d'EGFR pourrait offrir un moyen non invasif et rapide pour caractériser les tumeurs et de prédire la réponse au traitement.

L'apprentissage automatique, ou machine learning, est une branche de l'informatique qui se concentre sur l'utilisation d'algorithmes pour analyser et apprendre à partir de données. L'utilisation de techniques de machine learning en radiomique a révolutionné la façon dont nous analysons les images médicales, permettant aux chercheurs de trouver des caractéristiques radiomiques pertinentes qui seraient autrement difficiles à identifier. Les algorithmes de

machine learning peuvent également être utilisés pour créer des modèles prédictifs qui peuvent aider à identifier les patients qui sont les plus susceptibles de répondre à un traitement particulier.

Dans ce travail, nous avons commencé par une analyse clinico-radiologiques des données de nos patient, ensuite la segmentations d'images avant de conclure avec l'approche radiomique qui va nous permettre de prédire la mutation EGFR chez les patients atteints de cancers pulmonaires non à petite cellule

# OBJECTIFS

L'objectif principal de ce travail est d'évaluer la valeur de la CT radiomics pour prédire la mutation du récepteur du facteur de croissance épidermique (EGFR) chez les patients atteints de cancer du poumon non à petites cellules (NSCLC), et la combiner avec les caractéristiques cliniques pour construire un modèle performant de prédiction de mutation EGFR.

# MATERIELS ET METHODES

# I. Étude épidémiologique :

## A. Le type d'étude

Il s'agit d'une étude transversale, réalisée chez 149 patients suivis pour ADK pulmonaire ayant bénéficié d'une étude de mutation EGFR au service d'anatomopathologie du CHU de Fès, étalée sur une période de 4 ans (entre Janvier 2018 à Décembre 2021).

## B. Critères d'inclusion et d'exclusion

### 1. Critères d'inclusion

Sont inclus dans notre étude, tous les patients suivis pour ADK pulmonaire ayant bénéficié d'une étude de mutation EGFR et qui ont fait des examens d'imagerie au service de radiologie du CHU HASSAN II de Fès sur une période de 4 ans.

### 2. Critères d'exclusion

Les patients n'ayant pas d'imagerie pré-thérapeutique.

Les patients ayant des tumeurs volumineuses responsables d'envahissement bronchique avec atélectasie d'aval.

Les patients ayant des scanners sans injection du PDC.

Les patients n'ayant pas d'imagerie dans notre formation.

## C. Méthodes de recueil des données

Les données ont été recueillies à l'aide d'une fiche d'exploitation des dossiers informatisés des patients à partir du système d'archivage des données cliniques HOSIX.

Les données scanographiques ont été recueillies à partir du système d'archivage des données radiologiques PACS (Agfa Impax, version 6.6.1.3004, Agfa-Gevaert).

Une base de données a été constituée à partir de cette exploitation incluant les données sociodémographiques, cliniques, anatomopathologiques et les résultats radiologiques (annexe1).

**Les données sociodémographiques :**

L'âge au moment de l'étude, le sexe, et la Consommation du tabac.

**Les données anatomopathologiques** : étude immunohistochimique et recherche de Mutation EGFR

**Les données scanographiques :**

L'analyse des examens scanographiques inclut : La localisation centrale ou périphérique, le lobe pulmonaire atteint, les contours, la présence de spiculation, la forme, les limites, la densité (solide ou mixte), la présence de bronchogramme aérique, l'attachement scissurale, l'attachement pleural, la rétraction scissurale, la rétraction pleurale, le type de rehaussement hétérogène ou homogène, la présence ou non de cavitation, d'emphysème, de calcifications, de fibrose, la taille, la présence de nodule pulmonaire dans le même lobe, de nodule pulmonaire dans un autre lobe, les ADP médiastinales, l'épanchement pleural homolatéral, l'épanchement pleural controlatérale et enfin la présence ou non de métastase à distance.

## D. Analyse statistique :

Dans un premier temps, une analyse descriptive des caractéristiques sociodémographiques, cliniques et radiologiques des patients a été effectuée, les variables quantitatives ont été exprimées par des moyennes et écart-types et les variables qualitatives ont été exprimées par des pourcentages.

Ensuite une analyse univariée a été faite pour rechercher l'association entre la mutation EGFR et les différents paramètres. Lors de la comparaison de Groupes, nous avons utilisé les tests paramétriques classiques (test de Student, Chi-2) en fonction de la nature des variables à comparer.

Enfin une analyse multivariée par régression logistique binaire a été utilisée pour déterminer les facteurs associés à la mutation EGFR après l'ajustement sur les différents facteurs qui avaient une degré de signification $p<0,20$ dans l'analyse univariée.

Pour chaque test statistique utilisé, le test était considéré comme significatif lorsque p (degré de signification) était inférieur à 0,05. L'analyse statistique est effectuée en utilisant le logiciel SPSS version 26.

# II. Segmentation d'image :

## A. Protocole de tomodensitométrie :

Les examens tomodensitométriques ont été réalisés à l'aide d'un scanner GE 64 barrettes. Les acquisitions ont été réalisées en décubitus dorsal en fin d'inspiration.

Les paramètres du scanner étaient les suivants : épaisseur de coupe de 5 mm avec reconstruction de 1,25 mm ; tension du tube de 120 kV ; courant du tube de 100-200 mAs ; collimation de 3 mm ; pitch de 1-1,5 ; taille de la matrice de 512 x 512, FOV de 35 cm. Toutes les images CT ont été transmises à la station de travail de post-traitement

Les images ont été anonymisées et exportées hors ligne pour la segmentation des lésions et la radiomique sur un poste de travail autonome.

## B. Base de données : dataset

Les images ont été collectées à partir de la base des données PACS. Ainsi, nous avons produit une base de données radiomique composée de 149 patients atteints de cancer du poumon qui ont été suivis dans le service de radiologie du CHU de Fès entre janvier 2018 et décembre 2021.

## C. Prétraitement d'images :

Les images de l'ensemble des données sélectionnés étant au format DICOM, nous avons utilisé la bibliothèque [Pydicom] de Python pour extraire ces données. Les valeurs des pixels de l'image ont été converties en unité Hounsfield [UH] qui est une mesure quantitative relative de la densité utilisée par les radiologues dans l'interprétation des images scanographiques. Les valeurs UH sont obtenues à l'aide de la formule suivante UH = pixel × pente + intercept, où la pente et l'intercept ont été obtenus dans les données DICOM. Les valeurs UH de toutes les coupes ont été tronquées à [-2000, 0] puis normalisées à une valeur comprise entre 0 et 1. Les masques binaires respectifs ont été extraits des fichiers RTSTRUCT correspondants. Le volume CT a été rééchantillonné à la même résolution [1 × 1 × 3mm] et chaque coupe a été recadrée de manière aléatoire à une taille de 256 × 256. Pour traiter et redimensionner l'image, nous avons utilisé les bibliothèques SciPy et Ndimage qui fournissent un certain nombre de fonctions générales de traitement et d'analyse d'images pour fonctionner avec

des tableaux de dimension arbitraire. Pour sauvegarder le tableau en tant qu'image 2D, nous avons utilisé la bibliothèque Python Imaging Library [PIL] qui permet de manipuler les images en général et d'effectuer de nombreuses opérations de base sur les images, comme le redimensionnement, le recadrage, la rotation, la conversion des couleurs, etc.

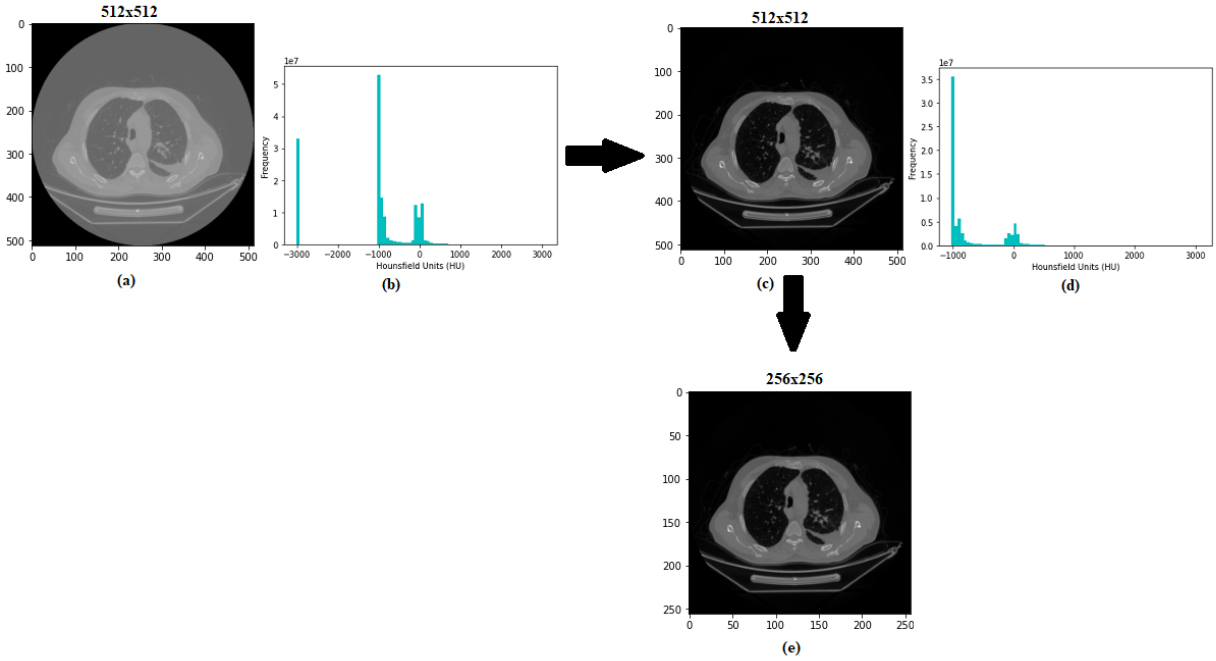



**Figure 1** : (a) L'image originale et (b) l'histogramme de l'unité de Hounsfield (UH) correspondante, (c) Résultat de l'image après application de l'unité de Hounsfield de [-2000,0] et (d) l'histogramme de l'unité de Hounsfield (UH) correspondante, (e) Résultat de l'image après redimensionnement.

## D. Segmentation :

Deux radiologues (un résident et un senior) ont segmenté séparément chaque tumeur pulmonaire manuellement sur des plans axiaux. Afin d'établir une segmentation " de référence ", les radiologues ont réévalué les différences de segmentation. Le logiciel ITK-SNAP a été utilisé pour segmenter la tumeur pulmonaire. Les images ont été importées dans le prototype d'analyse tumorale DECT sur un poste de travail autonome. Le prototype permet la génération d'images dérivées à partir d'ensembles de données DECT, ainsi que la segmentation semi-automatique des lésions, la génération automatique de ROI définies par l'utilisateur, y compris différents volumes et bords de lésions, et une estimation radiomique au sein de chaque ROI.

### 1. Réseau neuronal convolutif (Convolutional neural network:CNN)

La segmentation volumétrique de la tumeur pulmonaire doit être précise, afin de pouvoir utiliser cette segmentation dans le traitement et le suivi post-thérapeutique.

Dans la littérature, on trouve que chaque modèle a sa propre valeur de précision de segmentation. L'utilisation d'un seul réseau de segmentation limite la valeur de précision. Pour cela nous avons proposé un modèle d'empilement de réseaux CNN (MSCNN) formé par cinq réseaux CNN.

L'idée de base de cette approche est de considérer la prédiction de chaque réseau en lui attribuant un score. Une fois que les cinq réseaux ont terminé leur prédiction, le vote obtenu par chaque prédiction est comptabilisé pour chaque pixel. Seule la prédiction ayant le vote le plus élevé est sélectionnée comme sortie du modèle, comme le montre la **figure 2.**

Cependant l'apprentissage initial de ces différents modèles reste difficile devant le nombre élevé des paramètres nécessaires pour le processus de validation. Pour résoudre ce problème, nous avons utilisé la notion d'apprentissage par transfert profond, en se basant sur cinq architectures CNN : U-Net, ResU-Net, VGG16U-Net, VGG19U-Net et MobileNetV2U-Net qui sont déjà pré-entraînées sur des données **ImageNet** que nous avons adapté par la suite à notre base de données. Les cinq CNN sont tous des dérivés de l'architecture U-Net qui est une architecture de réseau neuronal bien adaptée à l'analyse d'images médicales (3). Les procédures de convolution d'image, d'activation ReLU et de normalisation 2D ont été réalisées sur le MSCNN en modifiant la composante de sous-échantillonnage des cinq modèles.

Ce modèle utilise les caractéristiques de plusieurs coupes axiales au même temps pour améliorer la reconstruction d'une seule image 2D. Sa conception est calquée sur celle d'un radiologue qui examine plusieurs images axiales avant de choisir une image 2D.

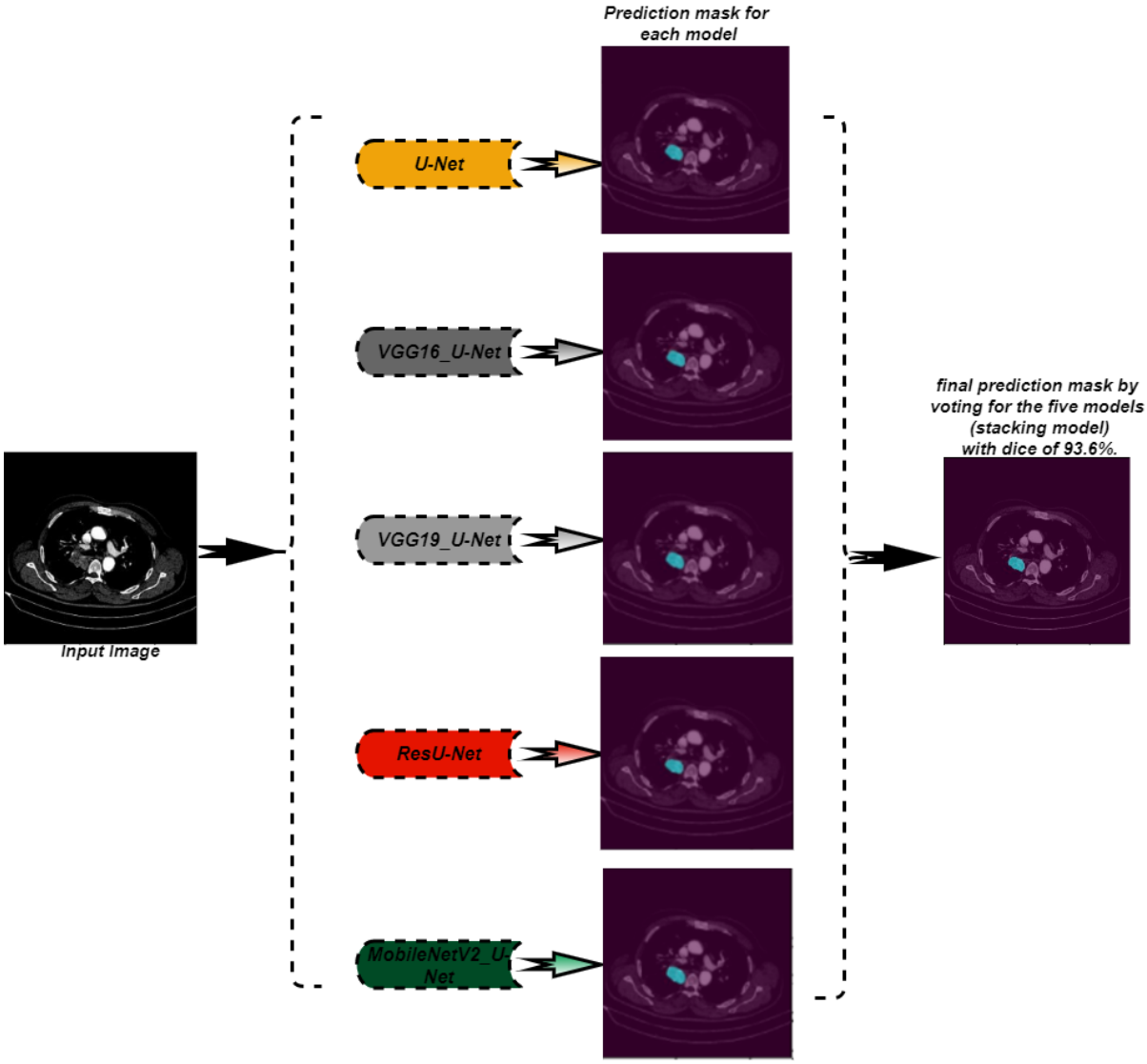



**Figure 2** : Modèle d'empilement utilisant cinq architectures différentes avec une image de sortie plus précise.

Le réseau neuronal a été formé et validé en utilisant la **validation croisée** à cinq facteurs. 80 % des données ont été attribuées comme données d'apprentissage, tandis que les 20 % restants ont été utilisés pour la validation. Chaque scanner thoracique a été utilisé une fois pour la validation et quatre fois pour l'apprentissage avant que le processus ne soit répété cinq fois (tableau1).

**Tableau 1** détaille les caractéristiques et les données démographiques des patients pour cinq validations croisées.

| (80 %) | | (20%) |
|---|---|---|
| Train (80%) | Validation (20%) | Test |
| 4671 | 934 | 1167 |

Le code logiciel de cette étude a été écrit en Python 3.6 en utilisant le logiciel libre TensorFlow r1.9 (licence Apache 2.0). Les réseaux neuronaux ont été entraînés sur une station de travail optimisée pour les processeurs graphiques (GPU) et équipée de 4 cartes GeForce GTX 1080 Ti (11 Go, Pascal microarch). (11Go, Pascal microarch ; NVIDIA, Santa Clara, Californie). Inférence Les statistiques de vitesse pour l'inférence ont été calculées en utilisant un seul GPU.

## 2. U-Net:

U-Net est un réseau neuronal convolutif créé au département informatique de l'université de Fribourg pour la segmentation d'images biomédicales (3). Le réseau est construit sur un réseau entièrement convolutif , mais son architecture a été modifiée et étendue pour lui permettre de fonctionner avec moins d'images d'apprentissage et de fournir des segmentations plus précises.

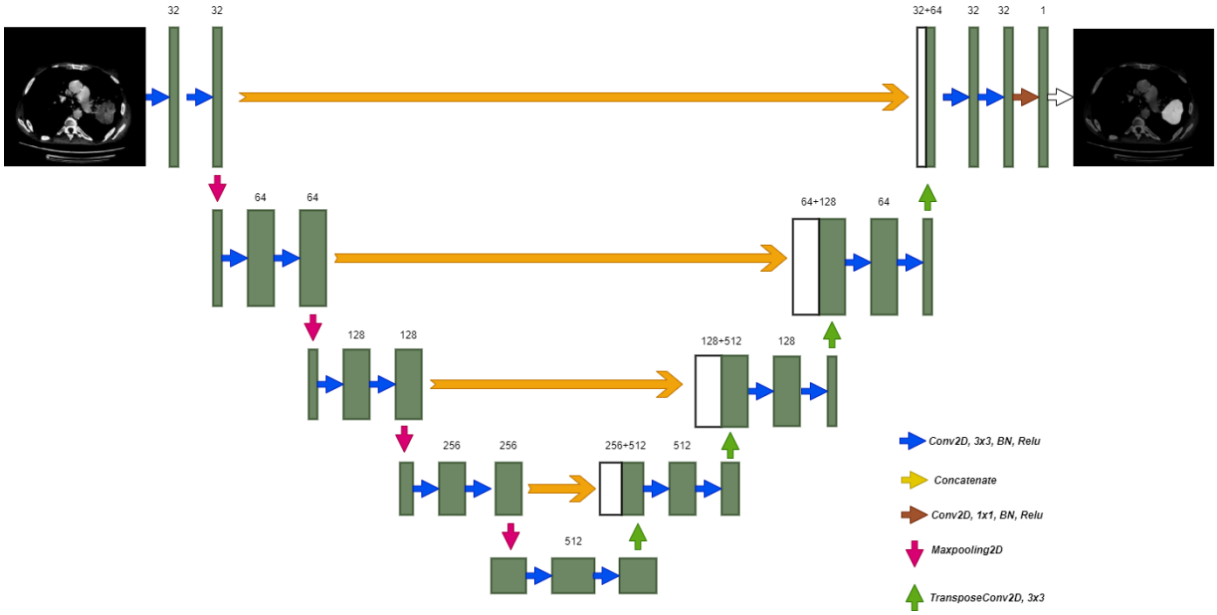



**Figure 3** *Architecture U-Net*. L'architecture est symétrique. L'image se contracte par sous-échantillonnage sur le côté gauche ; l'image s'étend par suréchantillonnage sur le côté droit.

### 3. VGG U-Net:

Dans ce travail, nous présentons une technique d'apprentissage en profondeur qui combine ces deux conceptions de réseau solides, le composant encodeur du U-Net étant VGG- (4). Comme le montre la figure 4.

La partie gauche du réseau est un encodeur et comporte 5 blocs. Il intègre les 13 couches convolutionnelles du VGG- original (5). Après chaque bloc de convolution, la flèche rose indique une opération de MaxPooling qui réduit les dimensions de l'image de 2 × 2. Le côté droit du réseau est un décodeur qui comporte également 5 blocs. La flèche verte après chaque bloc est une opération de suréchantillonnage qui rétablit les dimensions de l'image. Chaque opération de suréchantillonnage répète les lignes et les colonnes de l'image par 2 × 2 (deux lignes et deux colonnes). Les connexions de saut entre les blocs (connexions horizontales avec des flèches noires) sont utilisées pour restaurer les dimensions de l'image. Ces connexions de saut sont mises en œuvre à l'aide de l'opération de concaténation pour combiner les cartes de caractéristiques correspondantes.

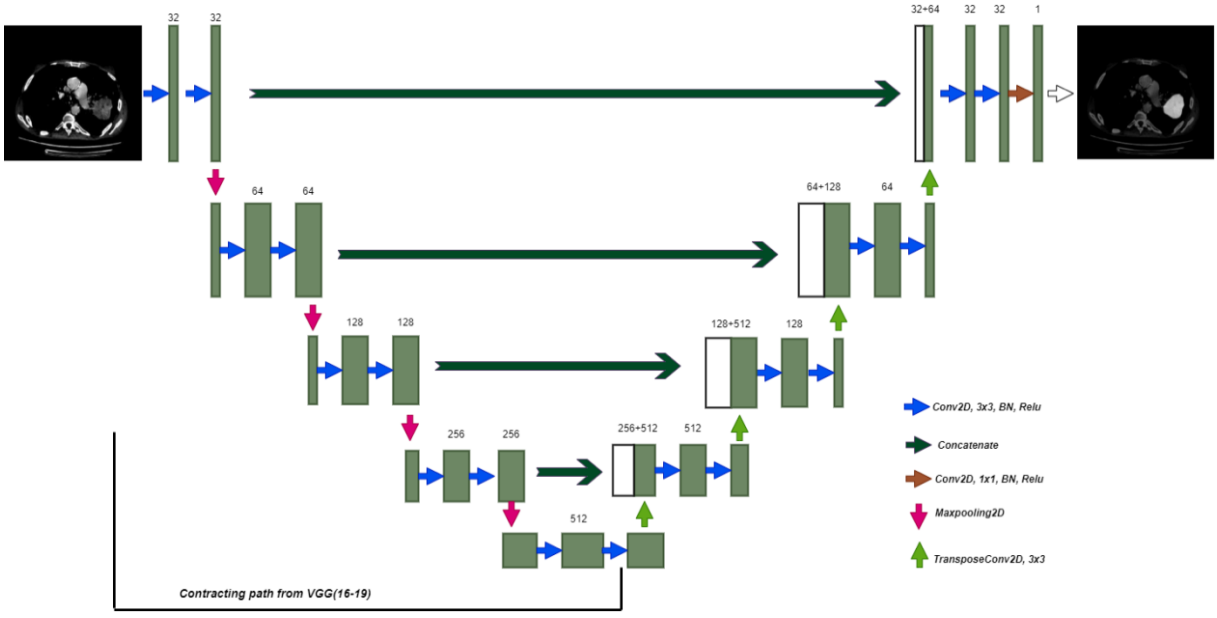



**Figure 4** : *Architecture VGG U-Net*. L'architecture est symétrique. L'image se contracte par sous-échantillonnage sur le côté gauche ; l'image s'étend par suréchantillonnage sur le côté droit.

## 4. ResU-Net :

ResU-Net fait référence à Deep Residual Unet. Il s'agit d'une architecture d'encodage-décodage développée par Zhengxin Zhang et al (6). Pour la segmentation sémantique. Il a été initialement utilisé pour l'extraction à partir d'images aériennes à haute résolution dans le domaine de l'analyse d'images de télédétection. Plus tard, il a été adopté par les chercheurs pour de multiples autres applications telles que la segmentation de polypes, la segmentation de tumeurs cérébrales, etc.

ResU-Net est un réseau neuronal entièrement convolutif conçu pour obtenir des performances élevées avec moins de paramètres. Il s'agit d'une amélioration de l'architecture UNET existante. ResU-NET fait parti à la fois de l'architecture UNET et de l'apprentissage résiduel profond.

Le ResU-Net se compose d'un réseau de codage, d'un réseau de décodage et d'un pont reliant ces deux réseaux, tout comme un U-Net. Le U-Net utilise deux convolutions 3 x 3, où chacune est suivie d'une fonction d'activation ReLU. Dans le cas de ResUNET, ces couches sont remplacées par un bloc résiduel pré-activé.

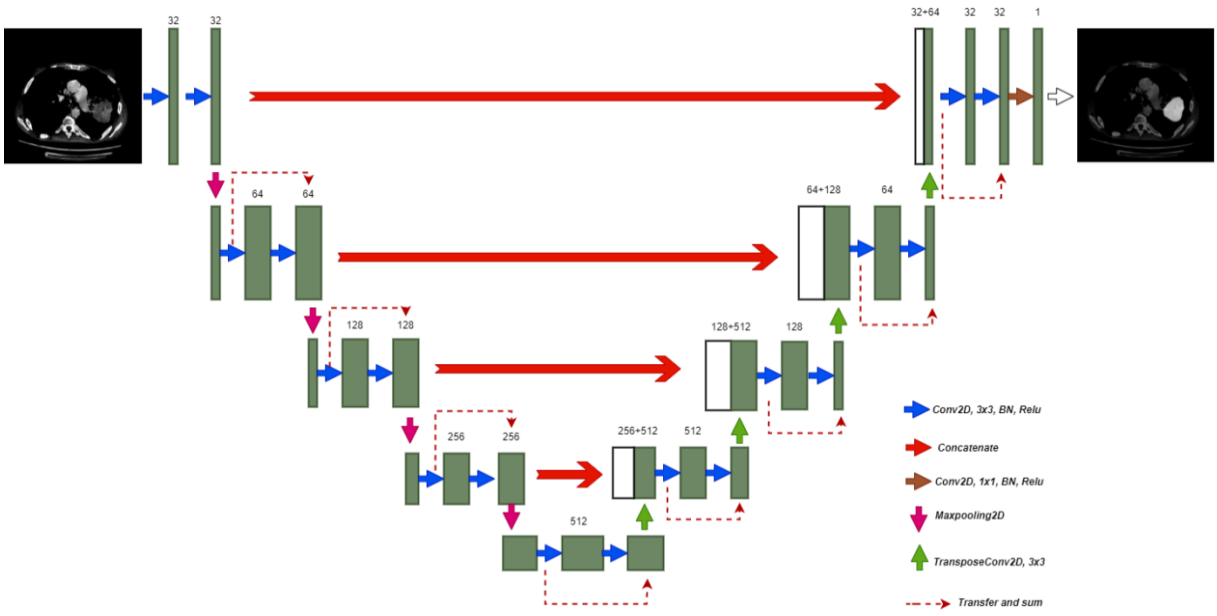



**Figure 5** *Architecture du ResU-Net*. L'architecture est symétrique. L'image se contracte par sous-échantillonnage sur le côté gauche ; l'image s'étend par suréchantillonnage sur le côté droit.

## 5. MobileNetV2U-Net :

L'étape de pré-entraînement comprend l'apprentissage de MobileNetV2 (7) L'étape de segmentation sémantique est l'étape de segmentation, dans laquelle nous transférons la couche de convolution de la première étape comme partie de l'encodeur, puis nous ajoutons cinq couches de déconvolution comme décodeur. Les données d'entraînement sémantique sont constituées d'images originales et d'images marquées manuellement au niveau du pixel.

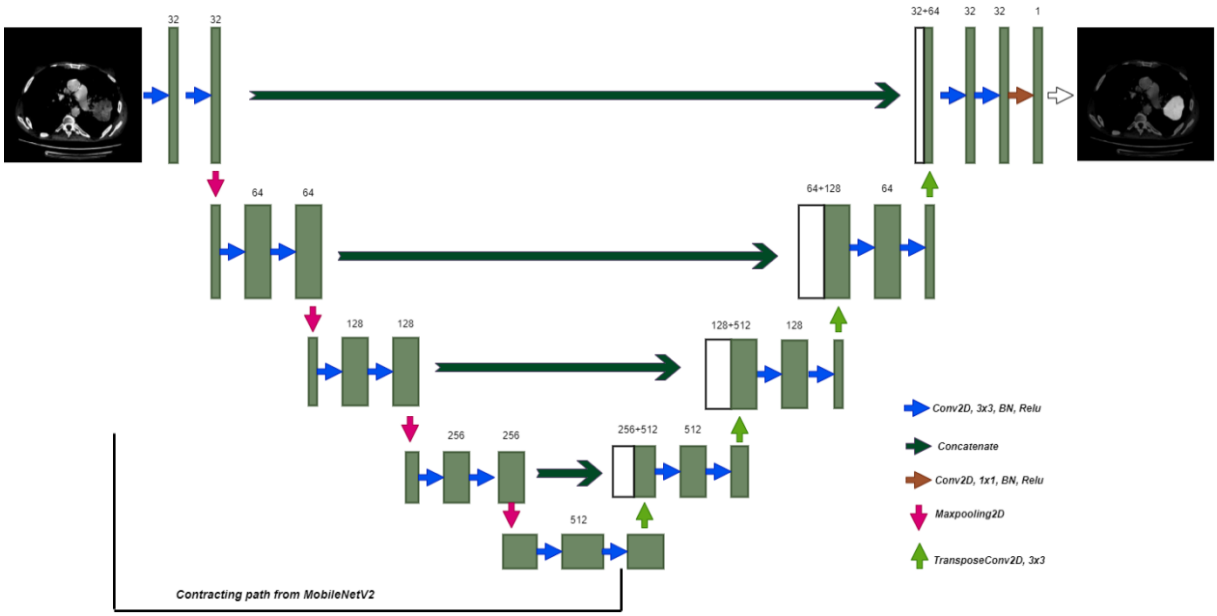



**Figure 6** *Architecture de MobileNetV2U-Net*. L'architecture est symétrique. L'image se contracte par sous-échantillonnage sur le côté gauche ; l'image s'étend par suréchantillonnage sur le côté droit

## 6. Mesures d'évaluation

Nous avons utilisé trois mesures (8) pour valider le processus de segmentation de notre modèle proposé : Le coefficient de similarité de Dice, le coefficient de similarité de Jaccard et le coefficient d'intersection sur l'union. Le premier critère est la mesure du coefficient de similarité de Dice, également connu sous le nom d'indice de Sorensen-Dice, qui est un outil statistique utilisé pour mesurer le taux de similarité entre deux ensembles de données. Le chevauchement spatial entre la segmentation manuelle et la segmentation dérivée du réseau neuronal est quantifié par le score de Dice. Un score de Dice proche de 1 indique que la segmentation manuelle et le réseau neuronal ont généré des contours qui sont en forte concordance. Actuellement, il est devenu l'outil le plus utilisé pour mesurer la puissance de segmentation d'un algorithme en imagerie médicale en quantifiant la similarité entre le volume spatial segmenté par le réseau et celui segmenté par le radiologue. L'équation de ce concept est donnée par :

$$Dice\ [X,Y] = 2 * |X \cap Y| / [|X| + |Y|]$$

Où X est l'ensemble des voxels du volume segmenté par notre modèle et Y est l'ensemble des voxels donné par le médecin. | X | indique le nombre d'éléments dans l'ensemble. Le symbole ∩ indique l'intersection de deux ensembles et le symbole ∪ indique l'union de deux ensembles.

Le deuxième critère est la mesure du coefficient de Jaccard qui mesure la similarité entre des ensembles d'échantillons X , Y et est défini par :

$$J\ [X,Y] = |X \cap Y| / [|X| + |Y|]$$

Le troisième critère est la mesure du coefficient d'intersection sur l'union [IoU], défini comme le rapport entre la zone de chevauchement de la segmentation du médecin, la zone prédite par le réseau et la zone totale :

$$IoU\ [X,Y] = |X \cap Y| / [|X| \cup |Y|]$$

# III. Radiomics :

## A. Segmentation et extraction des caractéristiques radiomiques :

Sur la base de l'ensemble des données, le volume d'intérêt (VOI) a été délimité manuellement sur chaque image CT prétraitée en utilisant ITK-SNAP. L'extraction des caractéristiques radiomiques a été effectuée à l'aide d'une bibliothèque python open source Pyradiomics (9). Un radiologue, ignorant les données cliniques et le statut mutationnel, a segmenté chaque tumeur. Un algorithme d'effet de traçage de niveau a été utilisé pour dessiner une région d'intérêt (ROI) autour de la tumeur ; lorsque l'algorithme ne pouvait pas reconnaître les limites d'une tumeur avec une précision suffisante, l'opérateur corrigeait manuellement le retour sur investissement. En raison de la difficulté de délimiter les contours des lésions en verre dépoli pur, ces lésions ont été exclues.

## B. Sélection des caractéristiques ; FeatureWiz

FeatureWiz est un outil utile pour la sélection de caractéristiques en vue de la construction de modèles de prédiction. Il utilise divers algorithmes pour évaluer les fonctionnalités en fonction de leur contribution à la performance du modèle, et peut aider à éliminer les caractéristiques peu informatives non spécifiques qui peuvent diminuée la performance de l'étude.

En utilisant FeatureWiz, on peut automatiser le processus de sélection de caractéristiques, économiser du temps et de l'effort, et améliorer la qualité des prédictions. Cela peut également aider à éviter l'overfitting, où un modèle est trop adapté aux données d'entraînement et qui ne peut pas être généralisé aux autres données des patients.

Finalement, la sélection de caractéristiques pertinentes est primordiale afin de construire des modèles de prédiction performants, et FeatureWiz est un outil précieux dans ce processus.

### C. Machine learning :

Après la sélection des caractéristiques, les cinq modèles Random Forest suivants ont été construits pour prédire les mutations de l'EGFR : un modèle avec les caractéristiques cliniques, un modèle avec les caractéristiques cliniques pertinentes, un modèle avec les caractéristiques radiomiques, un modèle avec les caractéristiques radiomiques pertinentes et un modèle combinant les caractéristiques radiomiques et cliniques pertinentes.

Random Forest est un algorithme d'apprentissage automatique d'ensemble composé de plusieurs arbres de décision, l'arbre de décision est une technique d'apprentissage supervisée avec une structure arborescente. Random Forest intègre toutes les prédictions faites par les arbres de décision en une seule prédiction. Dans cette étude, 100 arbres de décision ont été générés dans Random Forest. Pour chaque arbre, des données d'entraînement ayant le même nombre que l'ensemble des données (n = 100) ont été obtenues en échantillonnant avec remplacement de l'ensemble des données. Après avoir construit 100 arbres, la prédiction de chaque arbre a été combinée, et la probabilité prédite de mutations de l'EGFR a été déterminée pour chaque patient.

## IV. Ethique :

L'approbation du comité d'éthique pour cette étude, a été approuvée par le comité d'éthique pour la recherche biomédicale Casablanca, Maroc, conformément à la Déclaration d'Helsinki, sous la référence 17/15. Le consentement éclairé écrit a été obtenu pour tous les patients.

L'anonymat et la confidentialité ont été respectés pour tous les participants.

# RESULTATS

# I. Étude épidémiologique

Un total de 149 patients suivis pour ADK pulmonaire ayant bénéficié d'une étude de mutation EGFR au service d'anatomopathologie du CHU de Fès ont été inclus dans l'étude.

## A. Étude descriptive :

### 1. La description des paramètres sociodémographiques :

#### a) Âge :

Dans notre échantillon, l'âge moyen était de 61,05 +/- 11,095ans, avec des extrémités allant de 29 à 93ans.

#### b) Sexe :

Dans notre série, on note une prédominance masculine avec un pourcentage de 70,5% (n=105) (Figure 7). Le sex-ratio H/F était de 2,38

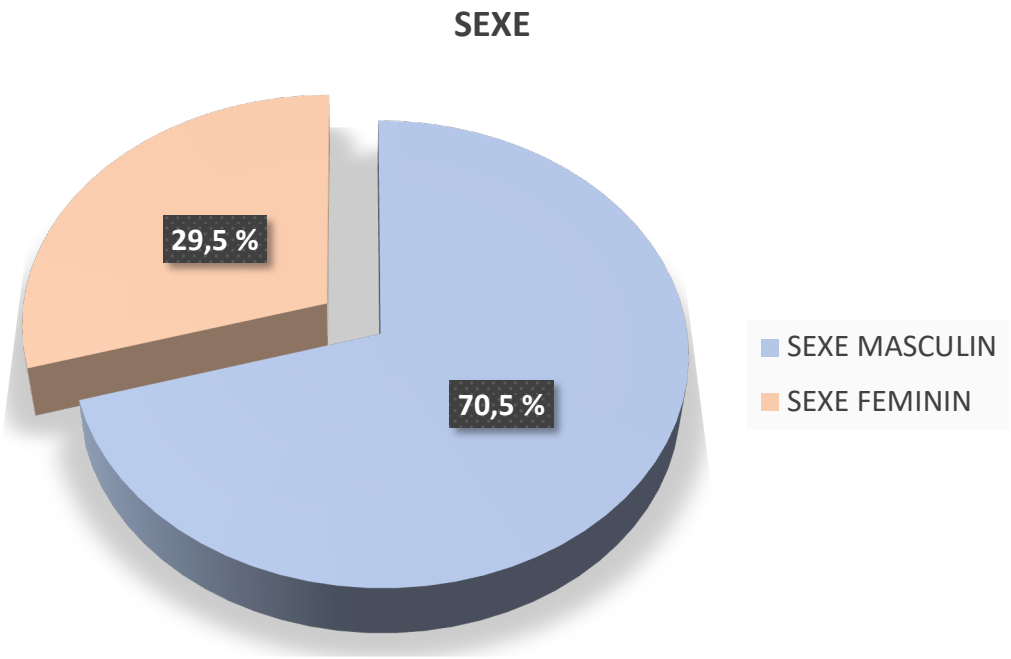



**Figure 7** : Répartition des patients en fonction du sexe

### c) Consommation du tabac :

Dans notre étude, 57,7% (N=86) des patients étaient fumeurs ou ex-fumeurs.

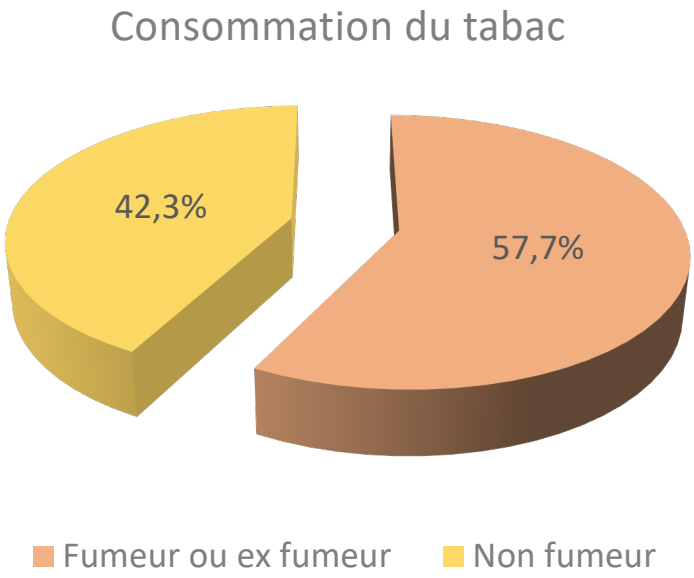



**Figure 8** : Répartition des patients en fonction de la consommation du tabac

## 2. La Mutation EGFR:

Dans notre étude, 41 patients présentaient une mutation EGFR soit un pourcentage de 27,5%.

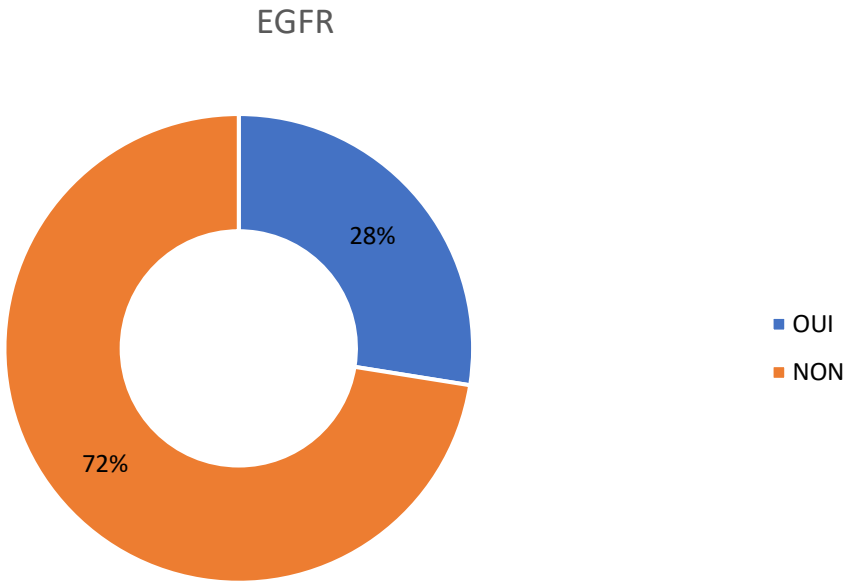



**Figure 9** : Répartition des patients en fonction de la mutation EGFR

## 3. La description des paramètres scanographiques :

### a) Localisation :

La localisation centrale des tumeurs pulmonaires était présente chez 86 de nos patient, avec un pourcentage de 57,7%, cependant 41,6% présentaient une localisation périphérique.

### b) Lobe :

Le lobe supérieur pulmonaire est le plus atteint avec un pourcentage de 68,5%.

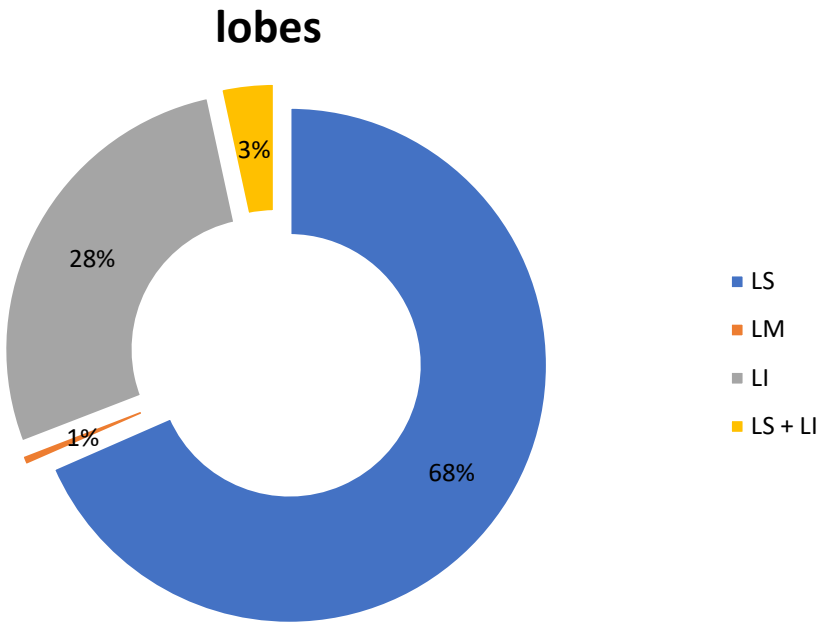



**Figure 10** : Répartition des lésions en fonction du lobe atteint

### c) Contours :

Concernant les contours des tumeurs, la quasi-totalité présentaient des contours irréguliers avec un pourcentage de 91,9% (n= 137), cependant 8,1% des tumeurs (n=12) présentaient des contours réguliers.

### d) Spiculation :

La majorité des tumeurs présentaient des contours spiculés avec un pourcentage de 66,4% (n=99).

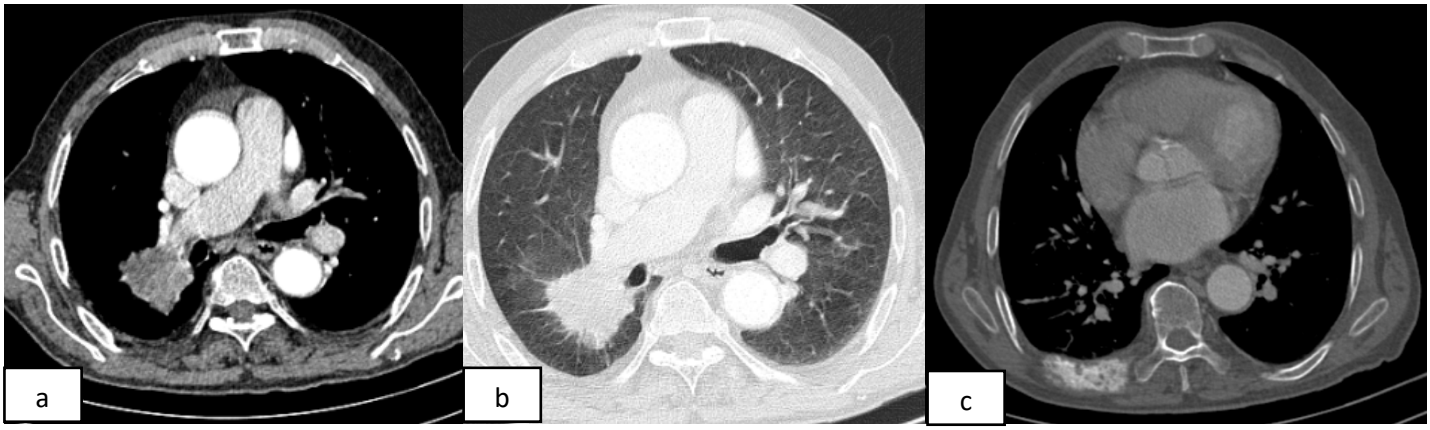



**Figure 11 :** homme de 81 ans, non tabagique, chez qui le scanner thoracique en fenêtre médiastinale (a), parenchymateuse (b) et osseuse (c) montre une masse parenchymateuse pulmonaire LID, à distribution centrale; de densité solide, de **contours spiculés**, avec un rehaussement hétérogène, et localisation secondaire osseuse costale.

### e) Forme :

La forme de la tumeur était indéfinie chez 72 patients, cependant elle était ovalaire chez 35 cas et ronde chez 41 cas.

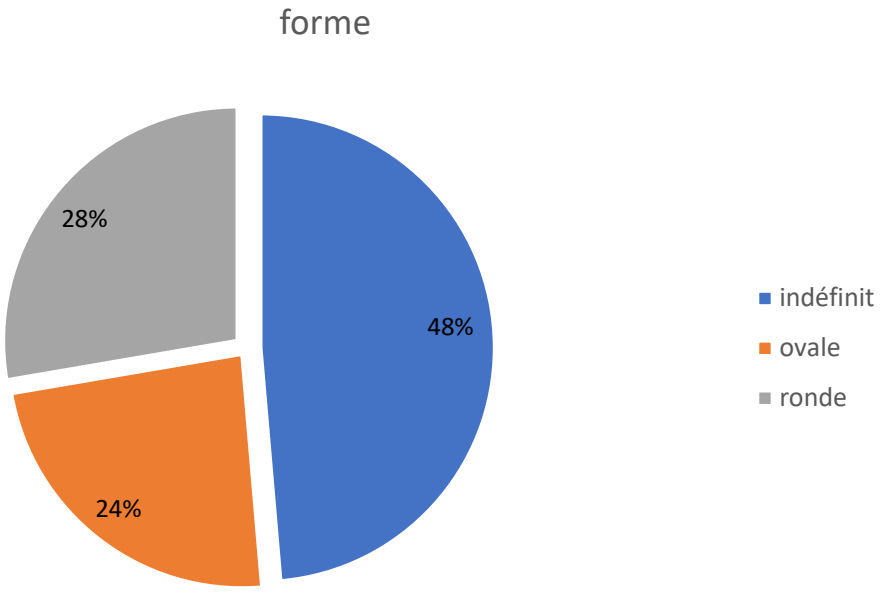



**Figure 12** : Répartition des lésions en fonction de la forme

### f) Limites :

Plus que la moitié des tumeurs présentaient des limites bien définit (51,7%).

### g) Densité :

La majorité des tumeurs étaient solides soit un pourcentage de 89,9%, cependant 10,1% des tumeurs présentaient une densité mixte.

### h) Bronchogramme aérique :

Moins de la moitié des tumeurs présentaient un bronchogramme aérique soit un pourcentage de 40,9%.

### i) Attachement scissurale :

Plus de la moitié des tumeurs présentaient un attachement scissurale soit un pourcentage de 64,8%.

### j) Attachement pleural :

La majorité des tumeurs présentaient un attachement pleural soit un pourcentage de 75,2%.

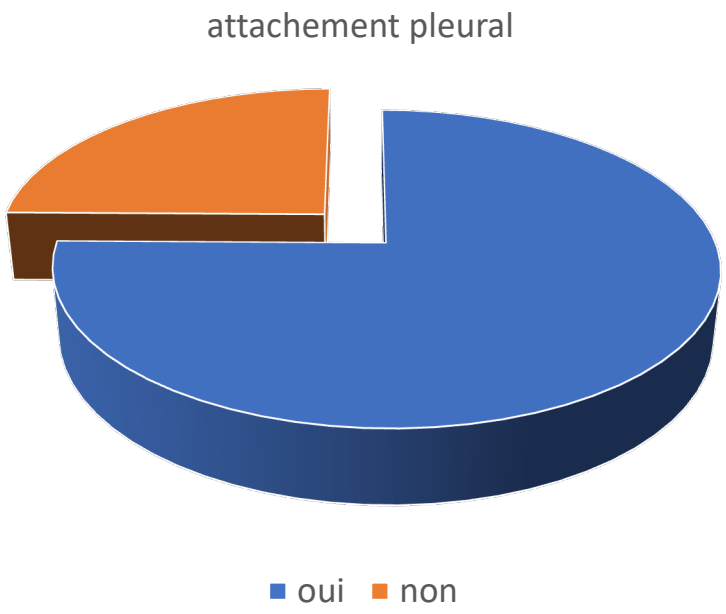



**Figure 13** : Répartition des lésions en fonction de la présence ou non d'attachement pleural

### k) Rétraction scissurale :

La rétraction scissurale était présente chez 33 patients soit un pourcentage de 22,1%.

### l) Rétraction pleurale :

La rétraction pleurale était présente chez 55 patients soit un pourcentage de 36,9%.

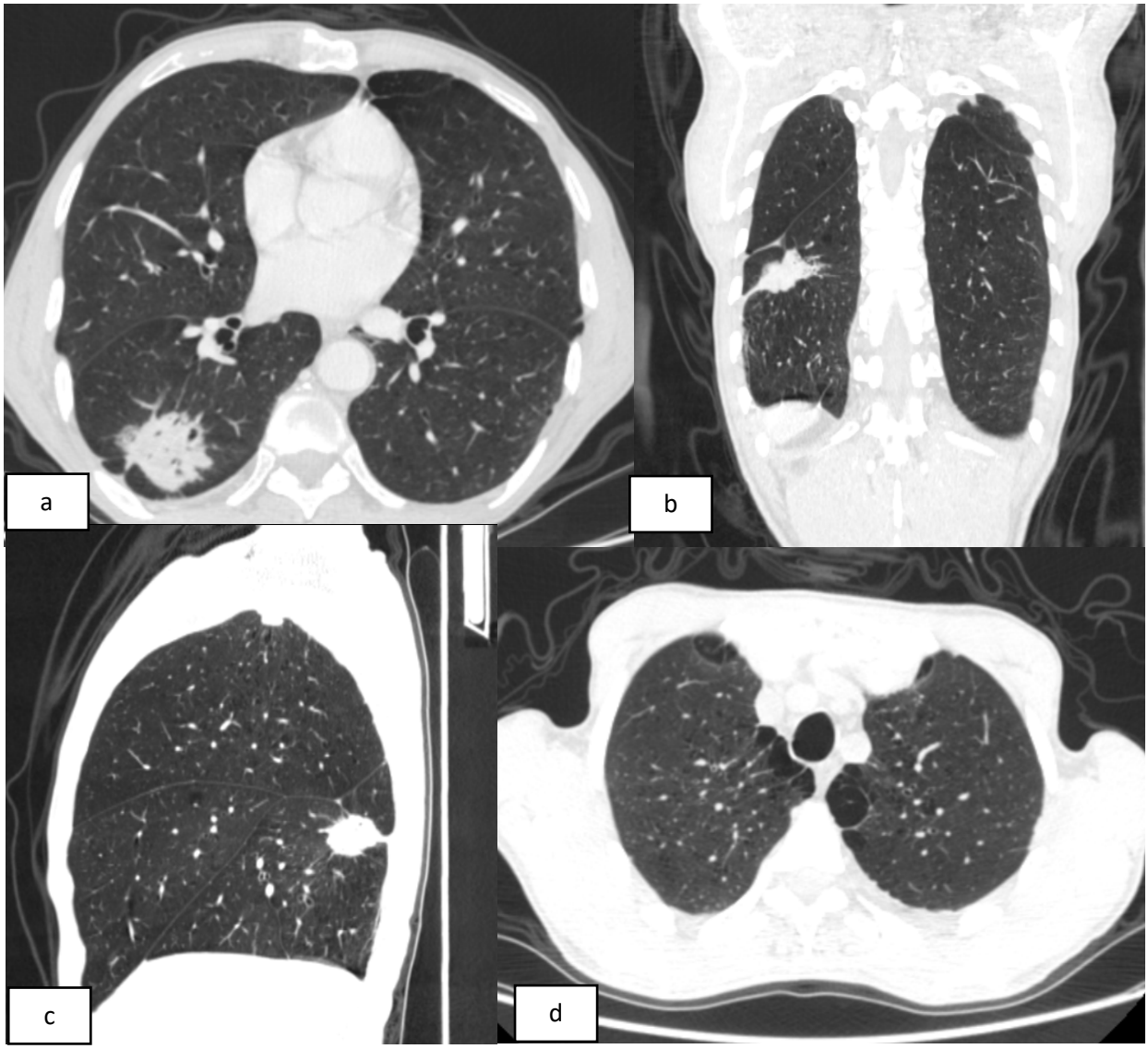



**Figure 14 :** homme de 64 ans, tabagique, chez qui le scanner thoracique en fenêtre parenchymateuse, coupes axiales (a,d), coronale (b), sagittal (c) montre une masse parenchymateuse pulmonaire LID, périphérique; de densité solide, de contours spiculés, avec un bronchogramme aérique, **une rétraction scissurale et pleurale**, sur un poumon emphysémateux (d).

### m) Rehaussement :

Le rehaussement était hétérogène dans la majorité des cas (N=133), et homogène pour 16 patients uniquement.

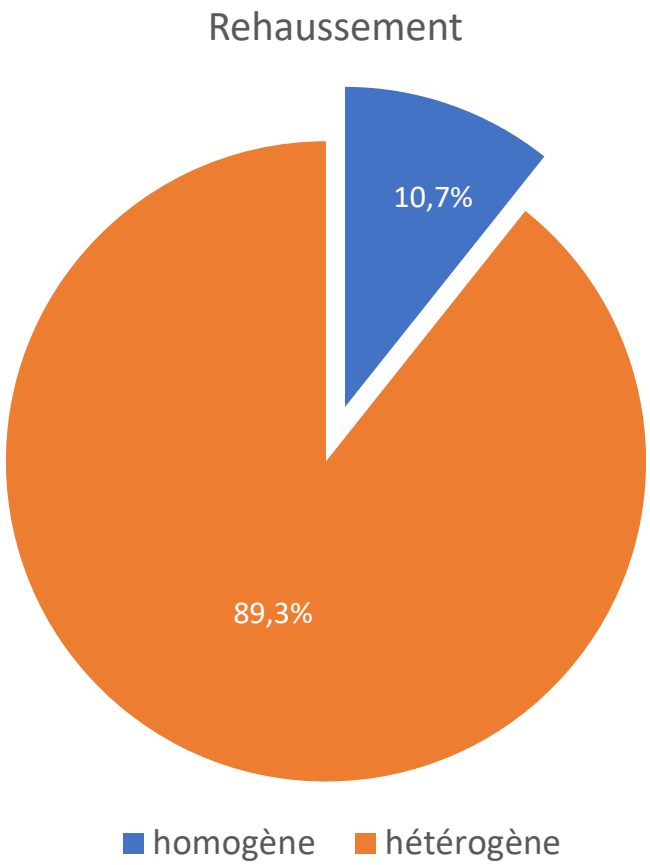



**Figure 15** : Répartition des lésions en fonction du type de rehaussement

**Figure 16** : Femme de 47 ans, non tabagique, chez qui le scanner thoracique en fenêtre médiastinale avant injection (a) et après injection du PDC en coupe axiale (c) et coronale (d) et en fenêtre parenchymateuse (b) : montre une masse parenchymateuse pulmonaire LID, périphérique; de densité solide, de contours spiculés, **rehaussée de façon hétérogène après contraste** (c,d), associé à une rétraction et attachement pleurale.

### n) Cavitation :

9,4% des tumeurs présentaient une cavitation. Tandis que la majorité des tumeurs n'étaient pas excavées (90,6%).

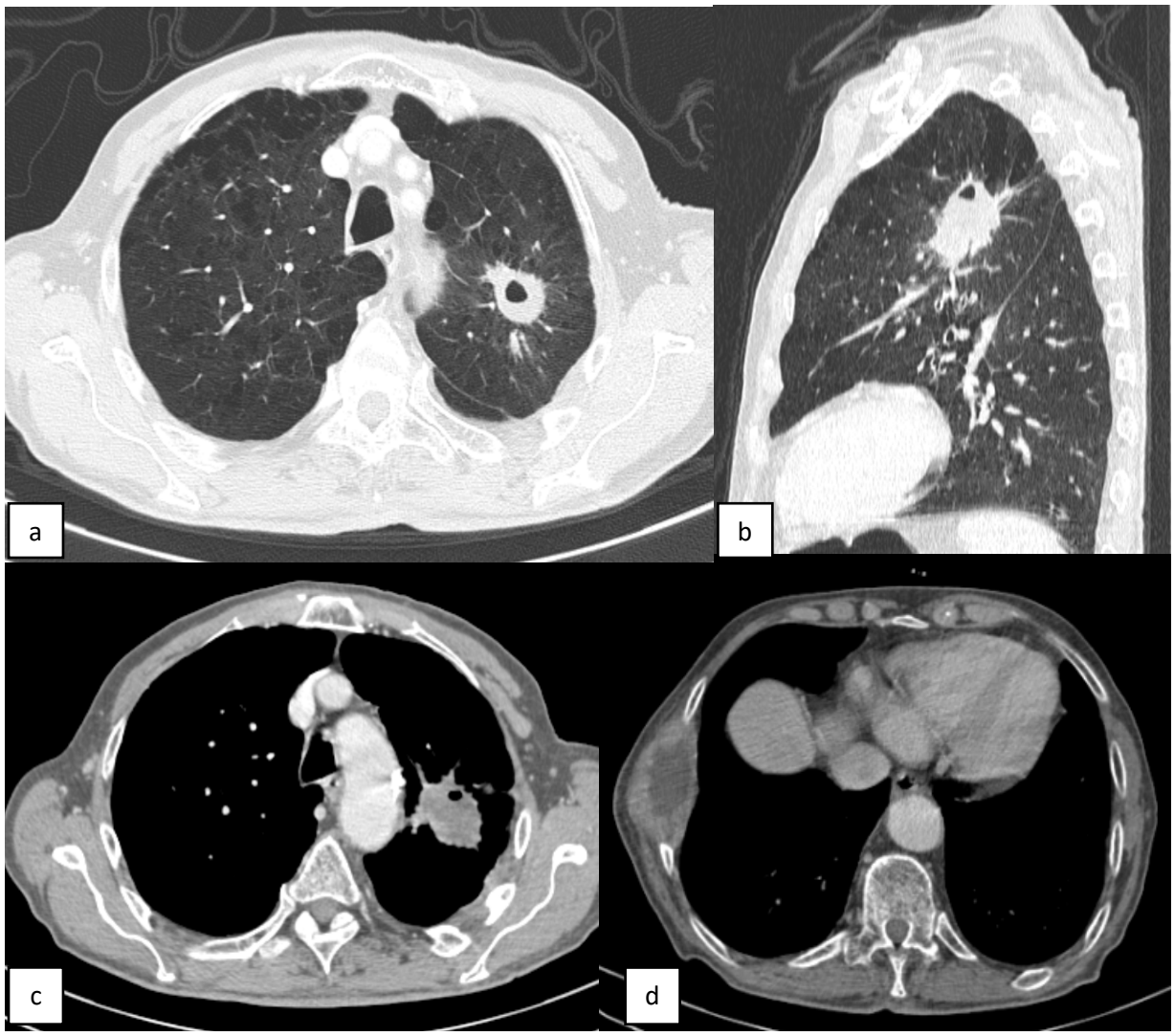



**Figure 17** : Homme de 84 ans, tabagique, chez qui le scanner thoracique en fenêtre médiastinale (c,d) et en fenêtre parenchymateuse (a,b) : montre une masse parenchymateuse pulmonaire LSG, périphérique; **excavée**, de densité solide, de contours spiculés, rehaussée de façon hétérogène après contraste délimitant des zones de nécrose, sur poumon emphysémateux, associé à une localisation secondaire osseuse.

### o) Emphysème :

65 patients présentaient un poumon emphysémateux soit un pourcentage de 43,6%.

### p) Calcifications :

Les calcifications était présente dans 22,8% des cas.

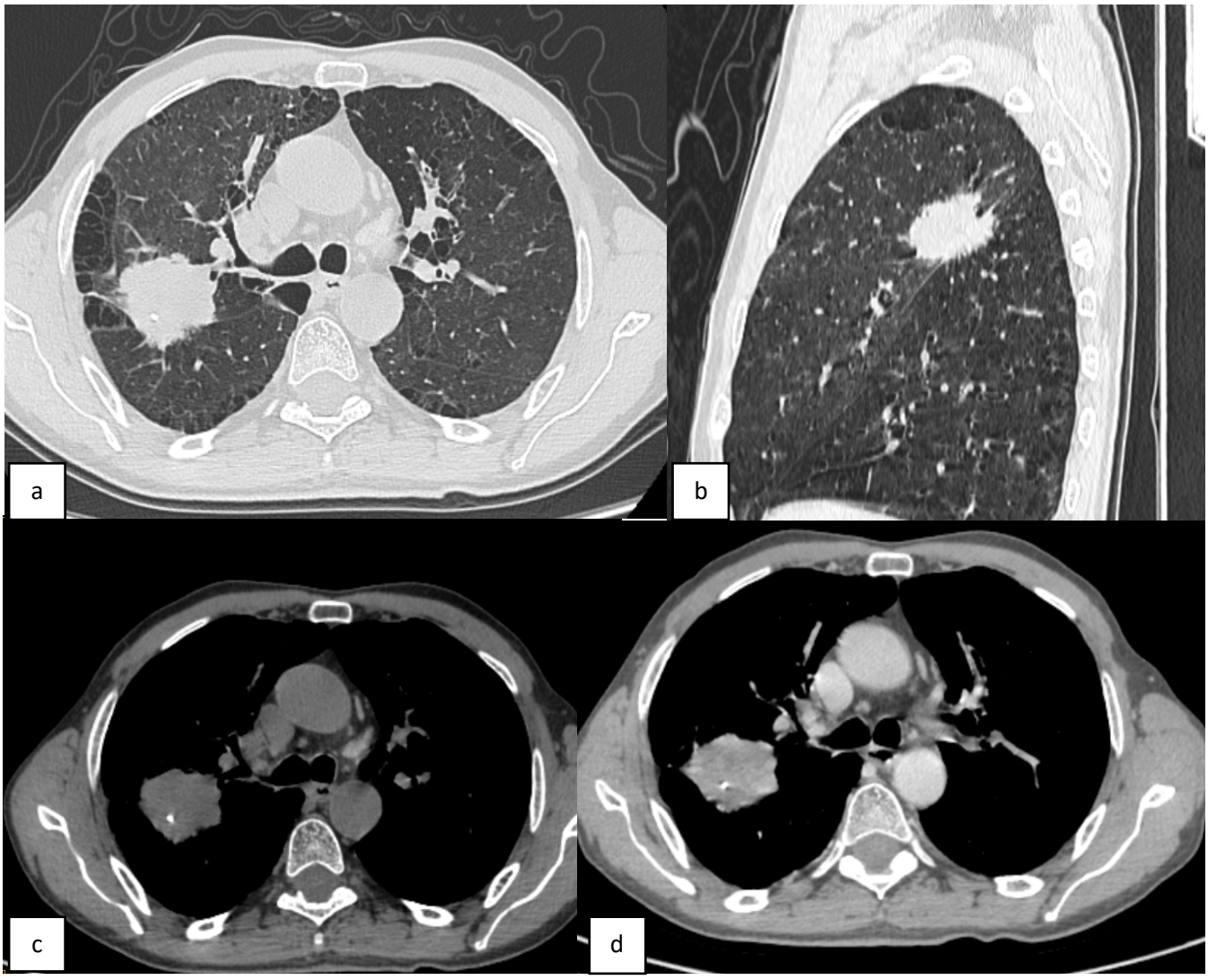



**Figure 18** : Homme de 42 ans, tabagique, chez qui le scanner thoracique en fenêtre médiastinale (c,d) et en fenêtre parenchymateuse (a,b) : montre une masse parenchymateuse pulmonaire LSD, périphérique; de densité solide, **contenant une calcification**, de contours spiculés, envahissant la grande scissure, rehaussée de façon hétérogène après contraste, sur poumon emphysémateux.

### q) Fibrose :

6% des patients présentaient une fibrose pulmonaire au moment du diagnostic de la tumeur.

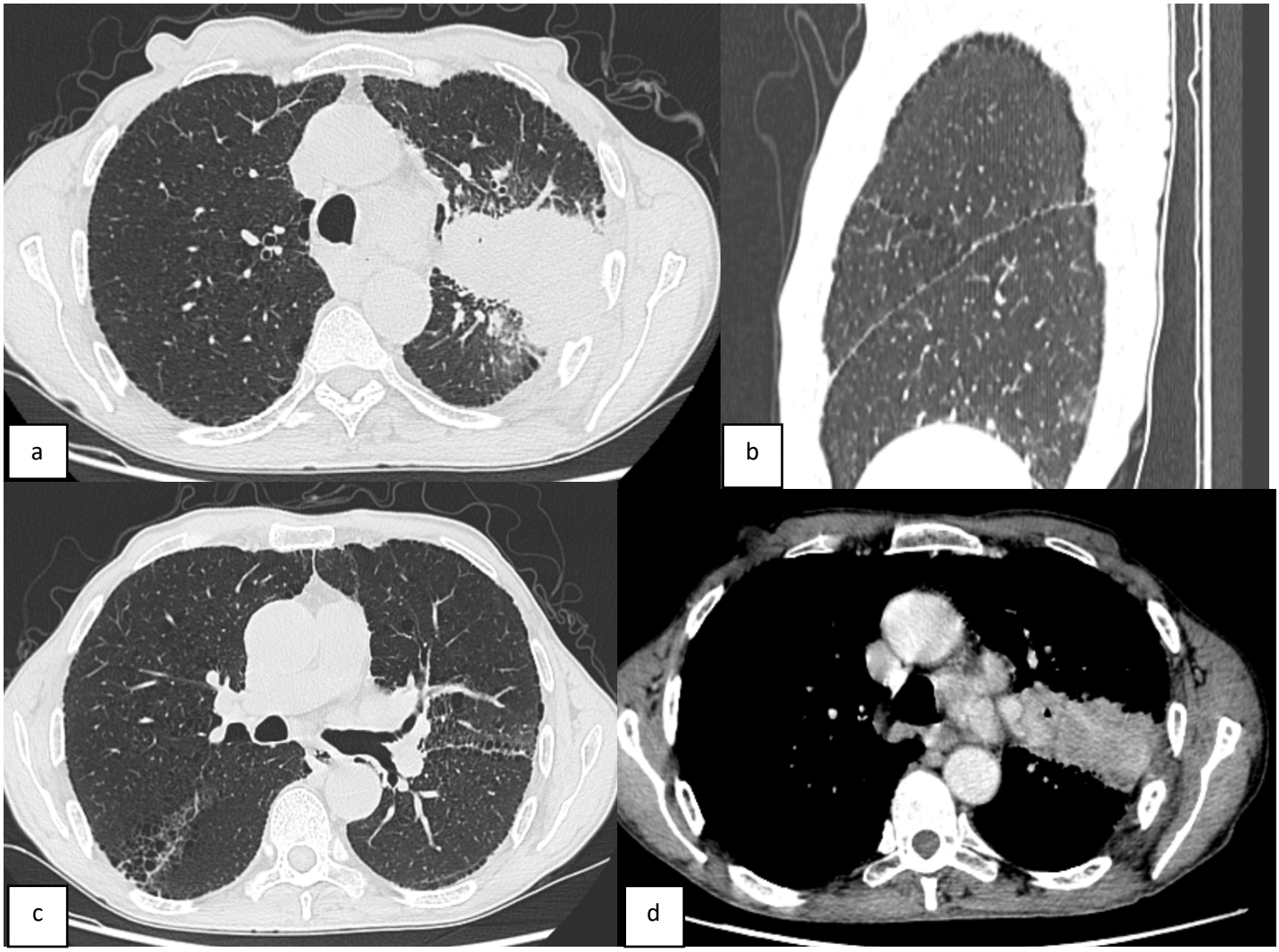



**Figure 19** : Homme de 60 ans, tabagique chronique, chez qui le scanner thoracique en fenêtre médiastinale (d) et en fenêtre parenchymateuse (a,b,c) : montre une masse parenchymateuse pulmonaire LSG, de densité solide, de contours irréguliers, envahissant la scissure, rehaussée de façon hétérogène après contraste. Il s'y associe la présence de multiples images kystiques jointives périphériques et scissurale réalisant l'aspect en rayon de miel en rapport avec une **fibrose pulmonaire**.

### r) Taille:

La majorité des tumeurs mesuraient plus de 3cm au moment du diagnostic soit un pourcentage de 87,2% des cas.

### s) Nodule pulmonaire dans le même lobe :

Presque la moitié des patients présentait au moment du diagnostic un nodule pulmonaire dans le même lobe soit un pourcentage de 43,6%.

### t) Nodule pulmonaire dans un autre lobe :

Presque la moitié des patients présentait au moment du diagnostic un nodule pulmonaire dans un autre lobe soit un pourcentage de 45%.

### u) ADP médiastinale :

La majorité des patients présentaient des ADPs médiastinales soit un pourcentage de 65,1%.

### v) Epanchement pleural homolatéral :

Presque la moitié des patients présentaient au moment du diagnostic un épanchement pleural dans le même côté atteint soit un pourcentage de 42,3%.

### w) Épanchement pleural controlatérale :

Uniquement 4,7% des patients présentaient au moment du diagnostic un épanchement pleural dans le côté controlatéral.

### x) Métastases à distance :

Plus de la moitié des tumeurs présentaient des métastases à distance au moment du diagnostic soit un pourcentage de 57,7%.

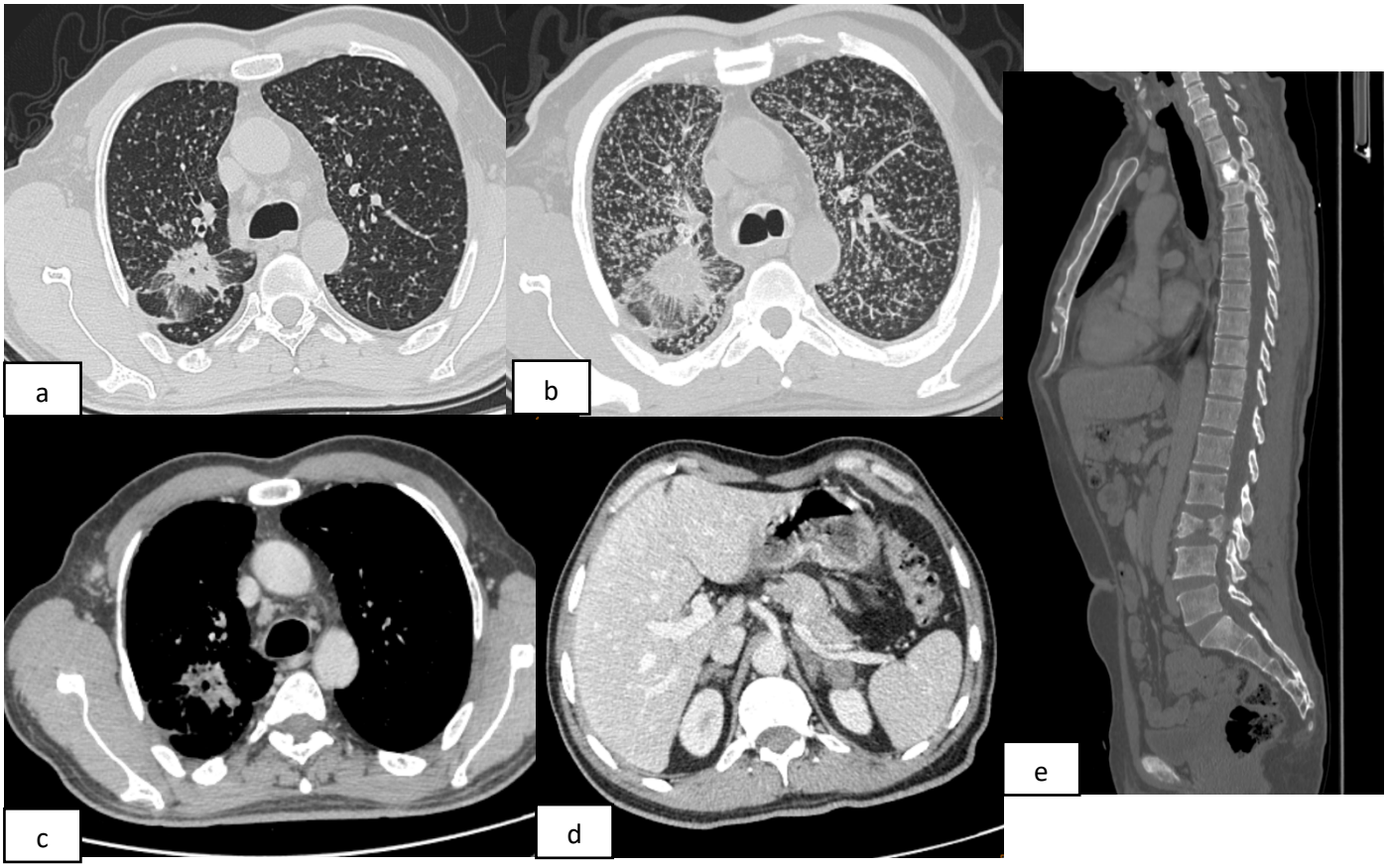



**Figure 20** homme de 58 ans, non tabagique, chez qui le scanner thoracique en fenêtre parenchymateuse (a,b), et médiastinale (c), montre une masse parenchymateuse pulmonaire LSD, à distribution centrale; de densité solide, de contours spiculés, avec bronchogramme aérique et une rétraction pleurale avec **miliaire carcinomateuse** (b) des **localisations secondaires surrénalienne gauche** sur le scanner abdominale (d) et **osseuse** (lésion lytique avec tassement vertébral de L3) sur la fenêtre osseuse (e).

## B. Étude analytique :

### 1. Les facteurs associés à la mutation EGFR par analyse univariée

#### a) Paramètres sociodémographiques:

L'âge avancé était associé à la mutation EGFR

Dans notre étude, les hommes présentaient plus de mutation EGFR (n=23) (56,1%), et cette différence est statistiquement significative (p=0,026).

Les patients non tabagiques présentaient plus de mutation EGFR (n=27) (65,9%), et cette différence est statistiquement significative (p<0,001).

**Tableau 2:** Les facteurs socio-démographiques associés à la mutation EGFR :

| **Variable** | | **EGFR** | | **p** |
|---|---|---|---|---|
| | | **Pas de mutation** | **Mutation** | |
| **Age** | | 59,84±11,09 | 64,24±10,51 | 0,030 |
| **Sexe** | | | | 0,026 |
| | Masculin | 82 (75,9) | 23 (56,1) | |
| | Féminin | 26 (24,1) | 18 (43,9) | |
| **Tabac** | | | | <0,001 |
| | Non | 36 (33,3) | 27 (65,9) | |
| | Oui | 72 (66,7) | 14 (34,1) | |

#### b) Paramètres scanographiques :

Dans l'analyse univarié , les facteurs associés à la mutation EGFR sont :

- **Contours spiculés** : la présence de contours spiculés était associée à la mutation EGFR, et cette différence était statistiquement significatif (p=0,007).
- **Attachement pleural** : la présence d'un attachement pleural était associé à la mutation EGFR, et cette différence était statistiquement significatif (p=0,006).
- **Rehaussement** : la présence d'un rehaussement hétérogène était associée à la mutation EGFR, et cette différence était statistiquement significatif (p<0,001).

- **Emphysème** : l'absence de l'emphysème était associée à la mutation EGFR, et cette différence était statistiquement significatif (p=0,001).
- **Nodule pulmonaire dans le même lobe** : la présence d'un nodule pulmonaire dans le même lobe que la tumeur primitive était associée à la mutation EGFR, et cette différence était statistiquement significatif (p=0,02).

**Tableau 3** : Les paramètres scanographiques associés à la mutation EGFR par analyse univariée :

| **Variable** | | **EGFR** | | ***p*** |
|---|---|---|---|---|
| | | **Pas de mutation** | **Mutation** | |
| **Localisation** | | | | 0,78 |
| | Centrale | 63 (58,3) | 23 (56,1) | |
| | Périphérique | 44 (40,7) | 18 (43,9) | |
| **Lobes** | | | | 0,35 |
| | LI | 26 (24,1) | 15 (36,6) | |
| | LM | 1 (0,9) | 0 | |
| | LS | 78 (72,2) | 24 (58,5) | |
| | LS + LI | 3 (2,8) | 2 (4,9) | |
| **Contours** | | | | 0,091 |
| | Irréguliers | 102 (94,4) | 35 (85,4) | |
| | Réguliers | 6 (5,6) | 6 (14,6) | |
| **Spiculations** | | | | **0,007** |
| | Non | 79 (73,1) | 20 (48,8) | |
| | Oui | 29 (26,9) | 21 (51,2) | |
| **Forme** | | | | 0,26 |
| | Arrondie | 26 (24,1) | 15 (36,6) | |
| | Ovalaire | 29 (26,9) | 6 (14,6) | |
| | Indéfinit | 52 (48,1) | 20 (48,8) | |
| **Limite** | | | | 0,85 |
| | Bien limitée | 55 (50,9) | 22 (53,7) | |
| | Mal limitée | 53 (49,1) | 19 (46,3) | |

| | | | | |
|---|---|---|---|---|
| **Densité** | | | | 1 |
| | Solide | 97 (89,8) | 37 (90,2) | |
| | Mixte | 11 (10,2) | 4 (9,8) | |
| **Bronchogramme aérique** | | | | 0,71 |
| | Oui | 43 (39,8) | 18 (43,9) | |
| | Non | 65 (60,2) | 23 (56,1) | |
| **Attachement scissural** | | | | 0,44 |
| | Oui | 71 (65,7) | 24 (58,5) | |
| | Non | 37 (34,3) | 17 (41,5) | |
| **Attachement pleural** | | | | **0,006** |
| | Oui | 88 (81,5) | 24 (58,5) | |
| | Non | 20 (18,5) | 17 (41,5) | |
| **Rétraction scissurale** | | | | 1 |
| | Oui | 24 (22,2) | 9 (22) | |
| | Non | 84 (77,8) | 32 (78) | |
| **Rétraction pleurale** | | | | 0,85 |
| | Oui | 39 (36,1) | 16 (39) | |
| | Non | 69 (63,9) | 25 (61) | |
| **Rehaussement** | | | | **<0,001** |
| | Homogène | 2 (1,9) | 14 (34,1) | |
| | Hétérogène | 106 (98,1) | 27 (65,9) | |
| **Cavitation** | | | | 1 |
| | Oui | 10 (9,3) | 4 (9,8) | |
| | Non | 98 (90,7) | 37 (90,2) | |
| **Emphysème** | | | | **0,001** |
| | Oui | 56 (51,9) | 9 (22) | |
| | Non | 52 (48,1) | 32 (78) | |
| **Calcifications** | | | | 0,27 |
| | Oui | 22 (20,4) | 12 (19,3) | |

| | | | | |
|---|---|---|---|---|
| | Non | 86 (79,6) | 29 (70,7) | |
| **Fibrose** | | | | 0,5 |
| | Oui | 7 (6,5) | 1 (2,4) | |
| | Non | 100 (92,6) | 40 (97,6) | |
| **Taille maximale** | | | | 0,78 |
| | < 3cm | 13 (12,3) | 4 (9,8) | |
| | > 3cm | 93 (87,7) | 37 (90,2) | |
| **Nodule pulmonaire dans le même lobe** | | | | **0,02** |
| | Oui | 41 (38) | 24 (58,5) | |
| | Non | 67 (62) | 17 (41,5) | |
| **Nodule pulmonaire dans au autre lobe** | | | | 0,2 |
| | Oui | 45 (41,7) | 22 (53,7) | |
| | Non | 63 (58,3) | 19 (46,3) | |
| **ADP médiastinale** | | | | 1 |
| | Oui | 70 (64,8) | 27 (65,9) | |
| | Non | 38 (35,2) | 14 (34,2) | |
| **Épanchement pleural homolatéral** | | | | 0,58 |
| | Oui | 44 (40,7) | 19 (46,3) | |
| | Non | 64 (59,3) | 22 (53,7) | |
| **Épanchement pleural controlatéral** | | | | 0,39 |
| | Oui | 4 (3,7) | 3 (7,3) | |
| | Non | 104 (96,3) | 38 (92,7) | |
| **Métastase à distance** | | | | 0,13 |
| | Oui | 58 (53,7) | 28 (68,3) | |
| | Non | 50 (46,3) | 13 (31,7) | |

### 2. Les facteurs associés à la mutation EGFR par analyse multivariée

L'analyse multivariée par régression logistique a mis en évidence une association de la mutation EGFR avec les contours irréguliers (ORA=6,43 ; IC95% 1,26 ;32,78), avec la présence des spiculations (ORA=5,81 ; IC95% 1,96 ;17,22), avec l'attachement pleural (ORA=3,53 ; IC95% 1,32 ;9,42), avec le rehaussement hétérogène (ORA=30,679 ; IC95% 5,149 ;182,778), avec l'absence de l'emphysème (ORA=4,815 ; IC95% 1,966 ;17,220), et avec la présence de métastases à distance (ORA=4,123 ; IC95% 1,373 ;12,383).

**Tableau 4** : Les facteurs associés à la mutation EGFR par analyse multivariée :

| | *p* | OR ajusté | Intervalle de confiance 95% pour ORA | |
|---|---|---|---|---|
| | | | Inférieur | Supérieur |
| **Contours** | | | | |
| Régulier | - | 1 | - | - |
| Irrégulier | 0,025 | 6,438 | 1,264 | 32,785 |
| **Spiculation** | | | | |
| Non | - | 1 | - | - |
| Oui | 0,001 | 5,819 | 1,966 | 17,220 |
| **Attachement pleural** | | | | |
| Non | - | 1 | - | - |
| Oui | 0,012 | 3,535 | 1,327 | 9,421 |
| **Rehaussement** | | | | |
| Homogène | - | 1 | - | - |
| Hétérogène | 0,000 | 30,679 | 5,149 | 182,778 |
| **Métastase à distance** | | | | |
| Non | - | 1 | - | - |
| Oui | 0,012 | 4,123 | 1,373 | 12,383 |
| **Emphysème** | | | | |
| Oui | - | 1 | - | - |
| Non | 0,005 | 4,815 | 1,966 | 17,220 |

## II. Segmentation des tumeurs pulmonaires :

Pour l'apprentissage et la validation, 5 838 images provenant des différents scanner ont été utilisées. Toutes les images de tumeurs pulmonaires provenant du même scanner ont été utilisées pour la randomisation. Le réseau neuronal a été formé et validé en utilisant la validation croisée à cinq facteurs. Lorsqu'on a évalué les masques CNN créés automatiquement de la tumeur pulmonaire, on a trouvé des résultats précis (figure 21). Le score de dé pour le CNN CHU-Net était de 0,936. Le coefficient d'intersection sur l'union était de 0,981.

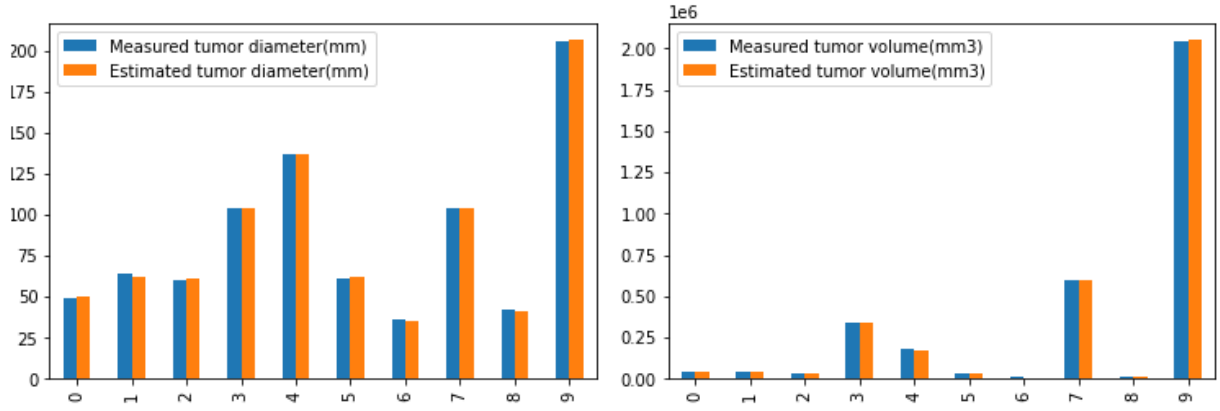



**Figure 21** : comparaison entre la mesure manuelle du volume et du diamètre tumoral par le radiologue (bleu) et la mesure automatique.

La validation croisée à 5 intervalles a été adoptée. Ainsi, l'ensemble des données de formation a été divisé en cinq intervalles de données égaux, dont quatre ont été utilisés pour l'apprentissage et le cinquième pour la validation. Cela nous a permis de donner les mesures des différentes performances de notre modèle sous la forme d'une valeur moyenne associée à un écart-type. Le processus d'optimisation des valeurs des différents poids de notre architecture a été réalisé grâce à l'algorithme d'Adam.

Un exemple du résultat de la segmentation d'une tumeur pulmonaire effectuée par chaque sous-réseau ainsi que par notre modèle de piquetage MSCNN est donné dans la figure 22.

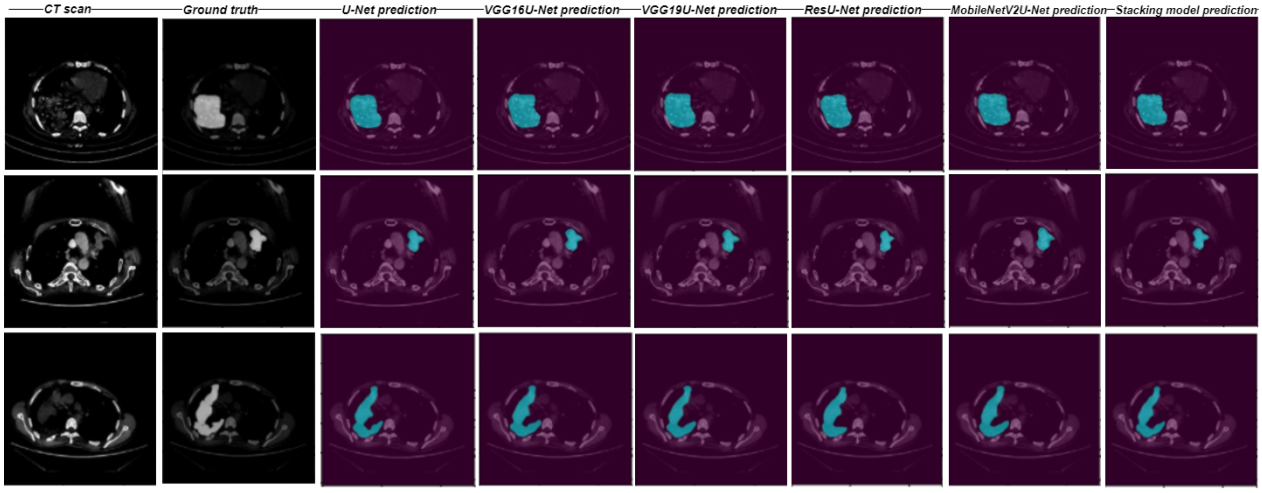



**Figure 22** : Comparaisons de segmentation. De gauche à droite, les colonnes représentent l'image originale, le masque original, les prédictions U-Net, les prédictions VGG16U-Net, les prédictions VGG19U-Net, les prédictions ResU-Net, les prédictions MobileNetV2U-Net et les prédictions du modèle d'empilement.

Les figures 23, 24, 25, 26, 27, 28, 29 et 30 montrent des exemples 3D illustrant le degré de ressemblance entre la segmentation manuelle et automatique.

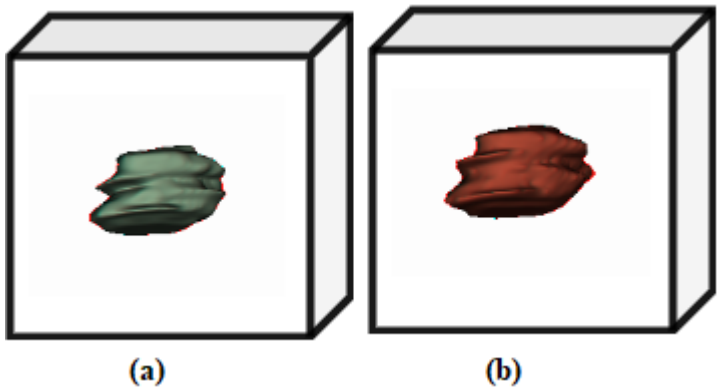



**Figure 23** : (a) segmentation manuelle et (b) segmentation automatique

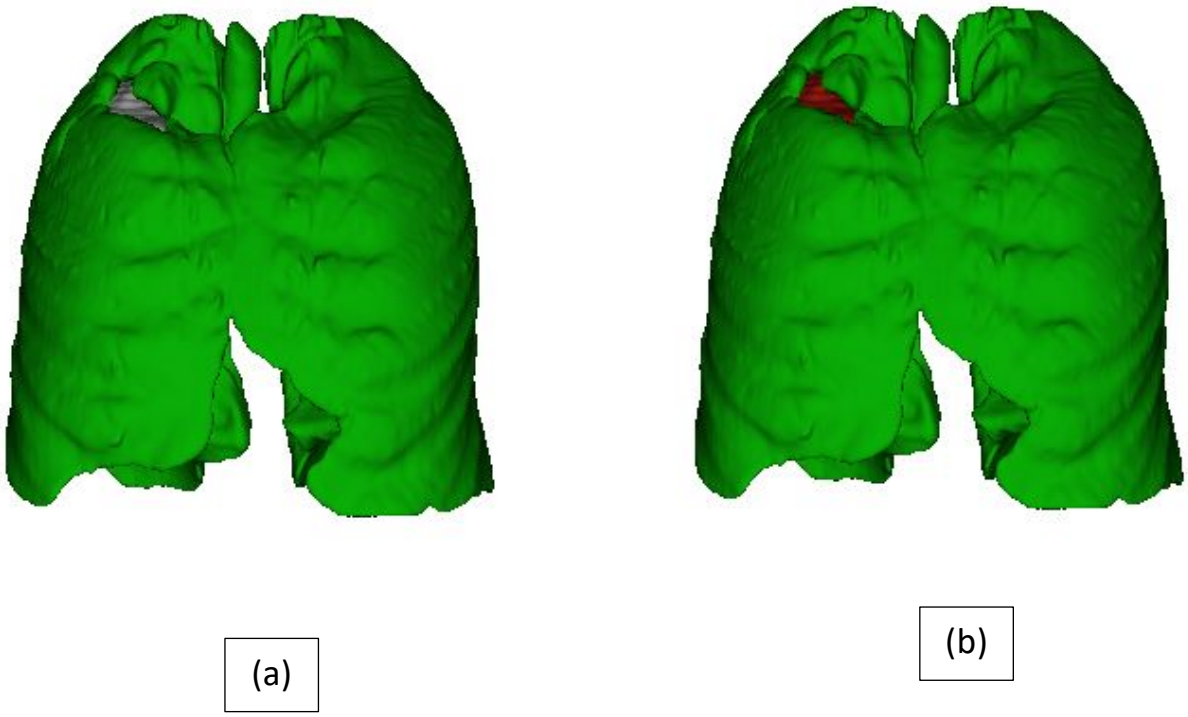



**Figure 24** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique apicale droite

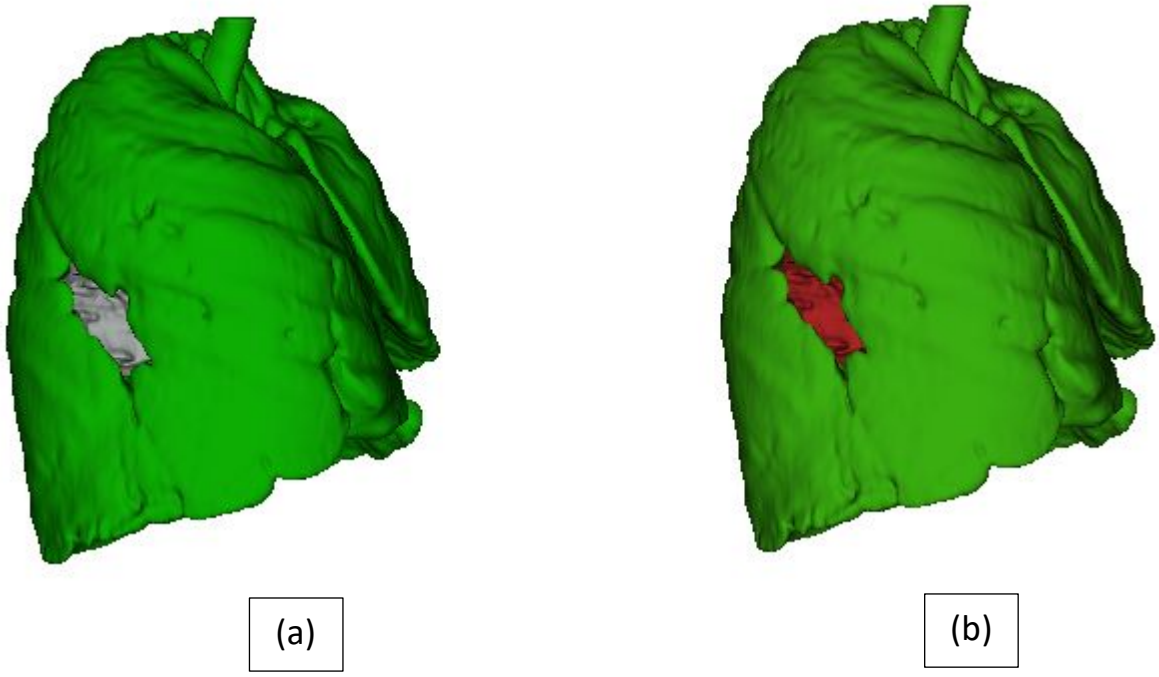



**Figure 25** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique lobaire moyenne

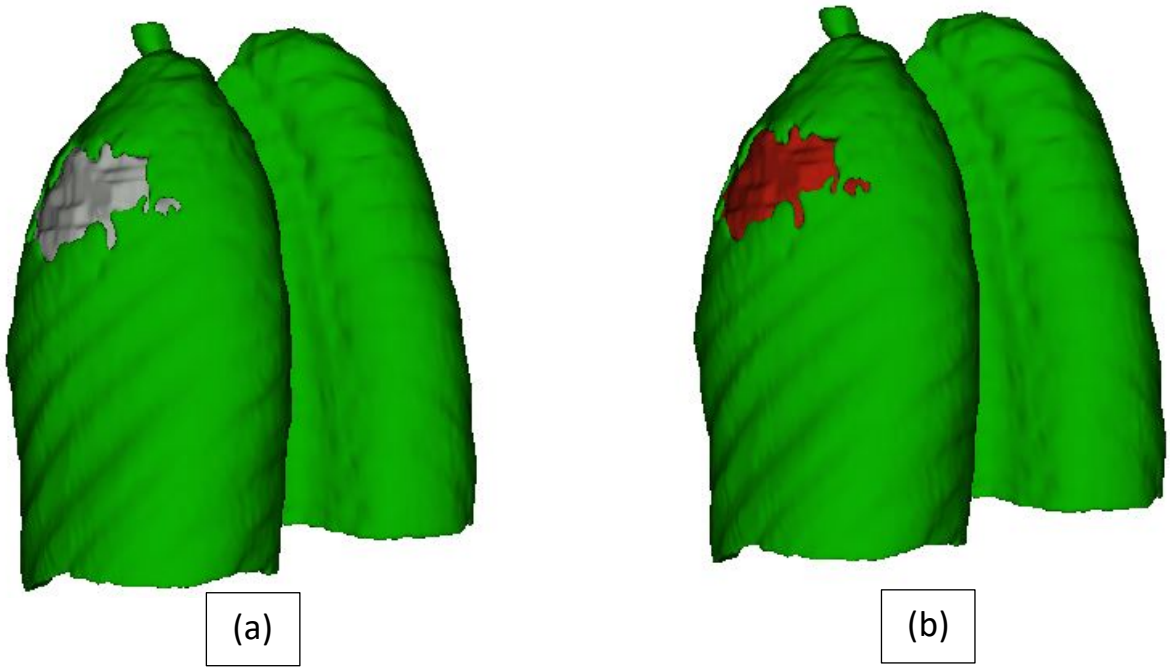



**Figure 26** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique apicale gauche

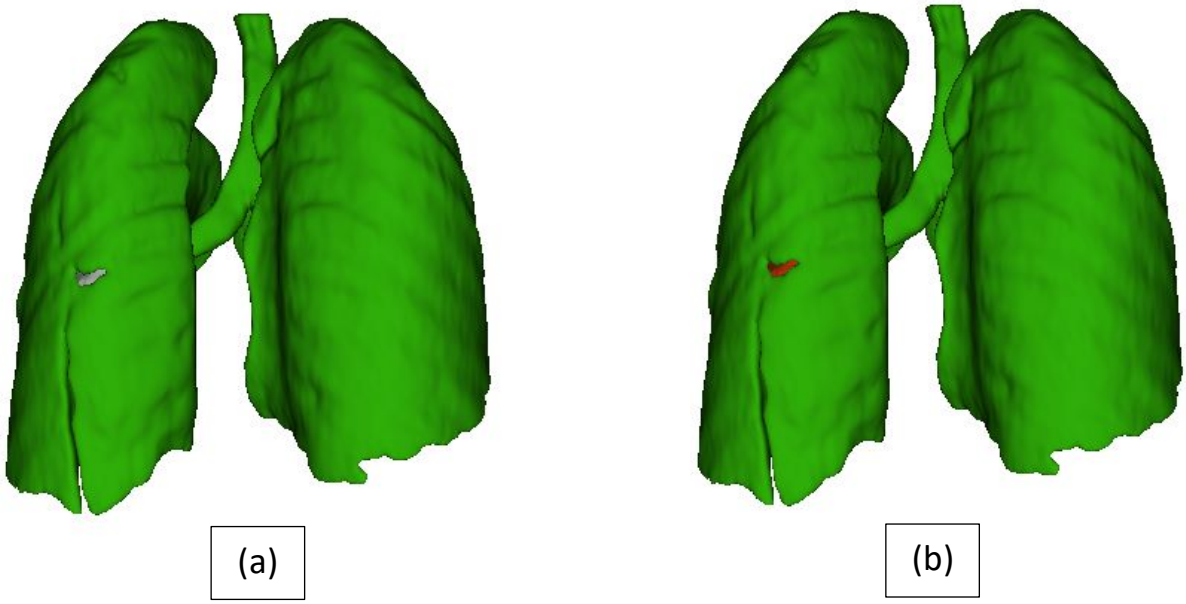
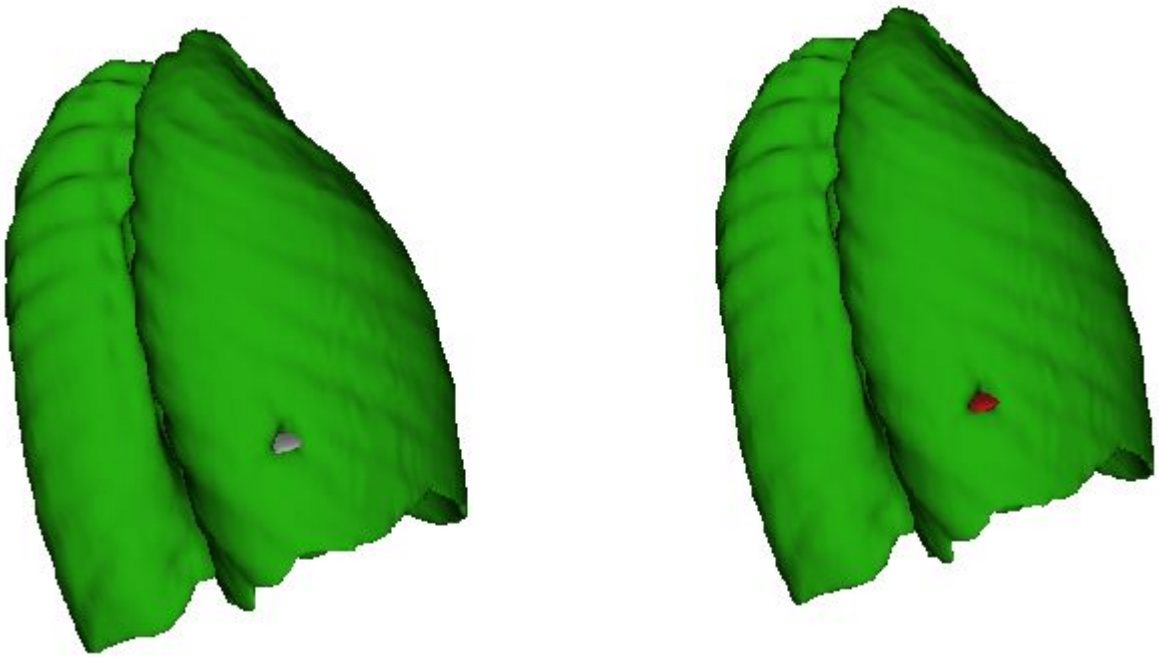
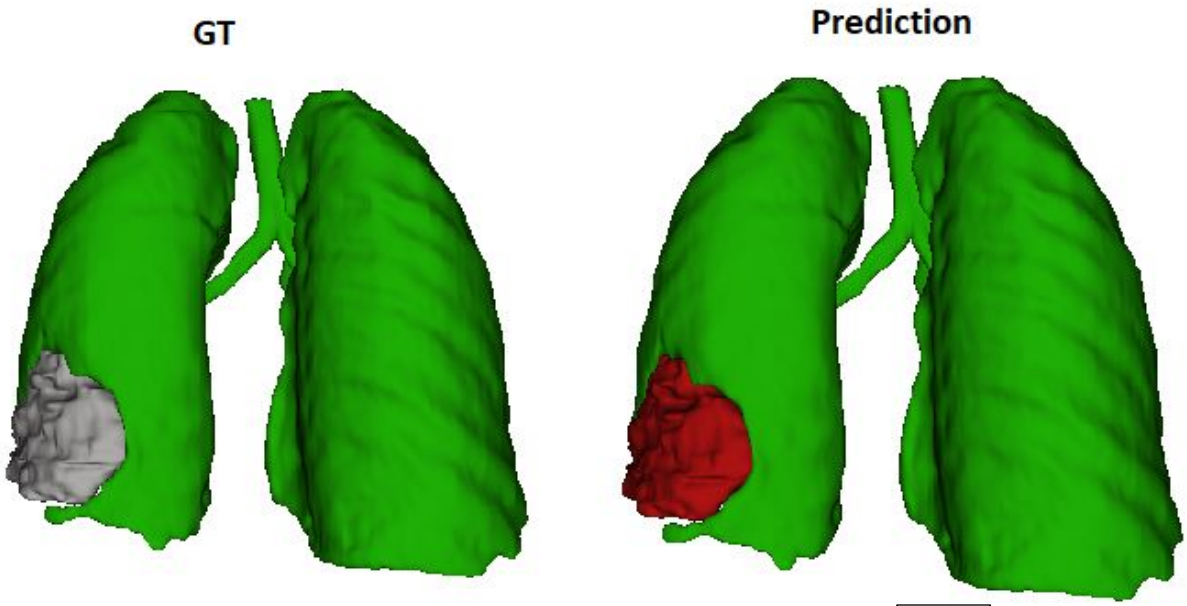



**Figure 27** (a) segmentation manuelle et (b) segmentation automatique d'une petite tumeur pulmonaire périphérique lobaire moyenne.



(a) (b)

**Figure 28** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique lobaire inférieure droite




(a) (b)

**Figure 29** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique lobaire inférieure droite

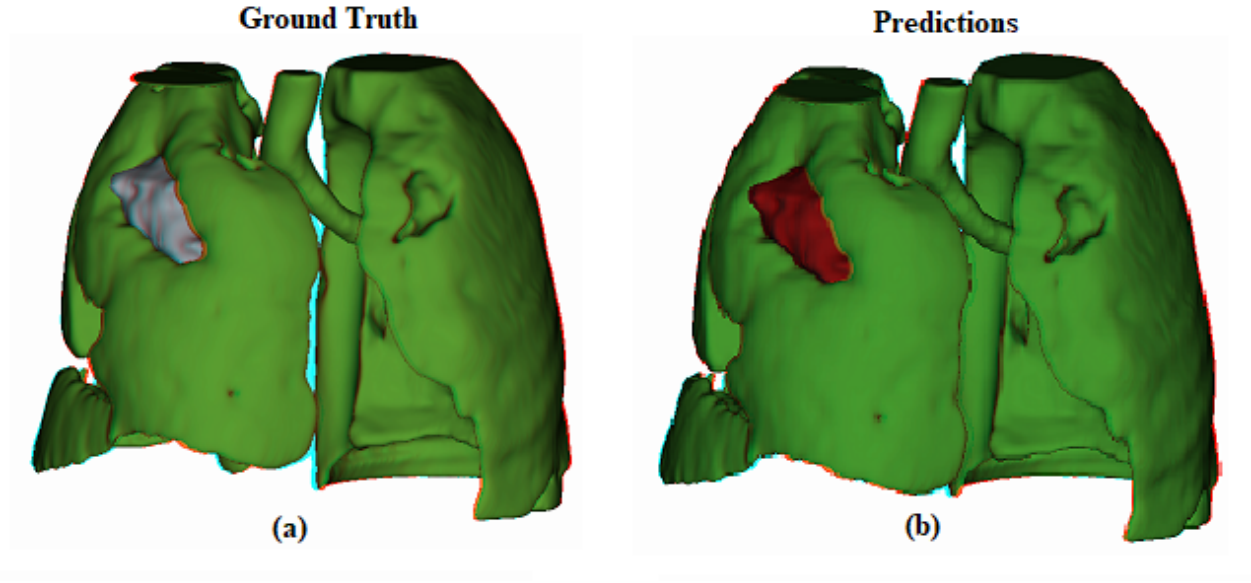



**Figure 30** (a) segmentation manuelle et (b) segmentation automatique d'une tumeur pulmonaire périphérique apicale droite

Nous avons utilisé trois mesures pour valider le processus de segmentation de notre modèle proposé : Le coefficient de similarité de Dice, le coefficient de similarité de Jaccard et le coefficient d'intersection sur l'union.
Ils ont été mesurés à chaque validation croisée et à chaque période des processus d'apprentissage et de validation.
Nous avons présenté uniquement les résultats de mesure du coefficient de Dice obtenus pour l'architecture ResU-Net.

La figure 31 présente les courbes de perte de formations obtenues en mesurant la moyenne du coefficient de Dice avec une validation croisée 5 fois du réseau ResU-Net lors du processus d'apprentissage et du processus de validation.

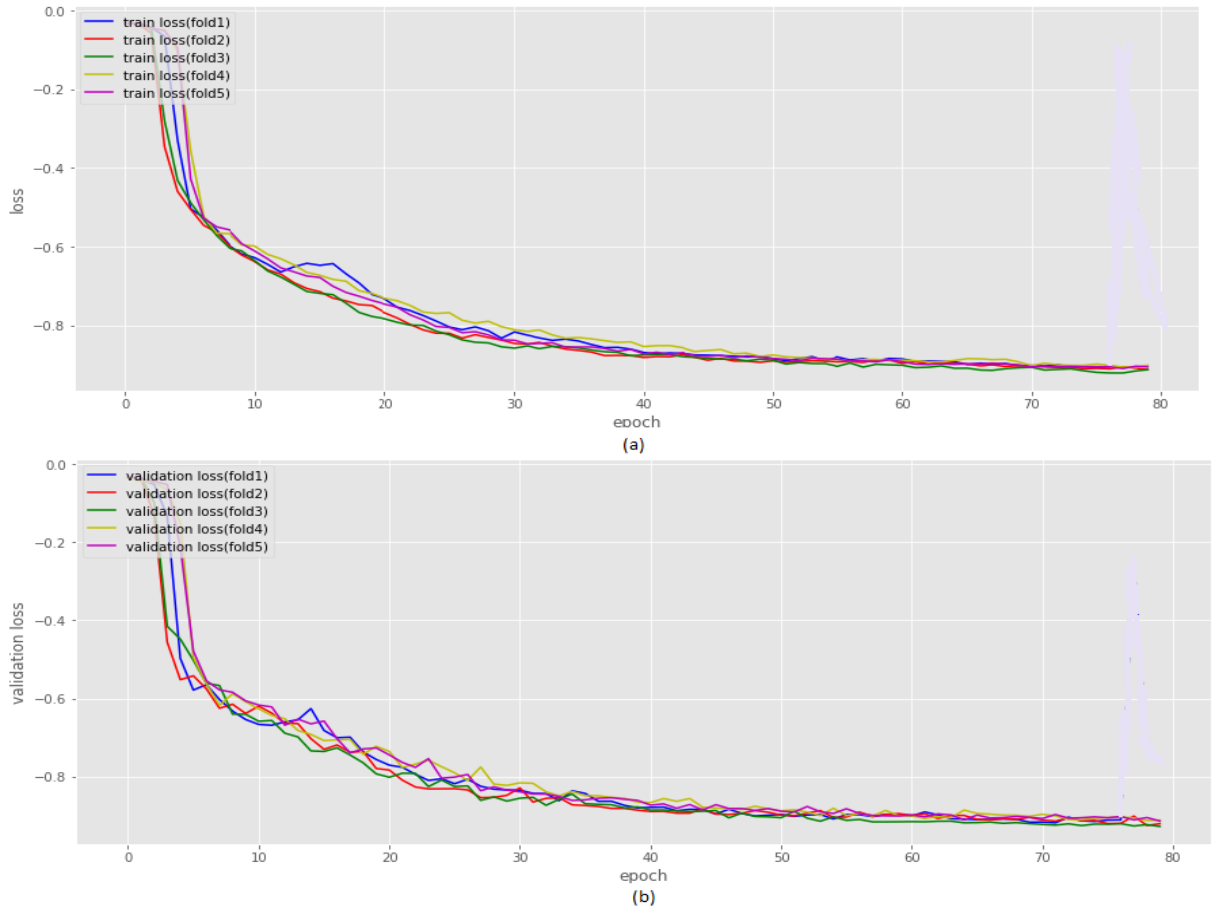



**Figure 31** : Courbes de perte de formation (a) et de validation (b) de la validation croisée (cinq fois).

La figure 32 présente les courbes de mesure de moyenne du coefficient de Dice avec une validation croisée 5 fois du réseau ResU-Net pendant le processus d'apprentissage et le processus de validation.

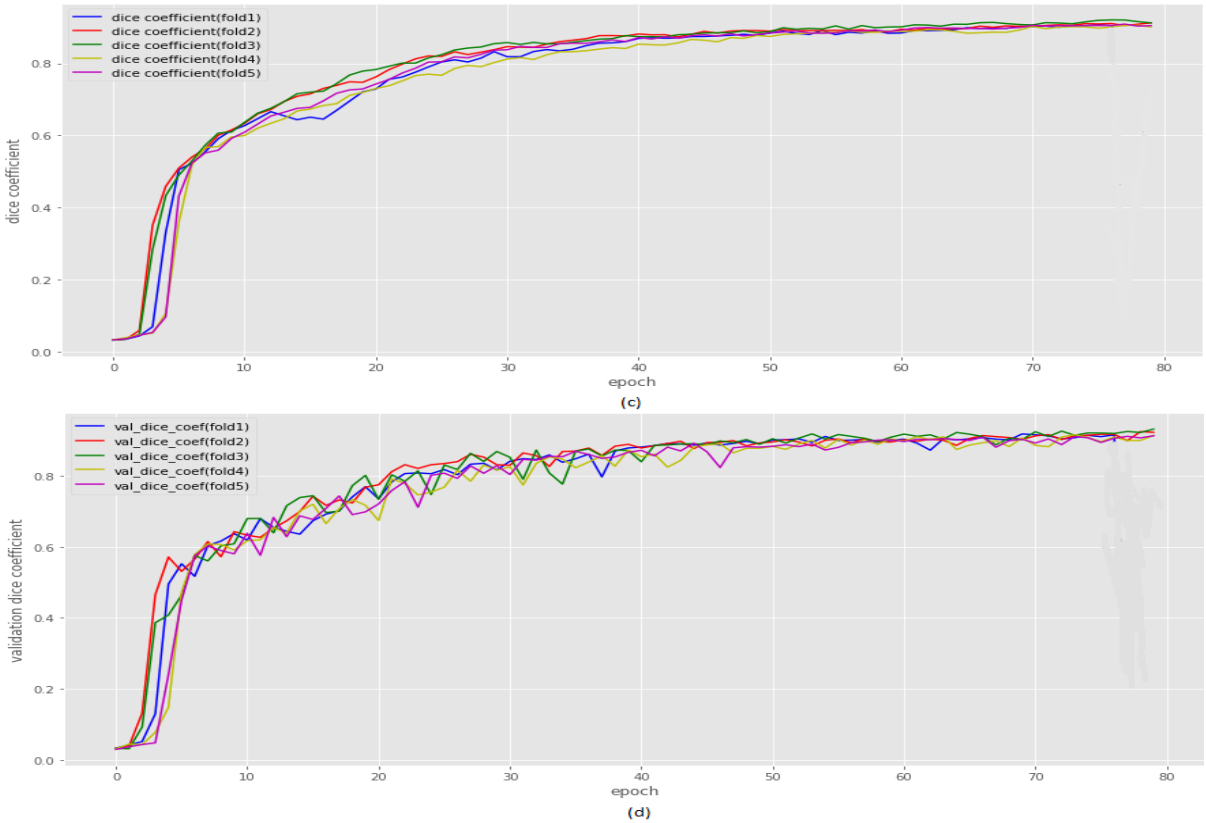



**Figure 32** : les courbes de Dice de formation (a) et de validation (b) de la validation croisée à cinq facteurs.

Nous avons comparé les mesures des trois critères d'évaluation : le coefficient de Dice, le coefficient de Jaccard et l'intersection, les résultats obtenues pour chaque modèle pris individuellement, avec les mesures obtenues avec notre modèle d'empilement. L'ensemble des mesures obtenues est présenté dans le tableau 5, qui montre la supériorité de notre modèle d'empilement proposé par rapport aux autres sous-modèles.

**Tableau 5** : comparaison des résultats des trois coefficients avec la validation croisée des cinq modèles.

| Modèle | Score de Dice [%] | Coefficient de Jaccard [%] | coefficient d'intersection sur l'union [%] |
|---|---|---|---|
| U-Net | 84.63 | 83.43 | 97.41 |
| VGG16U-Net | 88.74 | 89.78 | 98.37 |
| VGG19U-Net | 89.45 | 90.02 | 98.82 |
| ResU-Net | 86.67 | 86.96 | 97.72 |
| MobileNetV2U-Net | 89.06 | 89.27 | 98.17 |
| **Modèle Stacking proposé MSCNN** | **94.02** | **93.88** | **98.85** |

Ensuite, nous avons comparé la moyenne des trois critères d'évaluation : le coefficient de Dice, le coefficient de Jaccard et l'intersection. Les résultats obtenus sont présentés dans le tableau 6 et figure 20. Nous constatons que malgré les bonnes performances des réseaux VGG19U-Net et MobileNetV2U-Net, les résultats obtenues par notre modèle de staking restent meilleurs avec un coefficient de Dice de 93,64 et un coefficient d'Intersection sur Union de 98,91.

**Tableau 6** : comparaison des résultats des moyennes des trois critères d'évaluation.

| Model | Score de Dice moyen | Coefficient de Jaccard moyen [%] | coefficient d'intersection sur l'union moyen [%] |
|---|---|---|---|
| *U-Net* | *84.58 [*±0.007] | *82.02[*±0.032] | *97.07[*±0.002] |
| *VGG16U-Net* | *89.44[*±0.003] | *89.02[*±0.002] | *98.49[*±0.015] |
| *VGG19U-Net* | *90.13[*±0.003] | *90.53[*±0.001] | *98.84[*±0.011] |
| *ResU-Net* | *82.07 [*±0.008] | *81.89[*±0.008] | *97.92[*±0.006] |
| *MobileNetV2U-Net* | *89.69[*±0.009] | *88.85[*±0.014] | *98.83[*±0.007] |
| *Modèle Stacking* | ***93.64*** | ***92.86*** | ***98.91*** |

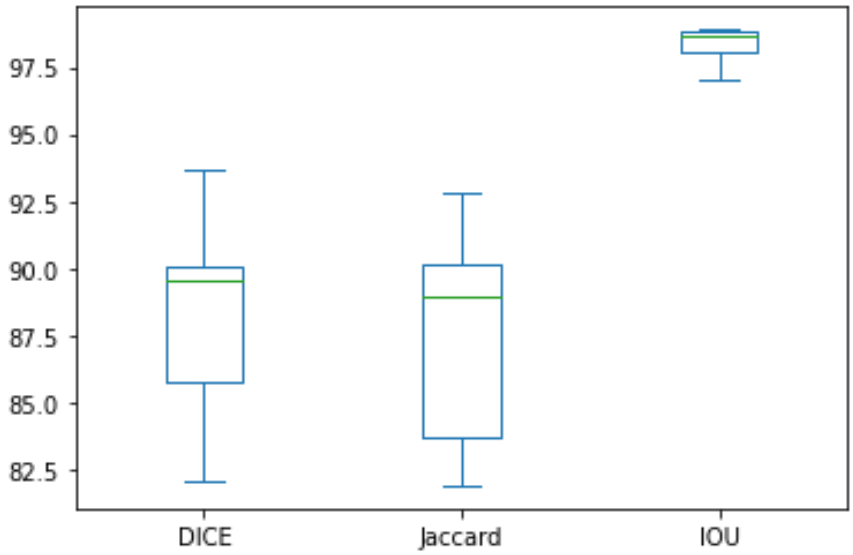



**Figure 33** : comparaison des résultats des moyennes des trois critères d'évaluation.

Afin de démontrer l'efficacité de notre modèle de piquetage dans la mesure de la taille volumétrique de la tumeur, nous avons sélectionné au hasard dix patients, pour lesquels, nous avons comparé le diamètre et le volume de la tumeur mesurés automatiquement par notre modèle à ceux mesurés manuellement par nos radiologues. Nous avons utilisé la moyenne absolue du pourcentage d'erreur MAPE comme critère d'évaluation, (elle est définie comme la moyenne des écarts en valeur absolue par rapport aux valeurs observées). Les résultats obtenus sont présentés dans le tableau 7. Nous constatons que les valeurs mesurées automatiquement sont concordantes avec les valeurs mesurées manuellement.

**Tableau 7** : comparaison des diamètres et des volumes tumoraux mesurés automatiquement, mesurés par le radiologue et avec la moyenne du pourcentage d'erreur MAPE chez 10 patients choisis au hasard.

| Patients | Mesure automatique du diametre tumoral [mm] | Mesure automatique du volume tumoral [$mm^3$] | Mesure du diametre tumoral Par le médecin [mm] | Mesure du volume tumoral Par le médecin [$mm^3$] | Mesure du diamètre tumoral MAPE | Mesure du volume tumoral MAPE |
|---|---|---|---|---|---|---|
| 1,00 | 49,71 | 43240,00 | 50,61 | 45380,00 | 0,02 | 0,05 |
| 2,00 | 64,52 | 40890,00 | 62,58 | 39060,00 | 0,03 | 0,04 |
| 3,00 | 60,63 | 32880,00 | 61,65 | 31560,00 | 0,02 | 0,04 |
| 4,00 | 104,30 | 338800,00 | 103,98 | 336100,00 | 0,00 | 0,01 |
| 5,00 | 137,40 | 179700,00 | 136,90 | 177400,00 | 0,00 | 0,01 |
| 6,00 | 61,52 | 31030,00 | 61,83 | 32000,00 | 0,01 | 0,03 |
| 7,00 | 36,75 | 10460,00 | 35,77 | 9080,00 | 0,03 | 0,13 |
| 8,00 | 104,50 | 601800,00 | 104,00 | 600700,00 | 0,00 | 0,00 |
| 9,00 | 42,23 | 14380,00 | 41,65 | 13130,00 | 0,01 | 0,09 |
| 10,00 | 205,90 | 2048000,00 | 206,50 | 2050000,00 | 0,00 | 0,00 |

# III. Radiomics :

1532 caractéristiques radiomics ont été extraites à partir des images scanographiques de chaque patient pour prédire les statuts mutationnels d'EGFR. Ceux-ci comprenaient les caractéristiques suivantes :

## Classes des caractéristiques

Statistiques du premier ordre

Basé sur la forme (2D et 3D)

Matrice de cooccurrence de niveau de gris (GLCM)

Matrice de longueur de course du niveau de gris (GLRLM)

Matrice des zones de taille du niveau de gris (GLSZM)

Matrice de dépendance des niveaux de gris (GLDM)

Matrice de différence des tons de gris voisins (NGTDM)

## Classes de filtres

En plus des classes de caractéristiques, il existe également des filtres optionnels intégrés :

Laplacien de Gauss (LoG, basé sur la fonctionnalité SimpleITK)

ondelette (en utilisant le paquet PyWavelets)

Carré

Racine carrée

Logarithme

Exponentiel

Gradient (magnitude)

Modèle binaire local (LBP) 2D / 3D

Le tableau 8 montre les caractéristiques radiomiques pertinentes restantes après l'utilisation de ***FeatureWiz***

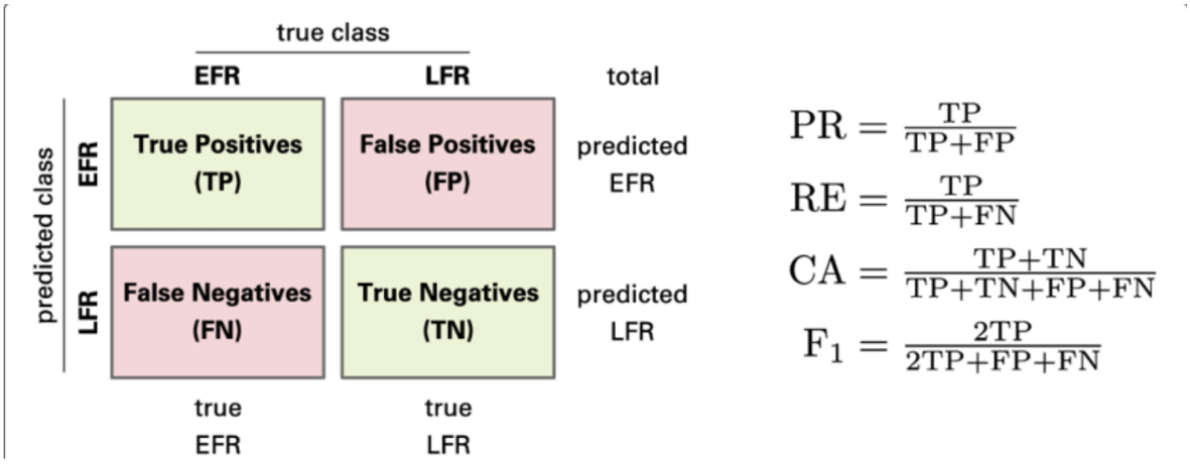
**Tableau 8** : les caractéristiques radiomiques les plus pertinentes (n=124)

| Noms des caractéristiques | Dimension | famille des caractéristiques |
|---|---|---|
| LargeAreaEmphasis | 2D | GLSZM |
| LargeAreaHighGrayLevelEmphasis | 2D | GLSZM |
| SmallAreaLowGrayLevelEmphasis | 2D | GLSZM |
| GrayLevelNonUniformityNormalized | 2D | GLSZM |

| | | |
|---|---|---|
| ZoneVariance | 2D | GLSZM |
| ZoneEntropy | 2D | GLSZM |
| GrayLevelNonUniformity | 2D | GLSZM |
| SizeZoneNonUniformity | 2D | GLSZM |
| ZoneEntropy | 3D | GLSZM |
| GrayLevelVariance | 3D | GLSZM |
| Elongation | 2D | Shape |
| Flatness | 2D | Shape |
| Sphericity | 2D | Shape |
| ShortRunLowGrayLevelEmphasis | 2D | GLRLM |
| LowGrayLevelRunEmphasis | 3D | GLRLM |
| RunLengthNonUniformityNormalized | 2D | GLRLM |
| LongRunHighGrayLevelEmphasis | 3D | GLRLM |
| HighGrayLevelRunEmphasis | 2D | GLRLM |
| LongRunEmphasis | 2D | GLRLM |
| ShortRunHighGrayLevelEmphasis | 3D | GLRLM |
| ShortRunLowGrayLevelEmphasis | 3D | GLRLM |
| GrayLevelNonUniformity | 2D | GLRLM |
| GrayLevelNonUniformityNormalized | 3D | GLRLM |
| Strength | 2D | NGTDM |
| Contrast | 2D | NGTDM |
| Strength | 3D | NGTDM |
| Busyness | 2D | NGTDM |
| Mean | 2D | First order |
| Median | 2D | First order |
| Skewness | 2D | First order |
| Uniformity | 2D | First order |
| Minimum | 2D | First order |
| Kurtosis | 2D | First order |
| Maximum | 2D | First order |
| Range | 2D | First order |
| Kurtosis | 3D | First order |
| RobustMeanAbsoluteDeviation | 2D | First order |
| 90Percentile | 3D | First order |
| ClusterShade | 2D | GLCM |

| | | |
|---|---|---|
| JointAverage | 3D | GLCM |
| Imc1 | 2D | GLCM |
| SumSquares | 2D | GLCM |
| Correlation | 2D | GLCM |
| SumSquares | 3D | GLCM |
| Idmn | 2D | GLCM |
| HighGrayLevelEmphasis | 2D | GLDM |
| LargeDependenceHighGrayLevelEmphasis | 2D | GLDM |
| SmallDependenceLowGrayLevelEmphasis | 2D | GLDM |
| DependenceNonUniformityNormalized | 2D | GLDM |
| DependenceVariance | 3D | GLDM |
| **Et 73 autres** | | |
| … | | |
| … | | |
| … | | |

Afin d'illustrer la capacité de prédiction de la mutation de l'EGFR, nous avons comparé les modèles développés à partir des caractéristiques cliniques, les caractéristiques cliniques pertinentes, les caractéristiques radiomiques et les caractéristiques radiomiques pertinentes, et de leur combinaison.
Nous avons utilisé les mesures de performances : précision (PR), recall (RE), F1 score et auccuracy (CA) comme montre le schéma ci-dessous :



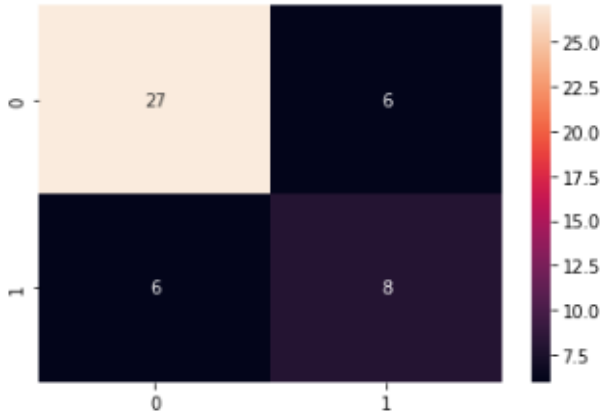

Matrice de confusion. CM illustré avec les formules de précision (PR), de rappel (RE), de précision (CA) et de F 1 mesure

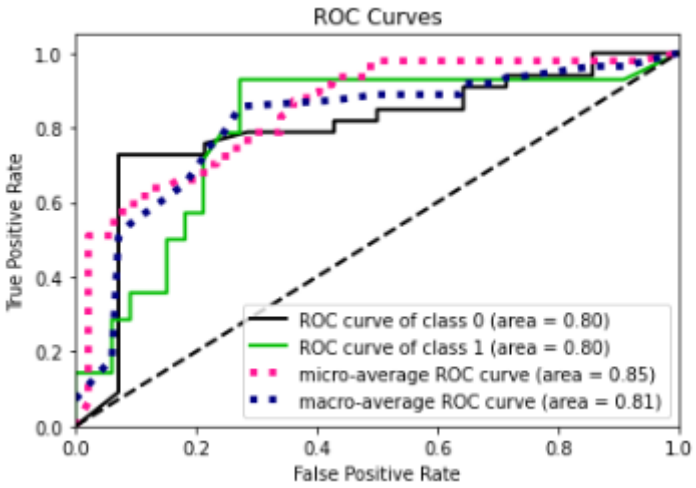
Le modèle 1, montre le modèle développé à partir de toutes les caractéristiques cliniques, chez 33 patients non mutés, on a eu 27 estimations correctes avec une précision de 82%. Cependant chez 14 patients mutés, on n'a eu que 8 estimations correctes avec une précision de 57%.
Le score de précision dans ce cas est de 0.7446.

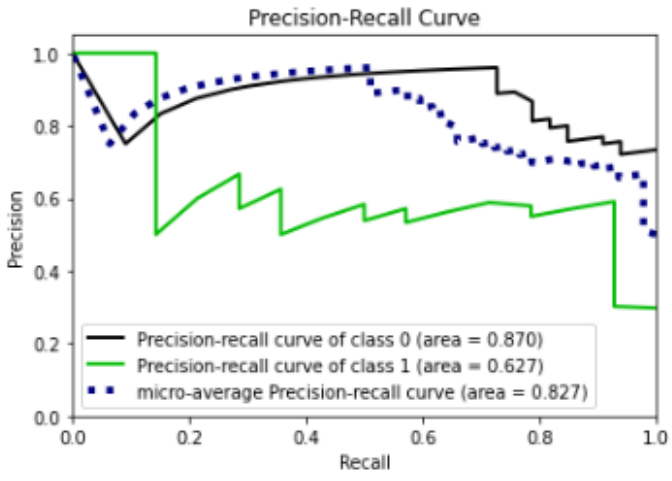
## Modèle1 : modèle développé à partir de toutes les caractéristiques cliniques :

accuracy_score : 0.7446

| Statut | precision | recall | f1-score | support |
|---|---|---|---|---|
| Non muté | 0.82 | 0.82 | 0.82 | 33 |
| Muté | 0.57 | 0.57 | 0.57 | 14 |
| accuracy | | | 0.74 | 47 |
| macro avg | 0.69 | 0.69 | 0.69 | 47 |
| weighted avg | 0.74 | 0.74 | 0.74 | 47 |



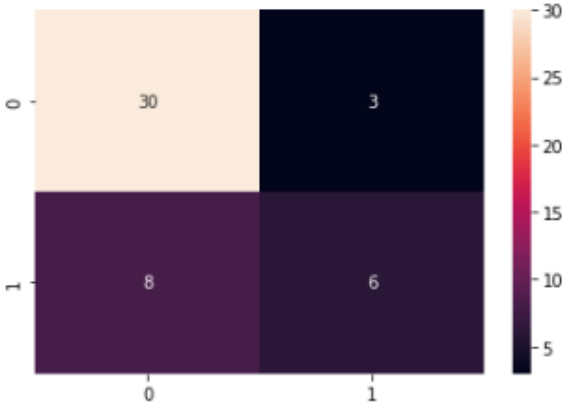



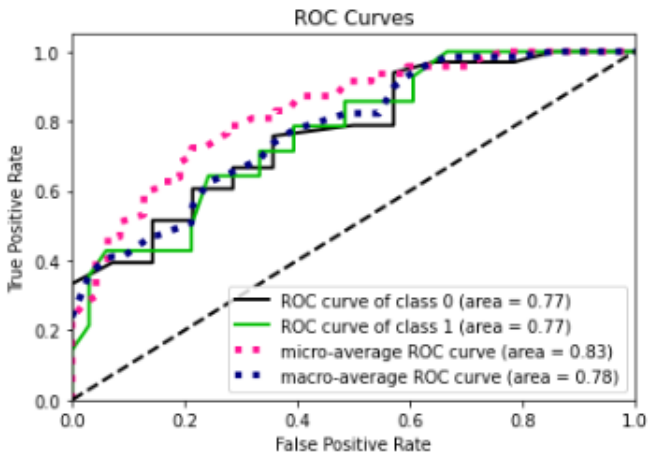
Le modèle 2, montre le modèle développé à partir des caractéristiques cliniques pertinentes en utilisant le test χ 2, chez 33 patients non mutés, on a eu 30 estimations correctes avec une précision de 79%. Cependant chez 14 patients mutés, on n'a eu que 6 estimations correctes avec une précision de 67%.
Le score de précision dans ce cas est de 0.7659.

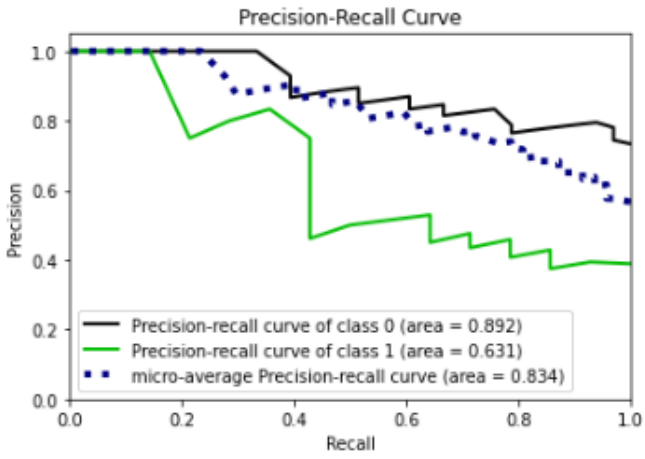
## Modèle 2 : modèle développé à partir des caractéristiques cliniques pertinentes :

accuracy_score : 0.7659

| Statut | precision | recall | f1-score | support |
|---|---|---|---|---|
| Non muté | 0.79 | 0.91 | 0.85 | 33 |
| Muté | 0.67 | 0.43 | 0.52 | 14 |
| accuracy | | | 0.77 | 47 |
| macro avg | 0.73 | 0.67 | 0.68 | 47 |
| weighted avg | 0.75 | 0.77 | 0.75 | 47 |



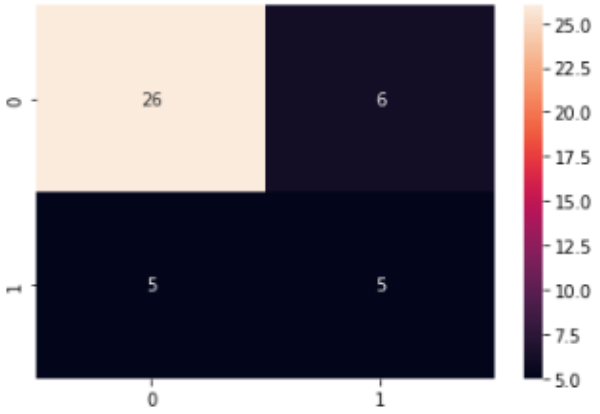



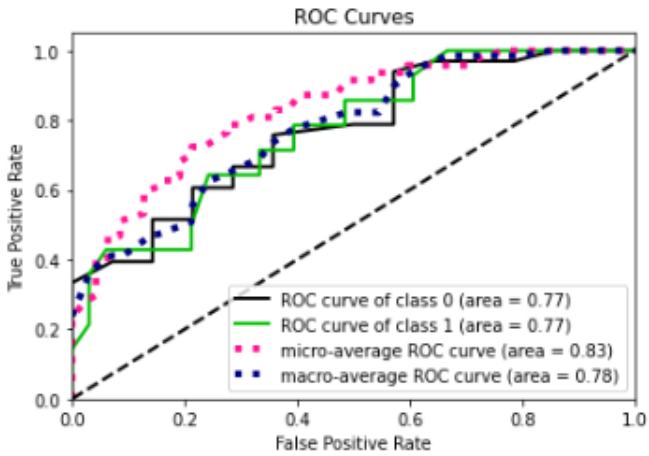
Le modèle 3, montre le modèle développé à partir de toutes caractéristiques radiomiques, chez 32 patients non mutés, on a eu 26 estimations correctes avec une précision de 84%. Cependant chez 10 patients mutés, on n'a eu que 5 estimations correctes avec une précision de 45%.
Le score de précision dans ce cas est de 0.7381.

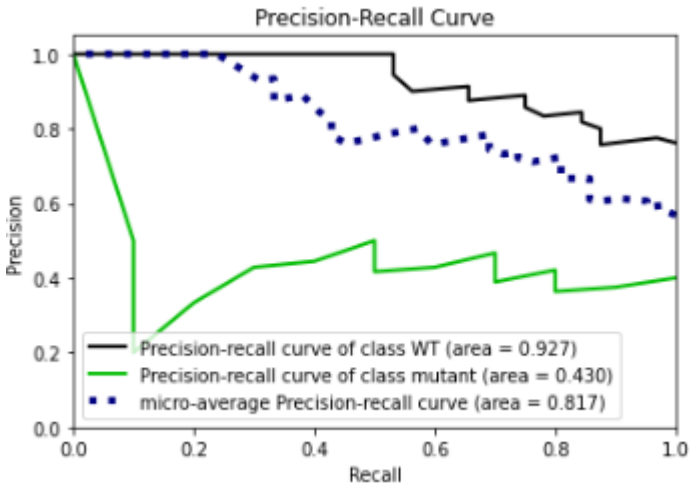
## Modèle 3 : modèle développé à partir de toutes les caractéristiques radiomiques :

accuracy_score : 0.7381

| Statut | précision | recall | f1-score | support |
|---|---|---|---|---|
| Non muté | 0.84 | 0.81 | 0.83 | 32 |
| Muté | 0.45 | 0.50 | 0.48 | 10 |
| accuracy | | | 0.74 | 42 |
| macro avg | 0.65 | 0.66 | 0.65 | 42 |
| weighted avg | 0.75 | 0.74 | 0.74 | 42 |



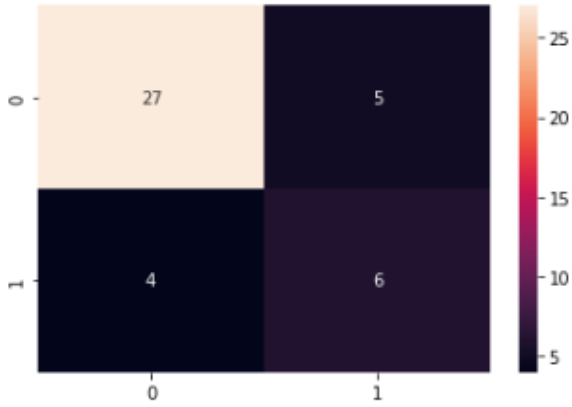



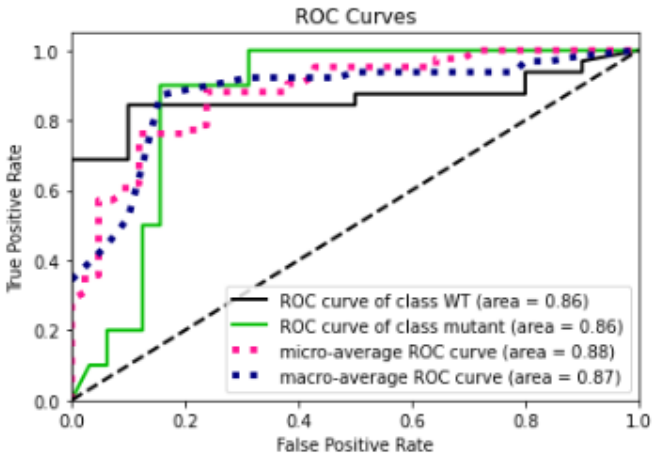
Le modèle 4, montre le modèle développé à partir des caractéristiques radiomiques pertinentes, chez 32 patients non mutés, on a eu 27 estimations correctes avec une précision de 87%. Cependant chez 10 patients mutés, on n'a eu que 6 estimations correctes avec une précision de 55%.
Le score de précision dans ce cas est de 0.7857.

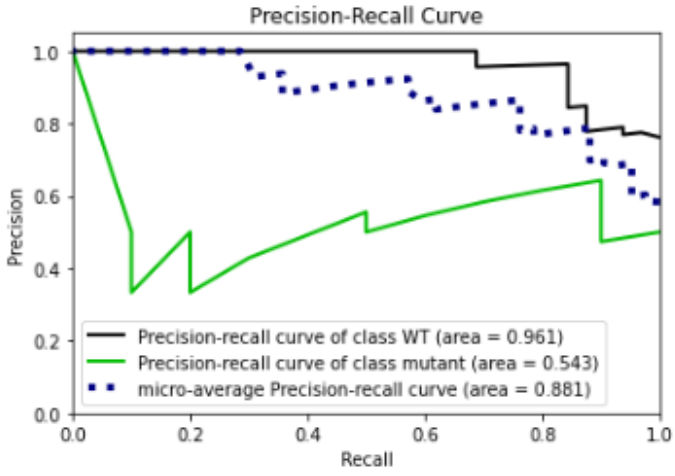
## Modèle 4 : modèle développé à partir des caractéristiques radiomiques pertinentes :

accuracy_score : 0.7857

| Statut | precision | recall | f1-score | support |
|---|---|---|---|---|
| Non muté | 0.87 | 0.84 | 0.86 | 32 |
| Muté | 0.55 | 0.60 | 0.57 | 10 |
| accuracy | | | 0.79 | 42 |
| macro avg | 0.71 | 0.72 | 0.71 | 42 |
| weighted avg | 0.79 | 0.79 | 0.79 | 42 |



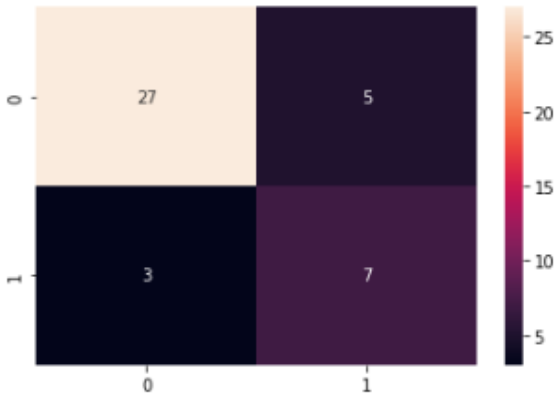



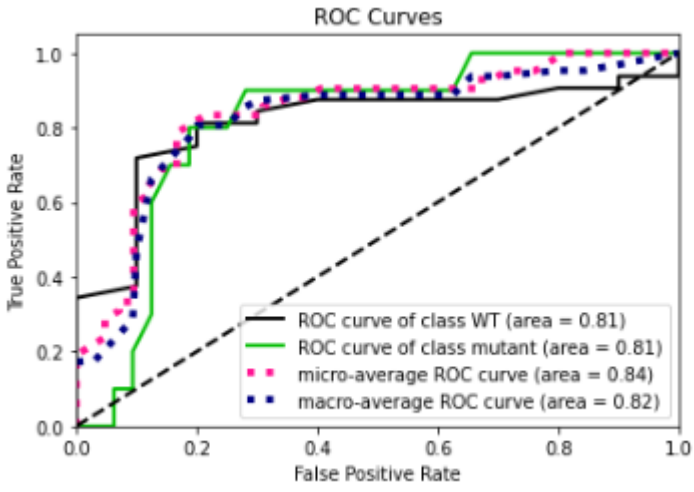
Le modèle 5, montre le modèle développé à partir des caractéristiques radiomiques pertinentes en utilisant FeatureWize, chez 32 patients non mutés, on a eu 27 estimations correctes avec une précision de 90%. Et chez 10 patients mutés, on a eu 7 estimations correctes avec une précision de 58%.
Le score de précision dans ce cas est de 0.8095.

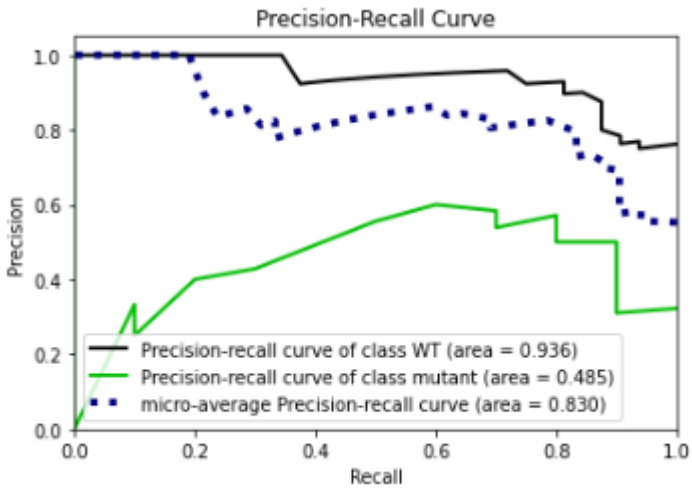
## Modèle 5 : modèle développé à partir des caractéristiques cliniques et radiomiques pertinentes :

accuracy_score : 0.8095




| Statut | precision | recall | f1-score | support |
|---|---|---|---|---|
| Non muté | 0.90 | 0.84 | 0.87 | 32 |
| Muté | 0.58 | 0.70 | 0.64 | 10 |
| accuracy | | | 0.81 | 42 |
| macro avg | 0.74 | 0.77 | 0.75 | 42 |
| weighted avg | 0.82 | 0.81 | 0.82 | 42 |



A partir de ces différents modèles, on a conclu que le modèle combiné entre les caractéristiques cliniques et radiomiques pertinentes est le plus performant pour prédire les mutations du gène EGFR. On a réalisé deux tableaux comparatifs entre les différents modèles illustrant cette performance (tableau 9, tableau 10).

**Tableau 9** : Matrice de confusion pour les 5 modèles étudiés

| | | | Predicted class | |
|---|---|---|---|---|
| | | | Muté | Non Muté |
| Les caractéristiques cliniques | | Non Muté | 27 | 6 |
| | True class | Muté | 6 | 8 |
| Les caractéristiques cliniques pertinentes en utilisant le test de $x^2$ | | Non Muté | 30 | 3 |
| | True class | Muté | 8 | 6 |
| Les caractéristiques radiomiques | | Non Muté | 26 | 6 |
| | True class | Muté | 5 | 5 |
| Les caractéristiques radiomiques pertinentes en utilisant FeatureWize | | Non Muté | 27 | 5 |
| | True class | Muté | 4 | 6 |
| Les caractéristiques cliniques pertinentes + Les caractéristiques radiomiques pertinentes | | Non Muté | 27 | 5 |
| | True class | Muté | 3 | 7 |

**Tableau 10** : comparaison de précision, recall et F1 score entre les 5 modèles étudiés

| | Précision | | Recall | | F1 score | |
|---|---|---|---|---|---|---|
| | Muté | Non Muté | Muté | Non Muté | Muté | Non Muté |
| Les caractéristiques cliniques | 0.57 | 0.82 | 0.58 | 0.81 | 0.57 | 0.82 |
| Les caractéristiques cliniques pertinentes en utilisant le test de $\chi^2$ | 0.67 | 0.79 | 0.44 | 0.91 | 0.52 | 0.85 |
| Les caractéristiques radiomiques | 0.45 | 0.84 | 0.50 | 0.81 | 0.49 | 0.83 |
| Les caractéristiques radiomiques pertinentes en utilisant FeatureWize | 0.55 | 0.87 | 0.60 | 0.84 | 0.57 | 0.86 |
| Les caractéristiques cliniques pertinentes + Les caractéristiques radiomiques pertinentes | 0.58 | 0.90 | 0.70 | 0.84 | 0.64 | 0.87 |

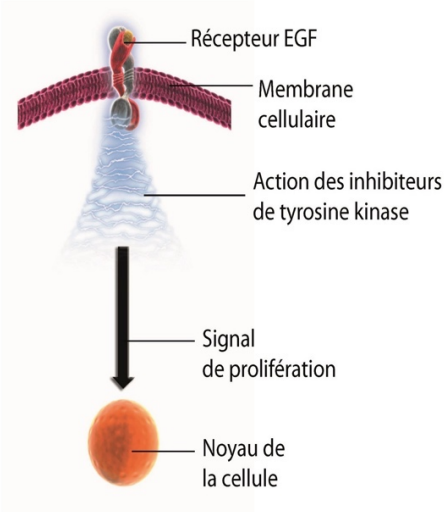
# DISCUSSION

# I. Généralité :

Le cancer du poumon est une pathologie néoplasique maligne fréquente qui représente 18 % de toutes les tumeurs, touchant principalement les hommes avec un âge moyen de 70 ans, et représentant la première cause de mortalité liée au cancer en 2020, avec 1,8 million de décès (1), (10). Les cancers pulmonaires sont divisés en carcinome pulmonaire non à petites cellules (NSCLC) et en carcinome pulmonaire à petites cellules (11). Environ 85 % des néoplasmes pulmonaires sont des NSCLC, comprenant trois sous-types principaux :

- les adénocarcinomes (40-60%)
- les carcinomes épidermoïdes (30-40%)
- les carcinomes à grandes cellules (10%).

Ces types peuvent être affectés par des mutations du récepteur du facteur de croissance épidermique (EGFR), l'homologue de l'oncogène viral du sarcome du rat Kirsten (KRAS), la kinase du lymphome anaplasiquemutation (ALK) ou autres mutations plus rare tels que ROS1 (ROS proto-oncogène 1), BRAF (v-Raf murine sarcoma viral onco-gene homolog B) V600E, NTRK (neurotrophic tyrosinereceptor kinase), MET (mesenchymal- epithelial transition), et le réarrangement RET (rearranged during transfection), (12), (13).




Au cours de la dernière décennie, le traitement du cancer du poumon non à petites cellules (NSCLC) a évolué d'une chimiothérapie cytotoxique à une thérapie ciblée basée sur des mutations moléculaires, et cela devant le développement significatif de la biologie moléculaire (14) , (15) . Les inhibiteurs de la tyrosine

kinase (ITK) à petites molécules ciblant des mutations spécifiques de l'EGFR sont les premiers médicaments ciblés pour le traitement du NSCLC.

Riley et al. (16) ont rapporté que le taux de réponse des ITK-EGFR chez les patients présentant des mutations de l'EGFR (60 à 80 %) était significativement plus élevé que celui des patients présentant un statut de mutation négatif ou inconnu de l'EGFR (10 à 20 %). Ainsi qu'un grand nombre d'essais cliniques de Schuler et al. (17) ont découvert que les traitements à base d'errotène, de gefitinib ou d'afatinibthe dans le NSCLC avec mutation de l'EGFR avait une survie (sans progression) plus longue et un taux de réponse plus élevé, par rapport à la chimiothérapie standard de première ligne. Cependant, ils ont également constaté que **si un patient atteint d'un cancer du poumon non mutant EGFR est traité avec des médicaments ciblés tels que le géfitinib, la survie était significativement plus courte que celle des patients recevant de chimiothérapie standard de première intention** (15). Ces données soulignent l'importance d'identifier avec précision l'état des mutations génétiques d'un patient pour guider le traitement.

Le diagnostic anatomo-pathologique du cancer de poumon sur les prélèvements biopsiques détecte principalement la présence et le type de la tumeur. Le diagnostic pathologique doit être classé en type et sous-types histologiques pour les patients présentant une inopérabilité avancée. Un marquage immunohistochimique est nécessaire pour les cas morphologiquement atypiques. Les prélèvements biopsiques des patients présentant un CBNPC avancé, en particulier un adénocarcinome, doivent être complété par des tests de pathologie moléculaire (18).

Pour certains patients, l'analyse des mutations de l'EGFR est très délicate voire impossible pour les anatomo-pathologistes en raison de petits fragments dans la biopsie, une erreur d'échantillonnage et une différence inter-observateur, ces paramètres peuvent entrainer l'apparition d'un résultat faux négatif (19). Deuxièmement, la biopsie est invasive, ce qui augmente le risque de complications. Enfin, le test génétique est souvent coûteux, ce qui limite l'application du test moléculaire. Malheureusement, jusqu'à présent, il n'y a pas de caractéristique clinique fiable qui puisse prédire avec précision le statut mutationnel de l'EGFR, d'où l'intérêt croissant de l'étude radiomique qui est

devenue une technique prometteuse pour identifier le phénotype d'un gène dans une tumeur. Elle peut intégrer les caractéristiques scanographiques et les données génomiques, et aider à explorer les caractéristiques de l'image qui peuvent refléter l'expression du gène.

Le radiomique commence par l'acquisition d'images de haute qualité. À partir de ces images, on peut identifier une région d'intérêt (ROI) qui contient soit la totalité de la tumeur, soit des sous-régions au sein de la tumeur. Celles-ci sont segmentées par le radiologue et sont finalement rendues en trois dimensions (3D). Les caractéristiques quantitatives sont extraites de ces volumes pour générer un rapport, qui est placé dans une base de données associé avec des données cliniques et scanographiques. Ces données sont ensuite exploitées pour développer des modèles de diagnostic, de prédiction ou de pronostic pour des résultats d'intérêt.

La figure 34 montre le processus de la radiomique et son utilisation dans la détection.

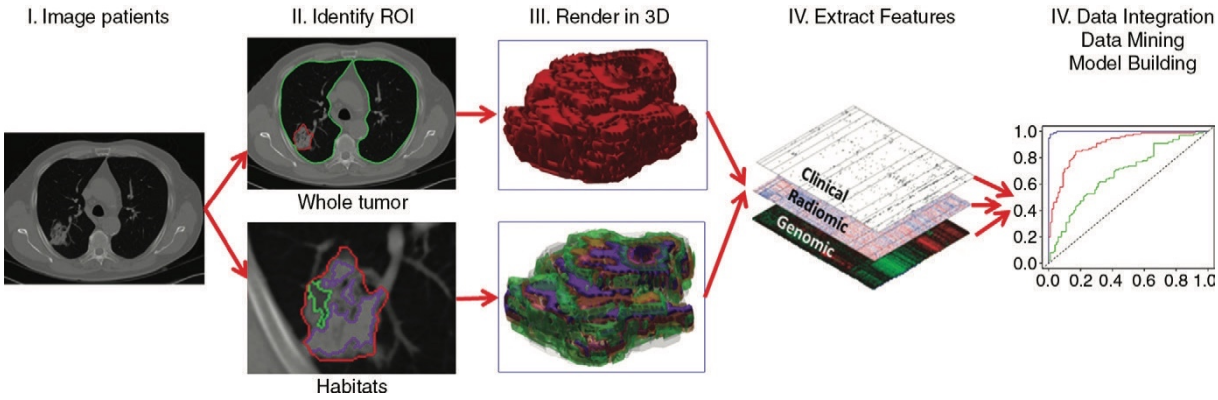



**Figure 34** : algorithme du radiomics dans la détection d'image.

Dans cette étude rétrospective, nous avons essayé d'explorer les relations entre les caractéristiques cliniques, radiomiques et la mutation de l'EGFR chez les patients atteints de NSCLC avant le traitement, et de construire un modèle prédictif combiné aux caractéristiques cliniques.

# II. Caractéristiques clinico-radiologiques :

## A. Caractérisations cliniques

Les caractéristiques cliniques sont facilement accessibles pour les patients atteints de NSCLC, et on s'attend généralement à les utiliser comme des prédicteurs de la mutation de l'EGFR. Des études cumulatives ont rapporté que plusieurs facteurs cliniques tels que le sexe féminin, les sujets non-fumeurs, l'histologie de l'adénocarcinome et l'origine asiatique étaient fréquemment associés à la mutation de l'EGFR (20). Par conséquent, dans la pratique clinique, ces facteurs sont les principaux paramètres utilisés pour la prédiction de la mutation EGFR et la confirmation est faite par la recherche de mutation par biologie moléculaire. *Dans notre étude, et conformément à la plupart des résultats précédents, nous avons constaté que les taux de mutation de l'EGFR étaient plus élevés chez les patients non-fumeurs, cependant on a trouvé que le taux de mutation est discrètement plus élevé chez le sexe masculin, et ceci pourrait être expliquer par la nette prédominance masculine dans notre échantillon étudié.*

### Mutation EGFR et sexe féminin

Une métanalyse incluant 31 études différentes portant sur 8965 patients dont l'objectif était d'évaluer l'association entre le **sexe féminin** et la **mutation EGFR** (21), a montré une hétérogénéité dans les données; cependant, une association entre le sexe féminin et la mutation EGFR a été retrouve dans la majorité des études avec un OR de 2,33 (IC à 95 % 1,97 - 2,75).

Un modèle récapitulatif des différents résultats de la métanalyse est détaillé dans la figure 35.

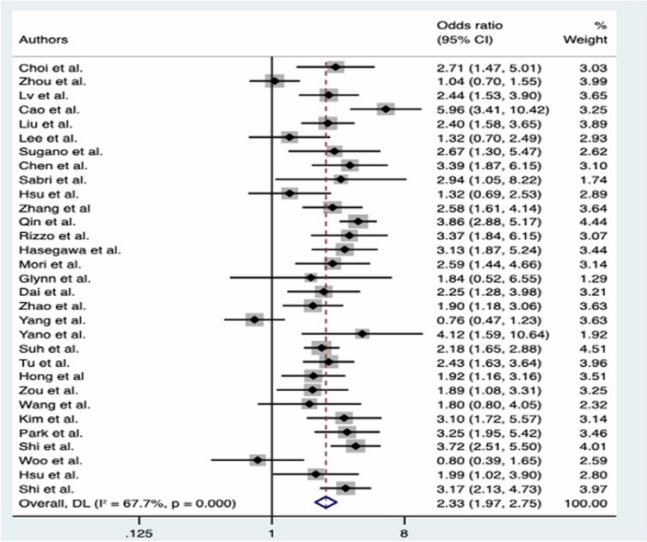



**Figure 35** : modèle récapitulatif de la métanalyse montrant la corrélation entre le sexe féminin et la mutation EGFR.

## Mutation EGFR et tabagisme

Une métanalyse incluant 27 études différentes portant sur 8214 patients dont l'objectif était d'évaluer l'association entre **le statut de non-fumeur** et la **mutation EGFR** (21), a montré une hétérogénéité dans les données ; cependant, une association entre le statut non-fumeur et la mutation EGFR a été retrouve dans la majorité des études incluses dans la métanalyse avec un OR de 2,79 (IC à 95 % 2,34 - 3,31)

Un modèle récapitulatif des différents résultats est détaillé dans la figure 36.

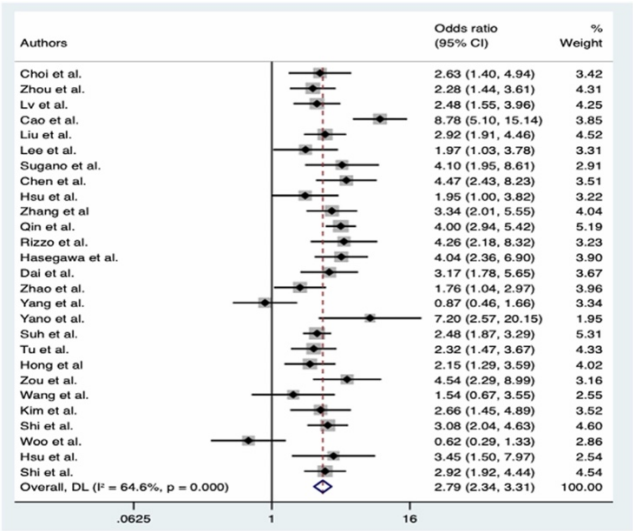



**Figure 36** : modèle récapitulatif des 27 études montrant la corrélation entre le statut non-fumeur et la mutation EGFR.

## B. Caractérisations scanographiques

Les premières études ont montré des résultats contradictoires concernant la corrélation entre les résultats scanographiques et le statut mutationnel de l'EGFR chez les patients atteints de NSCLC (19), (2). Selon une étude de Zhou et al. (20), il n'existe pas de différence dans les caractéristiques morphologiques scanographiques des tumeurs pulmonaires entre les patients mutés et non mutés. En revanche, Rizzo et al. (22) ont prouvé que les mutations de l'EGFR sont étroitement liées à quelques caractéristiques scanographiques telles que le bronchogramme aérique, la rétraction pleurale, la taille de la lésion et la présence ou non de fibrose pulmonaire.

Ces dernières années, de nombreuses études portant sur des cohortes larges avec un bon niveau de preuve ont mis en évidence une corrélation significative entre certaines caractéristiques scanographiques et le statut mutationnel de l'EGFR.

En effet, sur une étude récente, Ozkan et al. (23) ont montré que 44 % des tumeurs mutés présentaient une localisation apicale droite. Les tumeurs solides étaient prédominantes dans les adénocarcinomes mutants de l'exon 19 (88 %). Les tumeurs en verre dépoli avec une partie solide étaient plus fréquemment associées aux mutants de l'exon 21(52 %). Et 85 % des tumeurs mutés présentaient des marges irrégulières.

Mei et al. (24) ont également constaté que les caractéristiques de texture scanographique ne sont pas seulement liées au statut de mutation de l'EGFR, mais également elles peuvent aider à distinguer davantage les types de mutation des tumeurs.

Dans notre étude, nous avons utilisé 24 caractéristiques scanographiques dépassant la majorité des études ultérieures.

Les caractéristiques utilisées sont : La localisation centrale ou périphérique, le lobe pulmonaire atteint, les contours de la tumeur, la présence ou non de spiculations, la forme, les limites, la densité (solide ou mixte), la présence de bronchogramme aérique, l'attachement scissurale, l'attachement pleural, la rétraction scissurale, la rétraction pleurale, le rehaussement hétérogène ou homogène, la présence ou non de cavitation, d'emphysème, de calcifications, de fibrose, la taille, la présence de nodule pulmonaire dans le même lobe, dans un

autre lobe, les ADP médiastinales, l'épanchement pleural homolatéral, l'épanchement pleural controlatéral et enfin la présence ou non de métastase à distance.

L'analyse multivariée des données a mis en évidence une association de la mutation EGFR avec les **contours irréguliers** (ORA=6,43 ; IC95% 1,26 ;32,78), avec la **présence des spiculations** (ORA=5,81 ; IC95% 1,96 ;17,22), avec **l'attachement pleural** (ORA=3,53 ; IC95% 1,32 ;9,42), avec le **rehaussement hétérogène** (ORA=30,679 ; IC95% 5,149 ;182,778), avec l'absence de **l'emphysème** (ORA=4,815 ; IC95% 1,966 ;17,220), et avec la présence de **métastase à distance** (ORA=4,123 ; IC95% 1,373 ;12,383).

L'association avec les autres caractéristiques n'était pas significative dans notre étude.

## Mutation EGFR et contours de la tumeur

Une métanalyse incluant 21 études différentes portant sur 5871 patients (21), a montré une association significative entre la **présence de spiculation**s et la **mutation EGFR** dans la majorité des études avec un OR global de 1,42 (IC à 95 % 1,19 - 1,70)

Un modèle récapitulatif des différents résultats est détaillé dans la figure 37.

*Dans notre étude, les contours spiculés étaient significativement plus fréquents chez les patients mutés (ORA=5,81 ; IC95% 1,96 ;17,22) ce qui est concordant avec la majorité des études.*

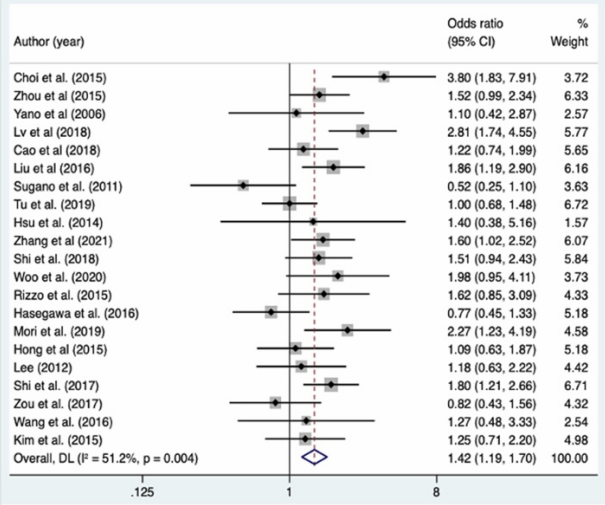



Figure 37 : modèle récapitulatif des 21 études montrant la corrélation entre la présence de contours spiculés et la mutation EGFR.

### Mutation EGFR et rehaussement hétérogène

Très peu d'études ont évalué la corrélation entre la mutation EGFR et le rehaussement tumoral.

Parmi lesquelles , une étude intéressante mené par Liu et al (25) portant sur 385 patients atteints de NSCLC a montré une association significative (P = 0.001) entre le rehaussement tumoral hétérogène et la mutation EGFR (OR, 0.23; 95% CI: 0.10, 0.53).

*Dans notre étude, nous avons trouvé que le rehaussement hétérogène était significativement plus fréquent chez les patients mutés (ORA=30,679 ; IC95% 5,149 ;182,778).*

### Mutation EGFR et absence d'emphysème

La majorité des études ont trouvé une association significative entre ces deux paramètres, ceci pourrait être expliqué par la nette prédominance de la mutation EGFR chez les patients non-fumeurs.

Rizzo S at al (22) ont trouvé une association significative entre la mutation EGFR et l'absence d'emphysème *(p=0,03)* lors de l'analyse univariée et multivariée.

Usui K et al (26) ont constaté que la fréquence de mutation EGFR était faible chez les patients atteints de maladies pulmonaires sous-jacentes tels que l'emphysème et la fibrose pulmonaire.

Une autre étude menée par Liu Y et al (25) a constaté que la mutation EGFR est élevée chez les **patients tabagiques n'ayant pas d'emphysème étendue au scanner**, Cela suggère qu'on pourrait prédire la présence de mutations EGFR chez les fumeurs qui n'ont pas ou peu de bulles d'emphysème au scanner.

*Dans notre étude, nous avons trouvé que l'absence d'emphysème était significativement plus fréquente chez les patients avec mutation EGFR (ORA=4,815 ; IC95% 1,966 ;17,220).*

### Mutation EGFR et attachement pleural :

Peu d'études dans la littérature ont évalué cette association avec des résultats discordants.

En effet, Zhang Shi et al (27) n'ont pas trouvé de corrélation significative entre l'attachement pleural et la mutation EGFR.

Zhou et al (20) ont constaté que les tumeurs pulmonaires présentant un attachement pleural étaient plus susceptibles d'être des tumeurs de type non muté.

Cependant, une étude plus récente mené par Liu Y et al (25) a constaté une association faible mais significative entre l'attachement pleural et la mutation EGFR, (P = .004), avec un OR = 0.54 (IC à 95%, 0.35 - 0.83).

*Dans notre étude, nous avons trouvé que la présence d'un attachement pleural était significativement plus fréquente chez les patients avec mutation EGFR (ORA=3,53 ; IC95% 1,32 ;9,42).*

### Mutation EGFR et métastase à distance

Très Peu d'études dans la littérature ont évalué cette association, et ceci pourrait être expliqué par le faite que la majorité des tumeurs étudiées étaient diagnostiquées à un stade précoce.

Zhang Shi et al (27) n'ont pas trouvé de corrélation significative entre les métastases à distance et la mutation EGFR.

Une autre étude mené par Rizzo S at al (22) n'a pas trouvé d'association significative entre ces deux paramètres *(P=0.64).*

*Dans notre étude, et contrairement aux études précédentes, nous avons trouvé que la présence de métastases à distance était significativement plus fréquente chez les patients mutés (ORA=4,123 ; IC95% 1,373 ;12,383).*

### Mutation EGFR et taille tumorale

Wenting Tu et al (28) ont constaté dans une étude récente que le diamètre maximal tumoral du groupe avec mutation EGFR était significativement plus petit que celui du type non mutant, par conséquent, selon ces auteurs, les tumeurs les plus petites suggèrent une mutation de l'EGFR, *ce qui est discordant avec notre étude, où on a trouvé que la majorité des patients mutés avaient une taille >3cm, ceci pourrait être due au retard diagnostique et à l'absence de dépistage précoce dans notre contexte.*

## Mutation EGFR et localisation tumorale

Wenting Tu et al (28) ont constaté également que la localisation périphérique, la densité, les ADPs, le bronchogramme aérique et la présence de cavitations étaient associés à la mutation EGFR. Le taux des mutations EGFR dans le cancer du poumon périphérique était plus élevé que dans la localisation centrale, ce qui peut être lié à la prédominance de l'adénocarcinome dans la localisation périphérique, *dans notre étude, on n'a pas trouvé cette corrélation.*

## Mutation EGFR et verre dépoli

Concernant l'association de la **mutation EGFR** et la présence de lésion en **verre dépoli (GGO)**; Une métanalyse incluant 23 études différentes portant sur 6893 patients a montré une association significative entre la présence de verre dépoli et la mutation EGFR avec un OR global de 1,86 (IC à 95 % 1,34 -2,57). Un modèle récapitulatif des différents résultats est détaillé dans la figure 38.

*Dans notre étude, la majorité des tumeurs mutées et non mutées présentaient une densité solide. Ainsi nous n'avons pas trouvé de corrélation significative entre la densité tumorale et la mutation EGFR.*

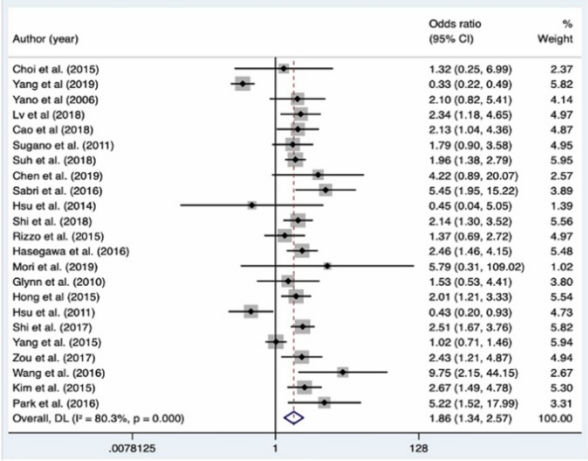



**Figure 38** : modèle récapitulatif des 23 études montrant la corrélation entre la présence de verre dépoli et la mutation EGFR.

## Mutation EGFR et bronchogramme aérique

Concernant l'association entre **le bronchogramme aérique** et la **mutation EGFR,** Une métanalyse incluant 27 études différentes portant sur 7630 patients (21), a montré une association significative entre la présence de **bronchogramme aérique** et la **mutation EGFR** dans la majorité des études avec un OR global de 1,60 (IC à 95 % 1,38 - 1,85).

Un modèle récapitulatif des différents résultats est détaillé dans la figure 39.

*Dans notre étude, le bronchogramme aérique ne présentait pas de corrélation significative avec la mutation EGFR.*

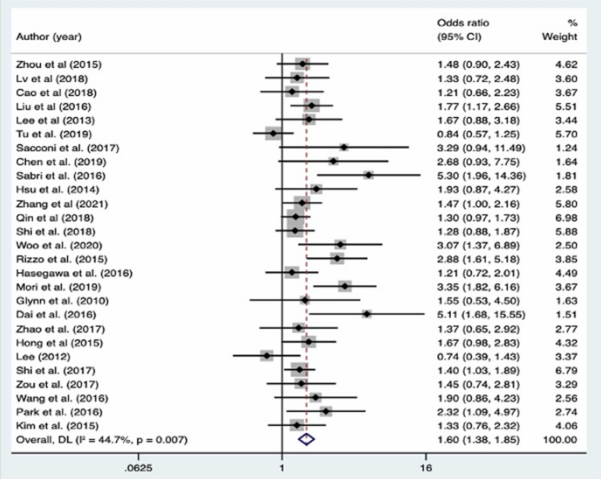



**Figure 39** : modèle récapitulatif des 27 études montrant la corrélation entre la présence de bronchogramme aérique et la mutation EGFR.

## Mutation EGFR et rétraction pleurale

Concernant l'association de la **mutation EGFR** et la présence de **rétraction pleurale**; Une métanalyse incluant 11 études différentes portant sur 3471 patients (21), a retrouvé une association entre la rétraction pleurale et la mutation EGFR dans la majorité des études avec un OR global de 1,99 (IC à 95 % 1,72 - 2,31). Un modèle récapitulatif des différents résultats est détaillé dans la figure 40.

*Dans notre étude, la majorité des tumeurs mutées EGFR ne présentait pas de rétraction pleurale.*

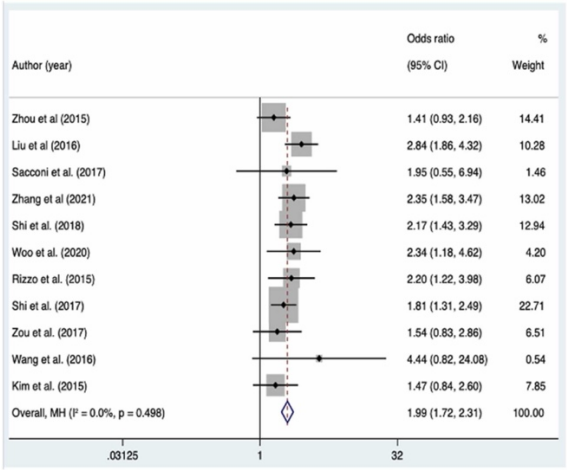



**Figure 40** : modèle récapitulatif des 11 études montrant la corrélation entre la présence de rétraction pleurale et la mutation EGFR.

## Mutation EGFR et cavitation

Et finalement, Une métanalyse incluant 15 études différentes portant sur 4891 patients dont l'objectif était d'évaluer l'association entre la **présence de cavitation tumorale** et la **mutation EGFR** (21). Sur la base des résultats de cette métanalyse, une **faible** association entre la présence **de cavitation tumorale** et la **mutation EGFR** a été retrouvée avec un OR global de 0,70 (IC à 95 % 0,57 - 0,86).

Un modèle récapitulatif des différents résultats est détaillé dans la figure 41.

*Dans notre étude, nous n'avons pas trouvé de corrélation significative entre la cavitation tumorale et la mutation EGFR.*

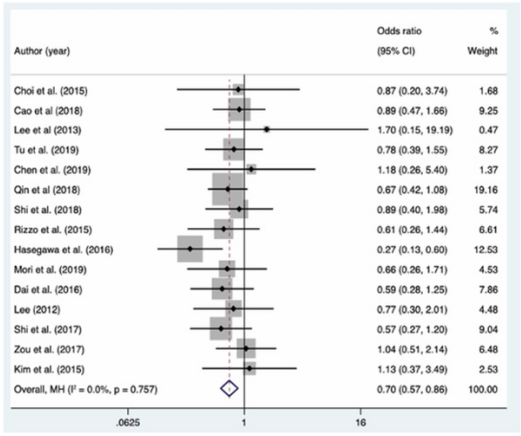



**Figure 41** : modèle récapitulatif des 15 études montrant la corrélation entre la présence de cavitation et la mutation EGFR.

## III. Segmentation d'images :

Afin de comparer nos résultats avec ceux publiés dans la littérature, nous avons détaillé dans le Tableau 11 les résultats de plusieurs publications dans le cadre de la segmentation des tumeurs pulmonaires.

Récemment, WUTIAN GAN et al (29) ont proposé une approche de réseau neuronal convolutionnel hybride 2D & 3D entraînée sur l'ensemble de données radiomiques de 260 patients atteints de cancer du poumon traités dans leur centre entre octobre 2019 et juin 2020, les résultats ont montré un score médian de dés de 0,67 pour le CNN 3D, 0,49 pour le CNN 2D et 0,73 pour l'Hybride (2D + 3D). Bing Wang et al (30) ont utilisé un réseau efficace pour la segmentation automatique des tumeurs pulmonaires complexes (CLT-Net) entraîné sur un ensemble de données de tumeurs pulmonaires difficilement segmentables DTS, l'ensemble des données se compose de 502 images scanographiques pulmonaires, chaque image contenant une tumeur DTS (tumeur difficile à segmenté), a été choisie par des radiologues spécialisés dans l'imagerie thoracique. Dans la première partie, 126 images scanographiques (fournies par l'hôpital affilié à l'université de Hebei) ont été examinées. Dans la deuxième partie, 207 images obtenues à partir de l'ensemble des données RIDER publié par TCIA en 2015 (31) ont été examinées. Dans la troisième partie, 169 images provenant de l'ensemble des données publiques du LIDCIDRI (32) ont été analysées. Cette méthode a permis d'obtenir un score médian de dés de 0,75. Les résultats ont montré que le coefficient de similarité de Dice et la métrique de distance de Hausdorff de la méthode proposée pour la segmentation DTS sont améliorés de 13,2 % et 8,5 % en moyenne, respectivement, par rapport aux modèles de segmentation classiques. Jinzhu Yang et al (33) ont utilisé un U-Net 3D équipé d'une architecture ResNet et d'un mécanisme de supervision à deux voies pour augmenter la capacité du réseau à apprendre des représentations plus riches des tumeurs pulmonaires. Des expériences approfondies ont été menées sur deux ensembles de données médicales réelles : l'ensemble de données scanographiques pulmonaires de l'hôpital de cancérologie de Liaoning en Chine avec 220 cas ; et l'ensemble de données publiques de TCIA avec 422 cas. Ces expériences ont permis d'obtenir un score moyen de 0,675 sur l'ensemble de données de l'hôpital de cancérologie de Liaoning et un score moyen de 0,691 sur

l'ensemble de données de TCIA. Jue Jiang et al (34) ont développés deux réseaux connectés à résolution multiple (MRRN) appelées incremental-MRRN et dense-MRRN. Ces réseaux combinent simultanément les caractéristiques de plusieurs résolutions d'images et les niveaux de caractéristiques par le biais de connexions résiduelles afin de détecter et de segmenter les tumeurs pulmonaires. Cette méthode a été évaluée sur un total de 1 210 tumeurs pulmonaires non à petites cellules (NSCLC) provenant de trois ensembles de données, 377 tumeurs provenant des archives d'imagerie du cancer en libre accès (TCIA), 304 de NSCLC au stade avancé traités par immunothérapie anti-PD-1 (checkpoint) provenant de l'ensemble de données de l'institution interne MSKCC, et 529 nodules pulmonaires provenant du Lung Image Database Consortium (LIDC). La méthode a produit le DSC le plus élevé de 0,74±0,13 pour TCIA, 0,75±0,12 pour MSKCC et 0,68±0,23 pour les ensembles de données LIDC.

*Le modèle que nous proposons a atteint une précision de 93,06% dans le processus de test, dépassant la plupart des résultats obtenus dans la littérature.*

**Tableau 11** : comparaison des différents résultats de segmentation des NSCLC

| | | Dataset | Dice (Mean±SD) | Jaccard Mean±SD) | HD(mm) (Mean±SD) |
|---|---|---|---|---|---|
| WUTIAN GAN et al. | Hybrid (2D + 3D) | -------- | 0.72 ± 0.10 | 0.58 ± 0.13 | 21.73 ± 13.30 |
| | 3D CNN | -------- | 0.65 ± 0.15 | 0.53 ± 0.14 | 26.73 ± 13.30 |
| | 2D CNN | -------- | 0.52 ± 0.16 | 0.40 ± 0.17 | 70.73 ± 23.30 |
| Bing Wang et al. | CLT-Net | DTS lung tumor dataset + TCIA(2015) + LIDCIDRI | 0.75 ± 0.07 | -------- | 2.49 ± 0.48 |
| Jinzhu Yang et al. | 2D U-Net | the dataset from Liaoning Cancer Hospital, | 0.578 | -------- | -------- |
| | 3D MSDS-UNet-GM | | 0.675 | | |
| | MSDS-UNet | TCIA dataset | 0.683 | -------- | -------- |
| | MSDS-UNet-GM | | 0.691 | | |
| Jue Jiang et al. | Unet | NSCLC(cross validation) | 0.68 ± 0.19 | -------- | 15.51 ± 11.74 |
| | | MSKCC(internal validation) | 0.65 ± 0.22 | | 7.87 ± 7.92 |
| | | LIDC(external validation) | 0.58 ± 0.30 | | 4.95 ±3.10 |
| | Segnet | NSCLC(cross validation) | 0.70 ± 0.20 | -------- | 15.24 ± 14.03 |
| | | MSKCC(internal validation) | 0.66 ± 0.22 | | 7.92±4.60 |
| | | LIDC(external validation) | 0.57 ± 0.29 | | 4.48±2.36 |
| | Incre | NSCLC(cross validation) | 0.74 ± 0.13 | -------- | 7.94 ± 4.96 |
| | | MSKCC(internal validation) | 0.75 ± 0.12 | | 5.85 ± 4.40 |
| | | LIDC(external validation) | 0.68 ± 0.23 | | 2.60 ± 1.71 |
| Proposed Method | stacking method | C.H.U dataset | 0.93± 0.11 | 0.83± 0.13 | 22.136± 0.34 |

Nous avons proposé dans cette étude une nouvelle architecture CNN combinant la notion de transfert de l'apprentissage profond avec un modèle d'empilement pour la segmentation volumétrique automatique du cancer du poumon à partir d'images scanographiques.

Les résultats ont montré que le CNN proposé a dépassé les autres modèles de segmentation publiés dans la littérature. Le réseau proposé a atteint une valeur de coefficient DICE supérieure à 0,93 avec une segmentation optimale et stable sur les contours de la tumeur, quelle que soit sa taille.

L'utilisation du transfert d'apprentissage et du modèle d'empilement a permis d'améliorer les deux concepts (contours flous, temps d'apprentissage) : Le modèle d'empilement a amélioré la détection tumorale en supprimant le flou qui entoure souvent le volume extrait de la tumeur. En effet, le modèle proposé a réuni les avantages de chaque sous-modèle CNN et la capacité à apprendre les informations des contours de la tumeur et à extraire localement les vraies lésions tumorales tout en éliminant, par le principe du vote, les faux positifs provoqués par les contours flous. Par conséquent, la précision de la segmentation volumétrique de la tumeur pulmonaire est nettement améliorée. En outre, l'utilisation de l'apprentissage par transfert profond élimine les paramètres complexes des couches et des longs processus de validation, réduisant le temps d'apprentissage.

Nous avons utilisé les réseaux 2D qui présentent des connexions permettant une forte mutualisation des cartes des caractéristiques de notre réseau, ce qui permet à chaque couche d'accéder aux couches précédentes et d'étendre le flux d'informations pour permettre une exploitation complète des contours tumoraux. Nous n'avons pas utilisé les réseaux de segmentation 3D qui sont très lents en temps d'apprentissage. De plus, ces réseaux 3D favorisent la détection des caractéristiques globales et conduisent à la détection de faux positifs.

# IV. Radiomics :

**Radiomics** est une méthode d'étude médicale qui vise à extraire un grand nombre de caractéristiques quantitatives d'images médicales à l'aide d'algorithmes de caractérisation des données.

**Elle** peut être appliquée à la plupart des modalités d'imagerie, y compris les radiographies standards, l'échographie, la TDM, l'IRM et le PET-scanner. Il peut être utilisé pour augmenter la précision du diagnostic, l'évaluation du pronostic et la prédiction de la réponse thérapeutique.

Dans cette étude, nous avons utilisé l'extracteur des caractéristiques radiomiques de la bibliothèque Pyradiomics de Python (35). Cette bibliothèque a été largement utilisé par des chercheurs du monde entier visant à prédire les résultats diagnostiques, génomiques, pronostiques et réponse thérapeutique pour un grand nombre de maladies.

L'objectif de notre travail était d'analyser les caractéristiques scanographiques et les caractéristiques cliniques de 149 patients, d'explorer la corrélation entre ces caractéristiques et le statut de mutation EGFR, afin de construire un modèle de prédiction combiné (associant les caractéristiques scanographiques et cliniques pertinents).

Dans notre étude. Nous avons pu extraire 1532 caractéristiques pour chaque patient, ce qui nous a fourni plus de détails et nous a aidé à compléter l'évaluation de la lésion de manière plus précise. Nous avons utilisé la bibliothèque "Featurewiz" de Python pour sélectionner les caractéristiques radiomiques pertinentes réduites à 124 caractéristiques seulement.

Nous avons pu construire cinq modèles prédictifs basés sur les données cliniques du patient, les données cliniques pertinentes, les données radiomiques, les données radiomiques pertinentes et le modèle combiné entre les données cliniques et radiomiques pertinentes. Les valeurs accuraccy = précision étaient respectivement de 0.7446, 0.7659, 0.7380, 0.7857 et 0.8095.

En intégrant les données cliniques et radiomiques pertinentes, on a pu construire un modèle combiné qui a montré une bonne performance avec un score de précision plus élevé. Ce résultat suggère que le modèle intégré basé sur la radiomique est plus efficace pour améliorer la prédiction personnalisée pré-thérapeutique de la mutation de l'EGFR (Figure 42).

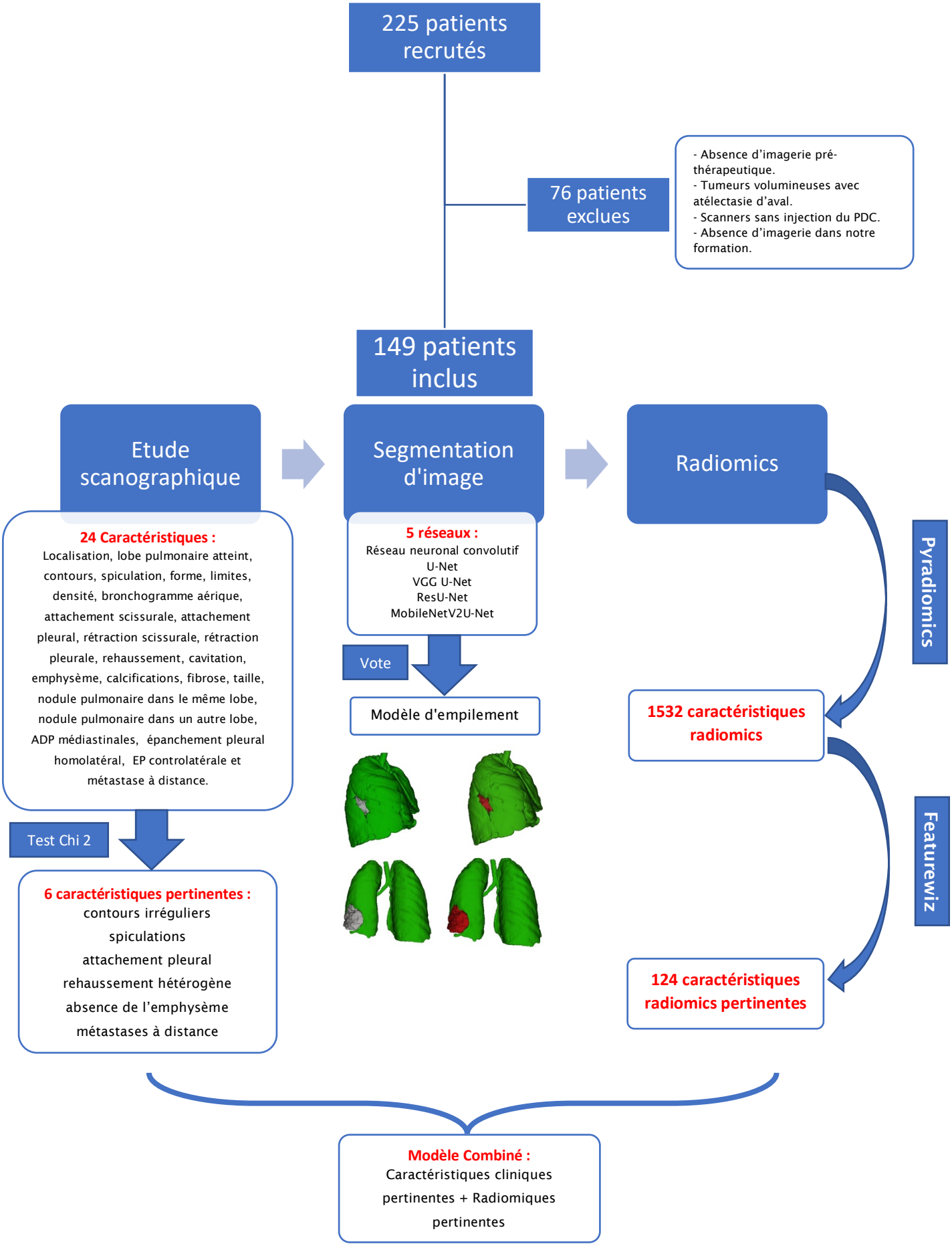



**Figure 42** Schéma de l'étude

Les résultats de notre étude sont cohérents avec l'étude de Liu et al. (36), qui ont utilisé un ensemble de 5 caractéristiques radiomiques réparties en 3 grands groupes : énergie d'atténuation CT, direction principale de la tumeur et texture définie selon les ondelettes et les lois. Cette étude a montré un score de précision AUC du ROC de 0,667 (IC à 95 %, 0,604-0,721) pour les caractéristiques cliniques seules. Un score de 0,647 (IC à 95 %, 0,576-0,701) pour caractéristiques radiomiques isolées, cependant lorsque les caractéristiques cliniques et radiomiques ont été combinées, l'ASC a été augmentée à 0,709 (IC à 95 %, 0,654-0,766).

Les résultats de notre étude sont aussi cohérents avec l'étude du Shanshan Wu et al(37) qui ont trouvé un AUC à 0,8387 pour le modèle clinique pertinent, un AUC à 0,8815 pour les caractéristiques radiomiques pertinentes, qui ont été extraites à partir de la bibliothèque **Pearson de Python**. Lorsque le modèle combine à la fois les caractéristiques radiomiques et les variables cliniques, l'ASC était de **0,9724**, ce qui était significativement plus élevé que le modèle radiomique ou clinique isolé, cela indique que le modèle combiné a montré les meilleures performances pour prédire les mutations du gène EGFR.

Zhang et al (38) ont également constaté que les résultats du modèle combiné est supérieur au modèle généré avec les caractéristiques cliniques seules et au modèle avec les caractéristiques radiomiques seules.

Bien que les critères d'inclusion, la taille de l'échantillon, l'approche de sélection des caractéristiques ou les méthodes de modélisation des études précédentes ont été différentes des nôtres, nous avons élaboré un modèle robuste de prédiction de l'EGFR avec une bonne performance.

Quelques études antérieures n'ont pas évalué les caractéristiques morphologiques de manière exhaustive. Rios et al. (39) n'ont intégré que deux caractéristiques basées sur la taille (diamètre maximal et volume tumoral) dans leur étude radiomique sur la prédiction de la mutation de l'EGFR. Le point fort de notre étude est d'inclure des caractéristiques morphologiques plus complètes que celles des études précédentes.

Cette étude comporte certaines limites : Tout d'abord, elle est transversale et ne concerne qu'une seule institution. Une étude prospective multi-institutionnelle devrait être réalisée pour permettre une généralisation du modèle et vérifier nos

résultats à grande échelle. Deuxièmement, les caractéristiques radiomiques ont été dérivées des ROIs de la segmentation manuelle et semi-automatique. Bien que la validité de la segmentation semi-automatique a été validée par nos radiologues, elle prend beaucoup de temps. Par conséquent, une méthode de segmentation automatique avec une précision élevée doit être développée. Si ces conditions sont réalisées, la probabilité de mutation de l'EGFR sera fournie automatiquement en quelques secondes. Cette méthode promet d'améliorer la gestion personnalisée de la thérapie ciblée.

# CONCLUSION

En conclusion, le cancer du poumon est la première cause de mortalité par cancer dans le monde. Les carcinomes broncho-pulmonaires non à petites cellules (NSCLC) et en particulier les adénocarcinomes représentent le type histologique le plus fréquent. Grâce aux avancées technologiques, la prise en charge des adénocarcinomes pulmonaires a profondément changé durant les 10 dernières années, passant de la chimiothérapie conventionnelle aux thérapies individuelles basées sur les caractéristiques histologiques et moléculaires de ces tumeurs. Les mutations EGFR représentent l'anomalie moléculaire la plus fréquente dans le NSCLC. La présence d'une mutation EGFR est corrélée à une réponse majeure aux inhibiteurs de tyrosine kinase (ITK) de l'EGFR.

Dans notre étude la présence d'une mutation EGFR a été détectée chez 27,50% des cas et tend à être plus fréquente chez les hommes que chez les femmes et chez les non-fumeurs que chez les fumeurs. Concernant les caractéristiques scanographiques, l'analyse multivariée par régression logistique a mis en évidence une association de la mutation EGFR avec les contours irréguliers, la présence des spiculations, l'attachement pleural, le rehaussement hétérogène, l'absence de l'emphysème et la présence de métastase à distance.

Après l'analyse des données clinico-radiologiques, nous avons proposé une nouvelle architecture CNN combinant la notion de transfert de l'apprentissage profond avec un modèle d'empilement pour une segmentation volumétrique précise du cancer du poumon à partir d'images scanographiques. Nos résultats montrent clairement l'amélioration de la précision de la segmentation sur notre ensemble de données par rapport aux autres modèles.

Ensuite, une analyse des caractéristiques radiomiques a été réalisée, et on a pu extraire 1532 caractéristiques pour chaque patient. Et enfin on a comparé les caractéristiques cliniques, cliniques pertinentes, radiomiques et radiomiques pertinentes.

En résumé, le radiomique serait un meilleur biomarqueur pour prédire le statut mutationnel de l'EGFR, en comparaison avec les caractéristiques cliniques. En outre, le modèle combiné des caractéristiques clinques et radiomiques pertinents a montré la performance prédictive la plus optimale, avec un score de précision évaluée à 0,81 surpassant les modèles cliniques et radiomiques isolés. Par

conséquent, le modèle combiné peut aider les médecins à choisir ou non un traitement ciblé par EGFR-TKI.

*Dans l'avenir, nous espérons obtenir des données plus complètes afin de prédire des sous-types de mutation de l'EGFR (exons 18, 19, 20 et 21) et des mutations types K-RAS, et ALK. Nous attendons avec impatience d'autres études plus complètes impliquant le développement de thérapies ciblées pour les différents sous-types de mutation de l'EGFR.*

# RESUMÉ

# RÉSUMÉ

***Introduction :*** Le cancer du poumon est la principale cause de décès par cancer dans le monde. Le cancer pulmonaire non à petites cellules (NSCLC) représente 85% des tumeurs pulmonaires primaires. Les récepteurs du facteur de croissance épidermique EGFR (Epidermal Growth Factor Receptor) se trouvent à la surface des cellules tumorales et leur rôle consiste à envoyer un signal de croissance au noyau de la cellule. Les cellules tumorales porteuses de mutations de l'EGFR sont très sensibles aux traitements anticancéreux appelés « Thérapies ciblées » ou « Inhibiteurs de la tyrosine kinase de l'EGFR». La recherche des mutations de l'EGFR nécessite la réalisation de biopsie pulmonaire. La lecture se fait à l'aide des techniques de la biologie moléculaire qui restent coûteuses. Radiomics est une méthode d'étude médicale qui vise à extraire un grand nombre de caractéristiques quantitatives d'images à l'aide d'algorithmes de caractérisation des données. On a utilisé dans ce travail le radiomique pour la prédiction de la mutation EGFR chez les patients atteints de cancer du poumon non à petites cellules (NSCLC).

***Objectifs*** :L'objectif principal de ce travail est d'évaluer la valeur de la CT radiomics pour prédire la mutation du récepteur du facteur de croissance épidermique (EGFR) chez les patients atteints de cancer du poumon non à petites cellules (NSCLC), et la combiner avec les caractéristiques cliniques pour construire le modèle de prédiction.

***Matériels et méthodes*** : Il s'agit d'une étude transversale, réalisée chez 149 patients suivis pour ADK pulmonaire ayant bénéficier d'une étude de mutation EGFR au service d'anatomopathologie du CHU de Fès, étalée sur une période de 4 ans (entre Janvier 2018 à Décembre 2021). Une base de données a été constituée à partir de cette exploitation incluant les données sociodémographiques, cliniques, anatomopathologiques et les résultats radiologiques. Le logiciel ITK-SNAP a été utilisé pour segmenter la tumeur pulmonaire. Les images ont été importées dans le prototype d'analyse tumorale DECT sur un poste de travail autonome. Une courbe ROC a été élaboré pour évaluer la précision diagnostique de notre modèle prédictif. En outre, une analyse statistique a été menée au sein afin d'identifier les caractéristiques radiomiques susceptibles de prédire les statuts de la mutation d'EGFR (caractéristiques pertinents).

***Résultat :*** Selon les critères d'inclusion et d'exclusion, sur les 149 patients inclus, l'âge moyen était de 61,05 +/- 11,095ans. Le sexe masculin était prédominant avec un pourcentage de 70,5% (n=105). Le sex-ratio H/F était de 2,38. 57,7% (N=86) des patients était fumeur ou ex-fumeur. 41 patients présentait une mutation EGFR soit un

pourcentage de 27,5%. L'analyse multivariée par régression logistique a mis en évidence une association de la mutation EGFR avec les contours irréguliers (ORA=6,43 ; IC95% 1,26 ;32,78), avec la présence des spiculations (ORA=5,81 ; IC95% 1,96 ;17,22), avec l'attachement pleural (ORA=3,53 ; IC95% 1,32 ;9,42), avec le rehaussement hétérogène (ORA=30,679 ; IC95% 5,149 ;182,778), avec l'absence de l'emphysème (ORA=4,815 ; IC95% 1,966 ;17,220), et avec la présence de métastase à distance (ORA=4,123 ; IC95% 1,373 ;12,383).

Nous avons pu extraire 1532 caractéristiques radiomiques pour chaque patient. Nous avons utilisé la méthode "Featurewize" pour sélectionner les caractéristiques radiomiques pertinents (N=124).
Nous avons pu construire cinq modèles prédictifs basés sur les données cliniques du patient, les données cliniques pertinents, les données radiomiques, les données radiomiques pertinents et le modèle combiné entre les données cliniques et radiomiques pertinents. Les valeurs AUC étaient respectivement de 0.7446, 0.7659, 0.7380, 0.7857 et 0.8095.

***Conclusion :*** cette étude a permis de prédire la mutation EGFR chez les patients suivis pour NSCLC, en utilisant un modèle combiné entre les caractéristiques radiomics et cliniques pertinents, ce qui pourrait éviter les biopsies invasives et la biologie moléculaire qui est couteuse.

***Mots clés :*** *cancer bronchique non à petites cellules, mutation EGFR, segmentation, radiomics, modèle de prédiction.*

# ABSTRACT

***Introduction***: Lung cancer is the leading cause of cancer death worldwide. Non-small cell lung cancer (NSCLC) accounts for 85% of primary lung tumors. Epidermal growth factor receptors (EGFR) are found on the surface of tumor cells and their role is to send a growth signal to the cell nucleus. Tumor cells with EGFR mutations are very sensitive to cancer treatments called "Targeted Therapies" or "EGFR Tyrosine Kinase Inhibitors". The search for EGFR mutations requires lung biopsies. The reading is done with the help of molecular biology techniques, which remain expensive. Radiomics is a medical study method that aims to extract a large number of quantitative features from images using data characterisation algorithms. In this work, radiomics was used for the prediction of EGFR mutation in patients with non-small cell lung cancer (NSCLC).

***Objectives:*** The main objective of this work is to evaluate the value of CT radiomics for predicting epidermal growth factor receptor (EGFR) mutation in non-small cell lung cancer (NSCLC) patients, and combine it with clinical features to build the prediction model.

***Materials and methods:*** This is a cross-sectional study, carried out in 149 patients followed up for pulmonary ADK who benefited from an EGFR mutation study at the anatomopathology department of the University Hospital of Fez, spread over a period of 4 years (between January 2018 and December 2021). A database was created from this study including socio-demographic, clinical, anatomopathological and radiological data. ITK-SNAP software was used to segment the lung tumor. The images were imported into the DECT tumor analysis prototype on a stand-alone workstation. An ROC curve was developed to assess the diagnostic accuracy of our predictive model. In addition, a statistical analysis was conducted within to identify radiomic features that could predict EGFR mutation status (relevant features).

***Result:*** According to the inclusion and exclusion criteria, of the 149 patients included, the mean age was 61.05 +/- 11.095 years. Male gender was predominant with a percentage of 70.5% (n=105). The sex ratio M/F was 2.38. 57.7% (N=86) of the patients were current or former smokers. 41 patients had an EGFR mutation, i.e. a percentage of 27.5%. Multivariate analysis by logistic regression showed an association of EGFR mutation with irregular contours (ORA=6.43; CI95% 1.26 ;32.78), with the presence of spiculations (ORA=5.81; CI95% 1.96 ;17.22), with pleural attachment (ORA=3.53 ; CI95% 1.32 ;9.42), with heterogeneous enhancement (ORA=30.679 ; CI95% 5.149 ;182.778), with the absence of emphysema (ORA=4.815 ; CI95% 1.966 ;17.220), and with the presence of distant metastasis (ORA=4.123 ; CI95% 1.373 ;12.383).

We were able to extract 1532 radiomic features for each patient. We used the "Featurewize" method to select the relevant radiomic features (N=124).
We were able to build five predictive models based on the patient's clinical data, the relevant clinical data, the radiomic data, the relevant radiomic data and the combined model between the relevant clinical and radiomic data. The AUC values were 0.7446, 0.7659, 0.7380, 0.7857 and 0.8095 respectively.
***Conclusion:*** This study was able to predict EGFR mutation in patients followed up for NSCLC, using a combined model between radiomics and clinically relevant features, which could avoid invasive biopsies and costly molecular biology.

***Keywords:*** *non-small cell lung cancer, EGFR mutation, segmentation, radiomics, prediction model.*

Annexe 1 :

| EGFR | age | sexe | tabac | localisation: central/ peripherique | lobe (LS/LM/LI) | contours: regulier, irregulier | spiculation : oui non | forme: ronde, ovalaire, indéfinit | limite: bien ou mal limité |
|---|---|---|---|---|---|---|---|---|---|
| muté | 81 | M | non | centrale | LI | irrégulière | oui | ovalaire | bien |
| muté | 71 | M | oui | périphérique | LI | régulière | non | ronde | bien |
| muté | examen avant chirurgie non touvé sur PACS | | | | | | | | |
| muté | examen avant chirurgie non touvé sur PACS | | | | | | | | |
| muté | examen non trouvé | | | | | | | | |
| muté | examen non trouvé | | | | | | | | |
| muté | 58 | M | non | centrale | LS | irrégulière | oui | ronde | bien |
| muté | 75 | F | non | centrale | LI | régulière | non | indéfinit | mal limité |
| muté | examen non trouvé | | | | | | | | |
| muté | 62 | M | oui | centrale | LS | régulière | non | indéfinit | mal limité |
| muté | examen non trouvé | | | | | | | | |
| muté | examen non trouvé | | | | | | | | |
| muté | 64 | M | oui | périphérique | LI | irrégulière | oui | ronde | bien |
| muté | examen non trouvé | | | | | | | | |
| muté | 65 | M | non | centrale | LI | irrégulière | non | ovalaire | bien |
| muté | 66 | M | oui | centrale | LS | irrégulière | oui | indéfinit | mal limité |
| muté | examen non trouvé | | | | | | | | |
| muté | examen non trouvé | | | | | | | | |
| muté | 63 | F | non | périphérique | LS | irrégulière | non | ronde | bien |
| muté | 62 | M | non | centrale | LS | irrégulière | oui | indéfinit | mal limité |
| muté | 79 | F | non | centrale | LI | irrégulière | non | indéfinit | bien |
| muté | 55 | M | oui | centrale | LI | régulière | non | indéfinit | mal limité |
| muté | 63 | F | non | centrale | LI | irrégulière | non | indéfinit | mal limité |
| muté | 77 | M | oui | périphérique | LS | irrégulière | oui | indéfinit | mal limité |
| muté | 75 | F | non | périphérique | LS | irrégulière | non | indéfinit | mal limité |
| muté | 43 | F | non | centrale | LS | irrégulière | oui | indéfinit | bien |
| muté | examen non trouvé | | | | | | | | |
| muté | 47 | F | non | périphérique | LI | irrégulière | oui | ronde | bien |
| muté | examen non trouvé | | | | | | | | |
| WT | 66 | M | non | périphérique | LI | irrégulière | oui | indéfinit | bien |
| muté | examen non trouvé | | | | | | | | |
| muté | 75 | F | non | périphérique | LS | irrégulière | oui | indéfinit | mal limité |
| muté | examen non trouvé | | | | | | | | |
| muté | examen non trouvé | | | | | | | | |
| muté | 65 | F | non | périphérique | LI | régulière | non | ronde | bien |
| muté | examen non trouvé | | | | | | | | |
| muté | 93 | M | non | centrale | LI | irrégulière | non | indéfinit | mal limité |

| densité : solide, VD, mixte | bronchogramme aerique | attachement scissural | attachement pleural | retraction scissural | retraction pleurale | rehaussement : homogene, heterogene |
|---|---|---|---|---|---|---|
| solide | oui | oui | non | oui | non | hétérogène |
| solide | oui | oui | non | non | non | homogène |
| | | | | | | |
| solide | oui | oui | non | oui | oui | hétérogène |
| solide | non | oui | oui | non | non | hétérogène |
| | | | | | | |
| solide | non | oui | oui | oui | non | hétérogène |
| | | | | | | |
| solide | oui | non | oui | oui | oui | homogène |
| | | | | | | |
| solide | non | oui | oui | non | non | hétérogène |
| solide | non | non | oui | non | non | hétérogène |
| | | | | | | |
| mixte | oui | non | non | non | oui | hétérogène |
| solide | oui | oui | non | non | non | hétérogène |
| solide | oui | non | oui | non | non | hétérogène |
| solide | oui | non | non | non | non | hétérogène |
| solide | non | oui | non | non | non | hétérogène |
| solide | non | oui | oui | non | non | hétérogène |
| solide | oui | oui | oui | non | non | hétérogène |
| solide | non | oui | non | oui | oui | hétérogène |
| | | | | | | |
| solide | non | non | oui | non | oui | hétérogène |
| | | | | | | |
| mixte | non | oui | oui | non | oui | hétérogène |
| | | | | | | |
| mixte | non | oui | oui | non | non | hétérogène |
| | | | | | | |
| solide | non | oui | non | non | non | hétérogène |
| | | | | | | |
| solide | non | non | non | non | non | hétérogène |

| cavitation | emphyseme | calcification | fibrose | taille max : <3cm, >3cm | nodule pulmonaire ds le meme lobe | nodule ds un autre lobe | ADP médiastinale | épanchement pleural homolat | EP controlateral | métastase à distance |
|---|---|---|---|---|---|---|---|---|---|---|
| non | non | non | non | 2 | non | non | non | non | non | oui |
| non | non | non | non | 2 | non | non | non | non | non | non |
| | | | | | | | | | | |
| oui | non | non | non | 2 | oui | oui | oui | non | non | oui |
| non | non | oui | non | 2 | non | oui | oui | oui | non | oui |
| | | | | | | | | | | |
| non | oui | non | non | 2 | non | non | oui | oui | non | non |
| | | | | | | | | | | |
| non | oui | non | non | 2 | non | non | non | non | non | non |
| | | | | | | | | | | |
| non | non | non | non | 2 | oui | oui | oui | non | non | oui |
| non | oui | oui | non | 2 | non | non | non | non | oui | non |
| | | | | | | | | | | |
| non | non | non | non | 1 | non | non | oui | non | non | oui |
| non | non | non | non | 2 | oui | oui | oui | oui | oui | oui |
| non | non | oui | non | 2 | oui | oui | oui | non | non | oui |
| non | oui | oui | non | 2 | non | oui | oui | oui | non | oui |
| non | non | non | non | 2 | oui | oui | oui | oui | non | non |
| non | non | oui | non | 2 | oui | oui | oui | non | non | oui |
| non | non | non | non | 2 | oui | oui | oui | non | non | oui |
| non | non | non | non | 2 | non | non | oui | non | non | oui |
| | | | | | | | | | | |
| non | non | non | non | 1 | non | non | oui | non | non | oui |
| | | | | | | | | | | |
| non | oui | non | non | 2 | oui | non | oui | non | non | oui |
| | | | | | | | | | | |
| non | non | non | oui | 2 | non | non | oui | oui | non | oui |
| | | | | | | | | | | |
| non | non | non | non | 1 | non | non | oui | oui | non | non |
| | | | | | | | | | | |
| non | non | non | non | 2 | non | non | oui | oui | oui | oui |

# BIBLIOGRAPHIE

1. Groot PM de, Wu CC, Carter BW, Munden RF. The epidemiology of lung cancer. Translational Lung Cancer Research [Internet]. juin 2018 [cité 12 mars 2023];7(3). Disponible sur: https://tlcr.amegroups.com/article/view/21996

2. Radiomics: Images Are More than Pictures, They Are Data | Radiology [Internet]. [cité 12 mars 2023]. Disponible sur: https://pubs.rsna.org/doi/full/10.1148/radiol.2015151169

3. U-Net : Réseaux convolutifs pour la segmentation d'images biomédicales | SpringerLink [Internet]. [cité 12 mars 2023]. Disponible sur: https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28

4. Wang Z, Chang Y, Peng Z, Lv Y, Shi W, Wang F, et al. Evaluation of deep learning-based auto-segmentation algorithms for delineating clinical target volume and organs at risk involving data for 125 cervical cancer patients. Journal of Applied Clinical Medical Physics. 2020;21(12):272-9.

5. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition [Internet]. arXiv; 2015 [cité 12 mars 2023]. Disponible sur: http://arxiv.org/abs/1409.1556

6. Attention-Based Personalized Encoder-Decoder Model for Local Citation Recommendation [Internet]. [cité 12 mars 2023]. Disponible sur: https://www.hindawi.com/journals/cin/2019/1232581/

7. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L. Identification of two methylated fragments of an SDC2 CpG island using a sliding window technique for early detection of colorectal cancer. In: Proc IEEE Conf Comput Vis Pattern Recognit. 2018. p. 4510-20.

8. Frontiers | Review of Deep Learning Based Automatic Segmentation for Lung Cancer Radiotherapy [Internet]. [cité 12 mars 2023]. Disponible sur: https://www.frontiersin.org/articles/10.3389/fonc.2021.717039/full

9. Wu S, Shen G, Mao J, Gao B. CT Radiomics in Predicting EGFR Mutation in Non-small Cell Lung Cancer: A Single Institutional Study. Frontiers in Oncology [Internet]. 2020 [cité 12 mars 2023];10. Disponible sur: https://www.frontiersin.org/articles/10.3389/fonc.2020.542957

10. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries - Sung - 2021 - CA: A Cancer Journal for Clinicians - Wiley Online Library [Internet]. [cité 12 mars 2023]. Disponible sur: https://acsjournals.onlinelibrary.wiley.com/doi/full/10.3322/caac.21660

11. Pape III : Cancer du poumon, mortalité cardiopulmonaire,... - Google Scholar [Internet]. [cité 12 mars 2023]. Disponible sur: https://scholar.google.com/scholar_lookup?title=Lung%20cancer&publication_year=2008&author=R.S.%20Herbst&author=J.V.%20Heymach&author=S.M.%20Lippman#d=gs_cit&t=1678626564784&u=%2Fscholar%3Fq%3Dinfo%3AHnLgAUuDfDAJ%3Ascholar.google.com%2F%26output%3Dcite%26scirp%3D0%26hl%3Dfr

12. Avrillon V, Pérol M. Alectinib for treatment of ALK-positive non-small-cell lung cancer. Future Oncology. févr 2017;13(4):321-35.

13. Bareschino MA, Schettino C, Rossi A, Maione P, Sacco PC, Zeppa R, et al. Treatment of advanced non small cell lung cancer. Journal of Thoracic Disease [Internet]. juin 2011 [cité 12 mars 2023];3(2). Disponible sur: https://jtd.amegroups.com/article/view/153

14. Jackman DM, Miller VA, Cioffredi LA, Yeap BY, Jänne PA, Riely GJ, et al. Impact of epidermal growth factor receptor and KRAS mutations on clinical outcomes in previously untreated non-small cell lung cancer patients: results of an online tumor registry of clinical trials. Clin Cancer Res. 15 août 2009;15(16):5267-73.

15. Gefitinib or Carboplatin–Paclitaxel in Pulmonary Adenocarcinoma | NEJM [Internet]. [cité 12 mars 2023]. Disponible sur: https://www.nejm.org/doi/full/10.1056/NEJMoa0810699

16. Clinical Course of Patients with Non–Small Cell Lung Cancer and Epidermal Growth Factor Receptor Exon 19 and Exon 21 Mutations Treated with Gefitinib or Erlotinib | Clinical Cancer Research | American Association for Cancer Research [Internet]. [cité 12 mars 2023]. Disponible sur: https://aacrjournals.org/clincancerres/article/12/3/839/194231/Clinical-Course-of-Patients-with-Non-Small-Cell

17. Schuler M, Wu YL, Hirsh V, O'Byrne K, Yamamoto N, Mok T, et al. First-Line Afatinib versus Chemotherapy in Patients with Non-Small Cell Lung Cancer and Common Epidermal Growth Factor Receptor Gene Mutations and Brain Metastases. J Thorac Oncol. mars 2016;11(3):380-90.

18. Health Commission of the People's Republic of China N. National guidelines for diagnosis and treatment of lung cancer 2022 in China (English version). Chin J Cancer Res. 30 juin 2022;34(3):176-206.

19. Devarakonda S, Morgensztern D, Govindan R. Genomic alterations in lung adenocarcinoma. Lancet Oncol. juill 2015;16(7):e342-351.

20. Zhou JY, Zheng J, Yu ZF, Xiao WB, Zhao J, Sun K, et al. Comparative analysis of clinicoradiologic characteristics of lung adenocarcinomas with ALK rearrangements or EGFR mutations. Eur Radiol. 1 mai 2015;25(5):1257-66.

21. Ortiz AFH, Camacho TC, Vásquez AF, del Castillo Herazo V, Neira JGA, Yepes MM, et al. Clinical and CT patterns to predict EGFR mutation in patients with non-small cell lung cancer: A systematic literature review and meta-analysis. European Journal of Radiology Open. 1 janv 2022;9:100400.

22. Rizzo S, Petrella F, Buscarino V, De Maria F, Raimondi S, Barberis M, et al. CT Radiogenomic Characterization of EGFR, K-RAS, and ALK Mutations in Non-Small Cell Lung Cancer. Eur Radiol. 1 janv 2016;26(1):32-42.

23. CT Gray-Level Texture Analysis as a Quantitative Imaging Biomarker of Epidermal Growth Factor Receptor Mutation Status in Adenocarcinoma of the Lung : American Journal of Roentgenology : Vol. 205, No. 5 (AJR) [Internet]. [cité 14 mars 2023]. Disponible sur: https://www.ajronline.org/doi/full/10.2214/AJR.14.14147

24. CT texture analysis of lung adenocarcinoma: can Radiomic features be surrogate biomarkers for EGFR mutation statuses | Cancer Imaging | Full Text [Internet]. [cité 14 mars 2023]. Disponible sur: https://cancerimagingjournal.biomedcentral.com/articles/10.1186/s40644-018-0184-2

25. Liu Y, Kim J, Qu F, Liu S, Wang H, Balagurunathan Y, et al. CT Features Associated with Epidermal Growth Factor Receptor Mutation Status in Patients with Lung Adenocarcinoma. Radiology. juill 2016;280(1):271-80.

26. Usui K, Ushijima T, Tanaka Y, Tanai C, Noda H, Abe N, et al. The Frequency of Epidermal Growth Factor Receptor Mutation of Nonsmall Cell Lung Cancer according to the Underlying Pulmonary Diseases. Pulmonary Medicine. 28 nov 2011;2011:e290132.

27. Radiological and Clinical Features associated with Epidermal Growth Factor Receptor Mutation Status of Exon 19 and 21 in Lung Adenocarcinoma | Scientific Reports [Internet]. [cité 8 avr 2023]. Disponible sur: https://www.nature.com/articles/s41598-017-00511-2

28. Tu W, Sun G, Fan L, Wang Y, Xia Y, Guan Y, et al. Radiomics signature: A potential and incremental predictor for EGFR mutation status in NSCLC patients, comparison with CT morphology.

Lung Cancer. 1 juin 2019;132:28-35.

29. Gan W, Wang H, Gu H, Duan Y, Shao Y, Chen H, et al. Automatic segmentation of lung tumors on CT images based on a 2D & 3D hybrid convolutional neural network. BJR. oct 2021;94(1126):20210038.

30. Wang B, Chen K, Tian X, Yang Y, Zhang X. An effective deep network for automatic segmentation of complex lung tumors in CT images. Medical Physics. 2021;48(9):5004-16.

31. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository | SpringerLink [Internet]. [cité 18 mars 2023]. Disponible sur: https://link.springer.com/article/10.1007/s10278-013-9622-7

32. Armato III SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans. Medical Physics. 2011;38(2):915-31.

33. Yang J, Wu B, Li L, Cao P, Zaiane O. MSDS-UNet: A multi-scale deeply supervised 3D U-Net for automatic segmentation of lung tumor in CT. Computerized Medical Imaging and Graphics. 1 sept 2021;92:101957.

34. Jiang J, Hu YC, Liu CJ, Halpenny D, Hellmann MD, Deasy JO, et al. Multiple Resolution Residually Connected Feature Streams for Automatic Lung Tumor Segmentation From CT Images. IEEE Transactions on Medical Imaging. janv 2019;38(1):134-44.

35. Computational Radiomics System to Decode the Radiographic Phenotype | Cancer Research | American Association for Cancer Research [Internet]. [cité 19 mars 2023]. Disponible sur: https://aacrjournals.org/cancerres/article/77/21/e104/662617/Computational-Radiomics-System-to-Decode-the

36. Liu Y, Kim J, Balagurunathan Y, Li Q, Garcia AL, Stringfield O, et al. Radiomic Features Are Associated With EGFR Mutation Status in Lung Adenocarcinomas. Clin Lung Cancer. sept 2016;17(5):441-448.e6.

37. Wu S, Shen G, Mao J, Gao B. CT Radiomics in Predicting EGFR Mutation in Non-small Cell Lung Cancer: A Single Institutional Study. Frontiers in Oncology [Internet]. 2020 [cité 12 mars 2023];10. Disponible sur: https://www.frontiersin.org/articles/10.3389/fonc.2020.542957

38. Zhang L, Chen B, Liu X, Song J, Fang M, Hu C, et al. Quantitative Biomarkers for Prediction of Epidermal Growth Factor Receptor Mutation in Non-Small Cell Lung Cancer. Translational Oncology. 1 févr 2018;11(1):94-101.

39. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, et al. Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer. Cancer Research. 13 juill 2017;77(14):3922-30.