MÉDECINE
Santé
Université Paris Cité

AVERTISSEMENT

Cette thèse d'exercice est le fruit d'un travail approuvé par le jury de soutenance et réalisé dans le but d'obtenir le diplôme d'Etat de docteur en médecine. Ce document est mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt toute poursuite pénale.

Code de la Propriété Intellectuelle. Articles L 122.4

Code de la Propriété Intellectuelle. Articles L 335.2 - L 335.1

MÉDECINE
Santé
Université Paris Cité

UNIVERSITÉ PARIS CITÉ

Faculté de Santé

UFR de Médecine

Année 2024/2025

THÈSE POUR LE DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE

Présentée et soutenue publiquement le : 14/10/2025

Par

MAXIME GASSIER

APPORT D'UN MODELE DE FONDATION POUR LE GRADING HISTOLOGIQUE DES ADENOCARCINOMES PULMONAIRES

Dirigée par M. le Professeur Nicolas Poté

JURY

Mme la Professeure Cécile Badoual	Présidente
M. le Professeur Nicolas Poté	Directeur
Mme la Professeure Audrey Lupo Mansuet	Membre du jury
M. le Professeur Nicolas Loménie	Membre du jury

REMERCIEMENTS

À la présidente du jury

Mme la Professeure Cécile Badoual,

Je tiens à vous exprimer toute ma gratitude pour votre transmission de savoir, vos enseignements et le soutien précieux que vous m'avez apporté dans cette belle discipline qu'est l'Anatomie-Pathologie. Vous avez su, en toutes circonstances, me guider avec bienveillance et rester à l'écoute. Votre expertise et votre professionnalisme resteront pour moi une source d'inspiration et un modèle à suivre.

Recevez ici l'expression de mon profond respect et de ma reconnaissance.

Au directeur de thèse

M. le Professeur Nicolas Poté,

Je vous exprime toute ma gratitude pour m'avoir accompagné dans la réalisation de ce travail si enrichissant, ainsi que pour le savoir-être et le savoir-faire que vous transmettez avec générosité. C'est un honneur de bénéficier de votre enseignement, de vos conseils éclairés et de votre disponibilité.

Recevez ici l'expression de mes sentiments respectueux.

Aux membres du jury

Mme la Professeure Audrey Lupo Mansuet,

Je vous suis reconnaissante de l'honneur que vous me faites en acceptant d'évaluer mon travail et d'y apporter votre précieuse expertise en Pathologie thoracique. Je vous remercie également pour votre disponibilité et votre bienveillance.

Je vous prie de recevoir, l'expression de mon profond respect et reconnaissance.

M. le Professeur Nicolas Loménie,

Je vous remercie vivement de l'honneur que vous me faites en siégeant au sein de ce jury et pour l'apport de votre expertise dans le domaine de l'intelligence artificielle, dont l'éclairage est essentiel à ce travail. Je vous suis également reconnaissante pour la confiance que vous m'accordez.

Veuillez recevoir l'assurance de ma profonde gratitude et de mon respect.

J'adresse mes remerciements les plus sincères à ma famille, en reprenant une citation de Napoléon Bonaparte : *« L'avenir d'un enfant est l'œuvre de sa mère»* Pour ma part, je dirais plutôt qu'il est l'œuvre de <u>mes parents</u>.

À toi, Papa : Merci d'être le père dont tout enfant rêverait. Tu as toujours été un exemple pour moi : un père aimant, dévoué à sa famille, empli d'amour et de bienveillance, présent aussi bien dans les rires que dans les larmes. Épanoui et passionné par ton travail, c'est toi qui m'as transmis la fibre médicale. Je me revois encore, enfant, déambuler à ta recherche dans ton laboratoire et venir m'asseoir à ton bureau. Merci d'avoir façonné l'homme que je suis aujourd'hui, grâce à ton soutien, ton équilibre et ton amour. Malgré les balades à vélo endiablées, accompagnées de la fameuse phrase « il y a du mou dans le genou » ou pendant nos folles parties de tennis à coup de « cours Forrest ! », je n'ai que ces mots à la bouche : Je t'aime Papa et merci.

A toi, Maman : Toi qui m'as donné la vie et sans qui je ne pourrais savourer tous ces instants de bonheur. Merci pour ton amour, ton soutien indéfectible, ton courage qui rassure, ta bienveillance qui réchauffe, et cette énergie qui, à elle seule, pourrait alimenter une centrale nucléaire. Tu es le pilier de notre famille, celui vers qui nous nous tournons lorsque le doute s'installe. Merci de m'avoir aidé à devenir la personne que je suis aujourd'hui. Je te garde toujours dans mon cœur : je t'aime, Maman.

A ma sœur, Émilie : Plus qu'une sœur, tu es une amie, une confidente, celle avec qui j'ai partagé le plus de moments dans cette vie. Merci pour tous ces instants de joie, pour ton soutien, pour nos soirées apéros et pour tous les merveilleux moments qui suivront, y compris ceux que nous ne raconterons pas ; car tant qu'un de nous connaît la vérité, c'est suffisant ! Je suis si fier de la femme que tu es devenue. Je t'aime, ma sœur.

A mon frère, Jules : Plus qu'un frère, tu es mon meilleur ami, celui avec qui je sais que tout devient possible. Merci d'être l'homme que tu es : honnête, bienveillant, dévoué à ta famille, et capable de me faire éclater de rire à tout moment… surtout quand on imite Papa ! Tu es l'une de mes plus grandes fiertés, et je serai toujours là pour toi. A jamais ensemble. Je t'aime, mon frère.

A toi, grand-père : Je sais que, de là où tu me regardes aujourd'hui, ton cœur est empli de bonheur. Merci pour tous ces instants précieux que nous avons partagés, pour ton soutien indéfectible, ta profonde gentillesse et l'amour que tu m'as offert. Je garde en mémoire tant de merveilleux souvenirs… Tu as bercé et illuminé mes premières années de vie, et mille mercis ne suffiraient pas à exprimer ma gratitude. Tu resteras à jamais dans mon cœur. Comme l'écrivait Lamartine : « Un seul être vous manque, et tout est dépeuplé ».

A toi, grand-mère : S'il ne fallait qu'un mot pour te définir, ce serait « gentillesse ». Merci pour tout l'amour que tu nous as donné, pour tous ces moments précieux partagés, et pour ceux encore à venir. Tu es un véritable rayon de soleil dans nos vies. Je me souviens encore des batailles avec Émilie pour avoir le privilège de nous allonger sur tes genoux. Merci pour tout. Je t'aime, grand-mère.

A ma tante, Kaky : Merci d'être tout simplement toi, maman de substitution avec ce petit grain de folie qui sait toujours me faire rire ! Merci pour tout ce que tu as fait pour moi, et pour tout ce que tu continues de faire aujourd'hui, avec générosité et tendresse.

C’est toujours un bonheur de partager des moments avec toi. Ma porte t’est et te sera toujours grande ouverte, et tu pourras, quoi qu’il arrive, compter sur moi.

À mon parrain, Jean-Noël : Tu es un véritable exemple de réussite et de gentillesse. Merci d’avoir toujours été là pour moi, aussi bien dans mes études que dans tous les autres aspects de ma vie. C’est toujours un plaisir de partager des moments de complicité avec toi. Sache que tu pourras toujours compter sur moi.

À mon tonton Alain et sa moitié Anne : Merci pour tous ces merveilleux moments passés ensemble et pour votre soutien indéfectible au fil des années. Vous n’avez été qu’amour et bienveillance. Je garde précieusement en mémoire nos vacances inoubliables à Hyères : les balades en bateau, les parties de pêche, les matchs de tennis endiablés, les sessions kitesurf… et, bien sûr, le fameux poulet au barbecue ! Sachez qu’il y aura toujours une place pour vous.

À Marie-Thérèse : Sans toi, rien de tout cela n’aurait été possible. Merci pour ton courage, ton soutien et ta générosité depuis toutes ces années. Merci aussi pour ces éclats de rire chaque matin et nos petites chamailleries, qui me permettent de passer une bonne journée. Je ne ferai que te citer : C’est pas toi, oh! Faut continuer, faut même pas t’arrêter !

À Babethouille : Merci pour tout ! Merci d’être cette personne qui, semaine après semaine, me permet de rester en forme. Merci pour tous ces fous rires partagés et pour ton soutien sans faille, présent à chaque instant. Je ne sais pas si on trouvera un jour une série qui surpasse « Yellowstone », mais ce dont je suis certain, c’est que tu pourras toujours compter sur moi.

À mes fidèles amis…

A Vincent, Lazo : Mon plus vieil ami, déjà presque 20 ans qu’on se connaît ! Merci pour ton soutien, ta « grande maturité » (qu’on a tous remarquée… ou pas), et pour tous ces fous rires et souvenirs. Franchement, on pourrait écrire un livre, voire un sketch, sur toutes les conneries qu’on a faites ado… et peut-être même en vendre assez pour financer nos prochaines bières. Je suis toujours plié quand on se remémore tous ces moments autour d’une pinte bien fraîche. Sache que tu pourras toujours compter sur moi.

A Arnaud : Mon ophtalmo préféré ! L’une de mes plus belles rencontres depuis mon arrivée à Paris. Merci pour toutes ces années d’amitié et de soutien. Je me souviens encore nos matchs de tennis… épiques pour moi, un peu moins pour toi, vu que l’issue était toujours la même : ma victoire (même quand tu portais ton slip porte-bonheur Superman). Et si jamais tu galères encore à repérer des ganglions sur tes lames histologiques, tu sais où me trouver ! On a encore un paquet de bons moments à partager. Merci d’être toujours là.

À mes deux Lillois préférés, Hugo et Julie : Déjà 10 ans qu’on se connaît… et que de super moments partagés depuis ! Au départ, je vous ai rencontrés séparément, puis ensemble et maintenant mariés, avec le petit Émile qui est venu compléter l’équipe ; la version 3.0, améliorée et certifiée « famille en or » ! C’est toujours un vrai plaisir de vous retrouver pour partager de bons moments. Sachez que ma porte vous sera toujours grande ouverte !

À mon TT : Ma plus belle rencontre de cet internat ! Un ami fidèle en toutes circonstances, avec qui un simple échange de regard se termine forcément en fou rire. Partout où on passe ensemble, il reste toujours une petite trace de nous : ce n'est pas pour rien qu'on nous appelle Tic et Tac. Je suis vraiment jalo…heureux pour tout ce qui se passe dans ta vie en ce moment, surtout depuis que tu as trouvé ta moitié et que tu t'épanouis… tel «un petit fruit à accueillir», comme dirait Camille. On a encore un paquet de bons moments à vivre, et sache que tu pourras toujours compter sur moi. Max et TT → BBF

À Nicolas : Mon Belge préféré… ou devrais-je dire *Chef Chaudard* ! Toujours un bonheur de passer du temps avec toi. Tu as le don de trouver la phrase parfaite pour me faire rire. Et plus je découvre ta famille, plus je comprends que je ne suis pas prêt d'arrêter de rigoler. Merci aussi d'être là pour ma sœur et de la rendre aussi heureuse. Sache que tu pourras toujours compter sur moi… sauf si tu m'invites à un repas 100% chicons au gratin. Là, j'invente une excuse, genre "j'ai piscine à Namur" ou "je dois nourrir mon Manneken Pis".

A M. Tihy, Matthieu : Mon ami suisse le plus barré, mais avec qui je passe toujours des moments mémorables ! Avec toi, on ne sait jamais à quoi s'attendre : ça peut commencer par une simple bière, se poursuivre à un concert symphonique, enchaîner sur un spectacle de danse et finir… en soirée techno psytrance. Crise de la quarantaine ? Mystère non résolu. Merci encore d'être venu à mon secours pour mon sujet de thèse — tu m'as sauvé la mise ! Et tu sais, dès que tu proposes quelque chose, je dis oui… alors n'hésite pas, mais préviens juste si ça finit en after à 8h du matin.

A Anaïs : Un grand merci à toi, Anaïs, pour ton aide précieuse dans la réalisation de cette thèse. Franchement, sans toi, je serais probablement encore en train de me battre avec mon premier tableau Excel et de chercher comment fonctionne QuPath… Merci de m'avoir partagé tous tes secrets (classés « top secret ») pour dompter ces logiciels, et surtout de m'avoir aidé à survivre au milieu de ces centaines de lames. C'est toujours un plaisir de te croiser à Bichat. Encore merci et courage pour finir tes travaux.

À mon groupe infernal, de l'hôpital Georges Pompidou : **Laure-Hélène**, alias LH, alias « l'indomptable ». Première complice de mes pires idées, toujours prête à la moindre virée déraisonnable, malgré ton ADN très "7e arrondissement chic". Ton rire ? Une arme de contagion massive. Par contre, s'il te plaît, comme le dit si bien Arnaud : « Rhabille-toi un peu, c'est limite ». **Arnaud**, justement… mélange improbable entre Albert Einstein et un personnage de cartoon sous caféine. Merci pour tous ces éclats de rire ; mais c'est fini ! Plus jamais tu ne t'allonges sur mon lit quand tu passes à l'appart. Et enfin, **Camille**. Merci d'être toi. Fidèle, unique, inclassable. Toujours un plaisir de passer des soirées ensemble. Mais hélas, malgré tous tes efforts, Théau restera à jamais ce petit fruit défendu que tu n'auras jamais cueilli.

A mon actuel binôme de choc, avec qui les sessions fous-rires s'enchaînent entre deux lectures de série NGS. **Yasmine**, notre déesse incontestée de la danse du ventre, capable d'enchaîner dans le même mois plus de Paris–Tunis que n'importe quel ambassadeur. Et **Nael**, la crème de la crème. La personne sur qui tu peux toujours compter, notamment sur ton pic glycémique du matin, parce qu'il a toujours la gentillesse de te ramener des petites douceurs pour accompagner ton café.

Il me semble que je n'ai oublié personne… enfin je crois… non, je rigole ! Bien sûr que je vais remercier ma grande famille, alias **« La Touche »** : Cette belle bande de loubards avec qui je traîne depuis déjà 10 ans. Dix ans de rires, de conneries, de soirées mémorables… et de lendemains beaucoup moins glorieux.

François, mon petit diable personnel, toujours prêt à m'attirer dans les vices les plus sombres… Un vrai guide spirituel… mais version *dark* ; **Guitou**, notre nounours éternel, capable de sculpter un barreau de chaise à la machette tout en sirotant son bourbon ; **Charvet**, notre surfeur/grimpeur qui vit pour l'extrême… et pour se foutre de ceux qui n'ont pas ses abdos ; **Dubourg**, notre jeune papa poule qui n'est plus le même depuis les folles soirées médecine ; **Dezouz**, notre Portugais national, aussi sucré qu'un Pastel de Nata, mais toujours prêt pour l'action ; **Jordan**, l'humoriste du groupe, dont la mission principale est de pourrir mon fil Instagram (et il réussit) ; **Gilly**, notre coureur de l'extrême et jeune marié, toujours là pour l'équipe, et qui prie pour que je sois prêt à l'heure aux prochaines vacances ; **Milan**, à la fois le plus fou et le plus sage, à potentiel hétérocurieux, je me rappelle encore cette fameuse photo qui avait été prise à Saint-Michel Escalus ; **Jacques**, notre japonais en herbe et roi du Death Valley, hâte que tu reviennes nous voir de tes contrées lointaines.

Et puis, impossible d'oublier **« la Touchette »** : Candice, Morgane, Camille, Juliette B., Juliette R., Séréna, Bertille, Aubane… chacune avec son petit truc qui rend notre groupe unique. Franchement, vous êtes incroyables… et surtout, vous rendez mes potes heureux, donc, par extension, vous me rendez heureux aussi.

En résumé : je vous aime tous !

Et je remercie toute cette grande famille de pathologistes Parisiens qui ont participé à ma formation pendant ces quatre dernières années, et avec qui j'ai partagé des moments inoubliables.

Résumé et indexation en français

Introduction : L'adénocarcinome (ADK) pulmonaire est la tumeur maligne la plus fréquente du poumon et présente un taux élevé de récidive post-opératoire. Le grade tumoral, basé sur l'évaluation par le Pathologiste de la proportion des différents patterns histologiques sur la pièce opératoire, est un critère histopronostique majeur. Cependant, le grading est une tâche complexe et chronophage, et présente une variabilité inter-observateur importante, incitant au développement d'outils d'intelligence artificielle. Récemment, les modèles de fondation ont montré de très bonnes performances dans diverses applications en Pathologie digitale. **Les objectifs principaux de ce travail étaient d'évaluer la performance d'un modèle de fondation (*H-optimus-0*) pour la classification des patterns et le grading des ADK pulmonaires.**
Matériel et Méthodes : Cette étude rétrospective monocentrique a inclus 122 ADK pulmonaires réséqués à l'hôpital Bichat, répartis en deux cohortes (entraînement n=36 ; validation n=86). Pour chaque cas, une lame HES représentative de la tumeur a été sélectionnée puis scannée. La proportion des différents patterns (lépidique, papillaire, micropapillaire, acinaire, cribriforme, solide, glandulaire complexe) et le grade (selon OMS) de chaque lame ont été évalués indépendemment par deux pathologistes (junior/senior). Les différents patterns ont été annotés sur la cohorte d'entrainement. Le modèle *H-optimus-0* a été appliqué sur les deux cohortes. Afin d'optimiser la performance du modèle, un lissage et une pondération des données ont été appliqués. Les prédictions du modèle ont été comparées à l'évaluation faite par le Pathologiste senior (PS), à l'aide de matrices de covariance et indice Kappa pondéré.
Résultats : Concernant la classification des patterns, les taux de concordance entre le modèle et le PS étaient de 0%, 8.3%, 16.7%, 28.6%, 42.9% ; 68% et 73.7% pour les patterns cribriforme, acinaire, micropapillaire, glandulaire complexe, papillaire, solide et lépidique, respectivement. Les taux de concordance entre le modèle et le PS concernant le grade étaient de 64.7% pour le grade 1, 43.8% pour le grade 2 et 98.1% pour le grade 3, avec un Kappa pondéré de 0.76. Seuls 4 cas présentaient une discordance de 2 grades (grade 1 vs grade 3) entre le modèle et le PS.
Conclusion : Ce travail montre que la performance du modèle de fondation *H-optimus-0* pour la classification des patterns est très variable selon le pattern, mais est satisfaisante pour le grading, avec une très bonne prédiction des grades 3. Ces résultats préliminaires sont à confirmer dans des cohortes indépendantes. A terme, l'intégration de cet outil à la pratique du Pathologiste pourrait améliorer la reproductibilité du grading des ADK pulmonaires, et ainsi améliorer l'évaluation du pronostic post-opératoire.

Spécialité : DES Anatomie et Cytologie Pathologiques
Mots clés français : adénocarcinome pulmonaire, classification histologique, intelligence artificielle, modèle de fondation, transformeur visuel, analyse d'images en microscopie, coupe histologique virtuelle, segmentation d'images médicales, pathologie numérique, concordance interobservateur.

Forme ou Genre :
fMeSH : Dissertation universitaire
Rameau : Thèses et écrits académiques

Résumé et indexation en anglais

Introduction : Pulmonary adenocarcinoma (ADK) is the most frequent malignant tumor of the lung and shows a high rate of post-operative recurrence. Tumor grade, based on the Pathologist's assessment of the proportion of different histological patterns on the surgical specimen, is a major histo-prognostic criterion. However, grading is a complex and time-consuming task, and presents significant inter-observer variability, prompting the development of artificial intelligence tools. Recently, foundation models have shown very good performance in various applications in Digital Pathology. **The main objectives of this work were to evaluate the performance of a foundation model (H-optimus-0) for the classification of patterns and the grading of pulmonary adenocarcinomas.**

Material and Methods : This monocentric retrospective study included 122 resected pulmonary adenocarcinomas at Bichat hospital, divided into two cohorts (training n=36; validation n=86). For each case, a representative H&E slide of the tumor was selected and scanned. The proportion of the different patterns (lepidic, papillary, micropapillary, acinar, cribriform, solid, complex glandular) and the grade (according to WHO) of each slide were independently assessed by two pathologists (junior/senior). The different patterns were annotated in the training cohort. The H-optimus-0 model was applied to both cohorts. To optimize model performance, data smoothing and weighting were applied. Model predictions were compared to the senior Pathologist's (SP) evaluation using confusion matrices and weighted Kappa index.

Results : For pattern classification, concordance rates between the model and the SP were 0%, 8.3%, 16.7%, 28.6%, 42.9%, 68% and 73.7% for the cribriform, acinar, micropapillary, complex glandular, papillary, solid and lepidic patterns, respectively. Concordance rates between the model and the SP regarding grading were 64.7% for grade 1, 43.8% for grade 2 and 98.1% for grade 3, with a weighted Kappa of 0.76. Only 4 cases showed a 2-grade discrepancy (grade 1 vs grade 3) between the model and the SP.

Conclusion : This work shows that the performance of the H-optimus-0 foundation model for pattern classification is highly variable depending on the pattern, but satisfactory for grading, with very good prediction of grade 3. These preliminary results need to be confirmed in independent cohorts. Ultimately, the integration of this tool into Pathologists' practice could improve the reproducibility of pulmonary adenocarcinoma grading, and thereby enhance post-operative prognostic evaluation.

Specialty : MD Pathology, Anatomical and Cytological Pathology
Keywords in English: pulmonary adenocarcinoma, histological classification, artificial intelligence, foundation model, visual transformer, microscopy image analysis, virtual histological slide, medical image segmentation, digital pathology, inter-observer agreement.

Publication type:
MeSH : Academic Dissertation

Liste des abréviations

- **ADK** : adénocarcinome
- **AI** : artificial intelligence (intelligence artificielle)
- **AUC** : area under the curve
- **CNN** : convolutional neural network
- **EGFR** : epidermal growth factor receptor
- **HES** : hématéine-éosine-safran
- **IASLC** : International Association for the Study of Lung Cancer
- **KRAS** : Kirsten rat sarcoma virus oncogene
- **LCC** : large cell carcinoma
- **MIL** : multiple instance learning
- **MSI** : microsatellite instability
- **NDCG** : normalized discounted cumulative gain
- **PA** : paquet-années
- **PD-L1** : programmed death-ligand 1
- **pT / N** : classification pTNM (taille tumorale et extension ganglionnaire)
- **ResNet** : residual network
- **SCC** : squamous cell carcinoma (carcinome épidermoïde)
- **SCLC** : small cell lung carcinoma
- **TCGA** : The Cancer Genome Atlas
- **TNM** : tumor, node, metastasis
- **UNI** : universal computational pathology foundation model
- **VADS** : voies aéro-digestives supérieures
- **WSI** : whole slide image

Table des matières

REMERCIEMENTS **2**

RESUME ET INDEXATION EN FRANÇAIS **7**

RESUME ET INDEXATION EN ANGLAIS **8**

LISTE DES ABREVIATIONS **9**

INTRODUCTION **12**

1. ADENOCARCINOMES PULMONAIRES : GENERALITES 12
   - *1.1. Épidémiologie, classification histologique et pronostic 12*
   - *1.2. Challenges dans l'évaluation des patterns et du grading des adénocarcinomes pulmonaires 19*
2. L'INTELLIGENCE ARTICIELLE EN MEDECINE 25
   - *2.1. Définition et historique 25*
   - *2.2. Principes généraux 26*
   - *2.3. Applications actuelles de l'IA en médecine et en cancérologie 29*
3. L'INTELLIGENCE ARTIFICIELLE EN PATHOLOGIE MEDICALE 29
   - *3.1. Différents modèles d'IA 29*
   - *3.2. IA et grading des adénocarcinomes pulmonaires 33*

PROBLEMATIQUE **35**

OBJECTIFS DE THÈSE **38**

MÉTHODOLOGIE **39**

1. PATIENTS ET ECHANTILLONS 39
2. ANNOTATION DES LAMES (COHORTE D'ENTRAINEMENT) 40
3. GRADING DES TUMEURS 42
5. PROTOCOLE DU MODELE DE CLASSIFICATION 43
6. ANALYSES STATISTIQUES 45

RÉSULTATS **47**

1. CARACTERISTIQUES ANATOMO-CLINIQUES DES COHORTES 47
2. CONCORDANCE INTER-OBSERVATEURS 48
3. PERFORMANCE DU MODELE POUR LA CLASSIFICATION DES PATTERNS ET LE GRADING 49
   - *3.1 Performances du modèle « brut » 49*
   - *3.2 Performances du modèle « lisse » 50*
   - *3.3 Performances du modèle après pondération 51*

3.4 Performances du modèle après lissage et pondération 52
3.5 Comparaison appariée des proportions par pattern sur la cohorte test (validation) 53
3.6 Comparaison appariée des proportions par pattern sur la cohorte d'entrainement 54
4. EXEMPLES DE CLASSIFICATION (HEATMAP) DES PATTERNS AVEC LE MODELE « H-OPTIMUS-0 » 55
5. EXEMPLES DE RESULTATS DISCORDANTS ENTRE LE MODELE ET LE PATHOLOGISTE SENIOR 60
6. EXEMPLES DE CLASSIFICATION (SEGMENTATION) DES PATTERNS AVEC « H-OPTIMUS-0 » 62

DISCUSSION 64

CONCLUSION ET PERSPECTIVES 68

BIBLIOGRAPHIE 69

TABLE DES FIGURES 73

TABLE DES TABLEAUX 75

INTRODUCTION

1. Adénocarcinomes pulmonaires : Généralités

1.1. Épidémiologie, classification histologique et pronostic

- **EPIDEMIOLOGIE**

Le cancer du poumon représente la première cause d'incidence et de mortalité par cancer, avec plus de 2,2 millions de nouveaux cas et 1,7 million de décès estimés dans le monde en 2018 (1). Ce taux de mortalité élevé s'explique en grande partie par un diagnostic souvent tardif et par la forte prévalence du tabagisme (2). Les taux de mortalité et d'incidence ont historiquement été les plus élevés dans les pays développés, notamment aux États-Unis et en Europe, mais ils diminuent désormais progressivement, en particulier chez les jeunes hommes et femmes. Le cancer du poumon a longtemps été plus fréquent chez les hommes, mais dans de nombreux pays (comme les États-Unis, le Danemark et la Suède), les taux d'incidence chez les hommes et les femmes tendent à converger.

Le risque de développer un cancer pulmonaire croît proportionnellement à la durée et à l'intensité du tabagisme, et diminue progressivement après l'arrêt, sans toutefois revenir au niveau observé chez les non-fumeurs. D'autres expositions environnementales et professionnelles sont également impliquées, telles que le radon (dans les habitations ou les mines), l'amiante, l'utilisation domestique du charbon, et la pollution atmosphérique extérieure (1). Par ailleurs, le cancer du poumon peut survenir chez des patients non-fumeurs, avec une incidence annuelle estimée entre 5 et 10 cas pour 100 000 personnes. -
La classification histologique des cancers pulmonaires est la pierre angulaire du diagnostic, de la prise en charge et du pronostic. La majeure partie des cancers du poumon sont des carcinomes, dont les principaux types histologiques sont :

- les carcinomes non à petites cellules (70%), regroupant principalement l'adénocarcinome (45%) et le carcinome épidermoïde (25%)
- le carcinome à petites cellules (15%)

Ainsi, l''adénocarcinome (ADK) pulmonaire est le type histologique le plus fréquent de carcinome pulmonaire, représentant environ 40 à 50 % des cas. Différents sous-types histologiques d'ADK pulmonaires sont identifiés, dont l'ADK non mucineux qui est de loin le plus fréquent. Son incidence est en augmentation, notamment chez les non-fumeurs et les femmes, ce qui en fait un enjeu majeur de santé publique (3–5). Cette prévalence croissante est en partie attribuée à une meilleure détection grâce aux avancées en imagerie thoracique, mais également à des facteurs environnementaux et génétiques encore mal compris.

Au cours des 20 dernières années, la prise en charge des cancers du poumon a considérablement évolué, se traduisant par une amélioration significative de la survie des patients. Ces améliorations sont en grande partie dues aux avancées rapides dans

la compréhension des altérations moléculaires associées à cette maladie, aux progrès dans leur détection, ainsi qu'au développement de nouveaux traitements, en particuliers les immunothérapies (1,2).

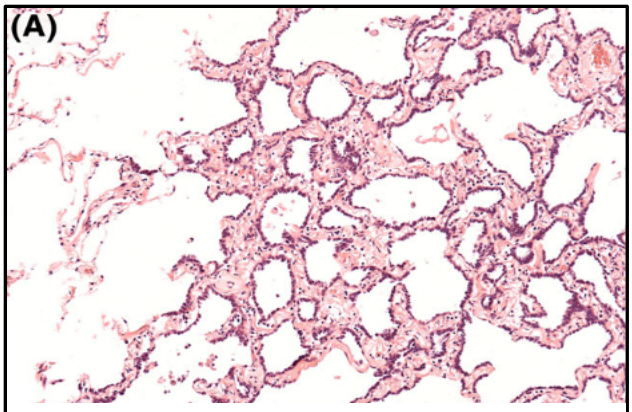
Malgré ces progrès, le pronostic du cancer pulmonaire reste sombre avec, en France, une survie globale à 5 ans tous stades confondus estimée à 17%. Seul 20 à 30% des patients présentent au diagnostic un cancer de stade localisé, pouvant bénéficier d'une résection chirurgicale qui reste le traitement le plus efficace. Malgré cela, le taux de récidive post-opératoire reste élevé et la survie globale 5 ans est faible, évaluée à 85%, 60%, 41% et 24% pour les stades I, II, IIIA et IIIB, respectivement (6).

Un des critères histopronotiques majeur post-opératoire des ADK pulmonaires non mucineux est le grade histologique, basé sur l'évaluation par le pathologiste de la proportion des différents patterns architecturaux présents sur l'ensemble de la tumeur. En effet, l'ADK pulmonaire présente une grande hétérogénéité histopathologique, avec la coexistence de plusieurs motifs ou « patterns » architecturaux au sein d'une même tumeur.

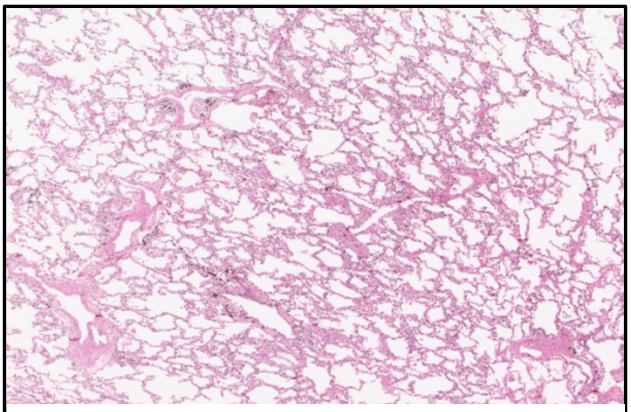
- **CLASSIFICATION HISTOLOGIQUE**

Selon la classification de l'Organisation Mondiale de la Santé (OMS) et celle de l'IASLC (2020), les principaux patterns histologiques sont :

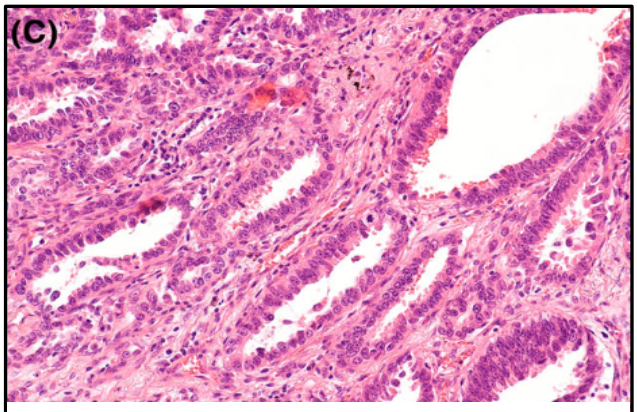
- Le pattern **LEPIDIQUE** : caractérisé par une croissance tumorale le long des structures alvéolaires, sans invasion stromale (ADK in situ) ni destruction de l'architecture pulmonaire.



Pattern Lépidique (2)


Image issue de la cohorte Hôpital Bichat

- Le pattern **ACINAIRE** : caractérisé par la présence de structures glandulaires bien définies.



Pattern Acinaire (2)

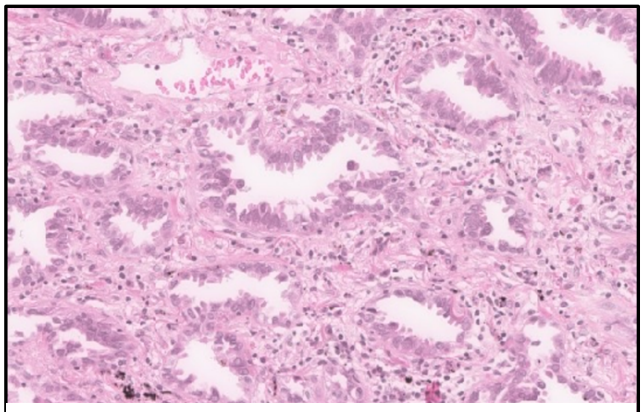
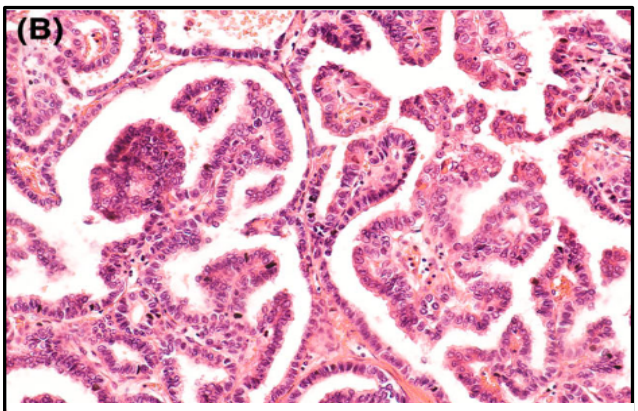


Image issue de la cohorte Hôpital Bichat

- Le pattern **PAPILLAIRE** : caractérisé par la formation de structures ayant une forme de papille et comprenant un axe fibrovasculaire.



Pattern Papillaire (2)

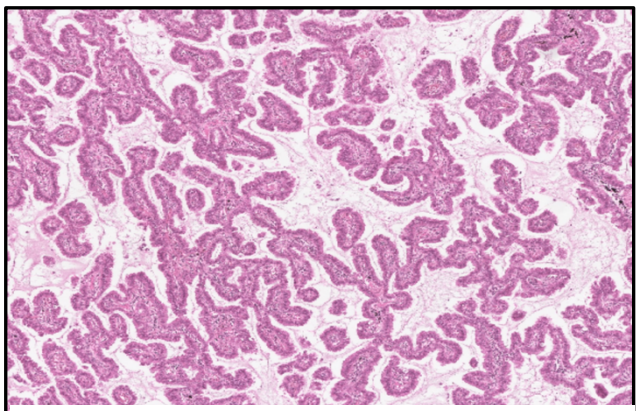
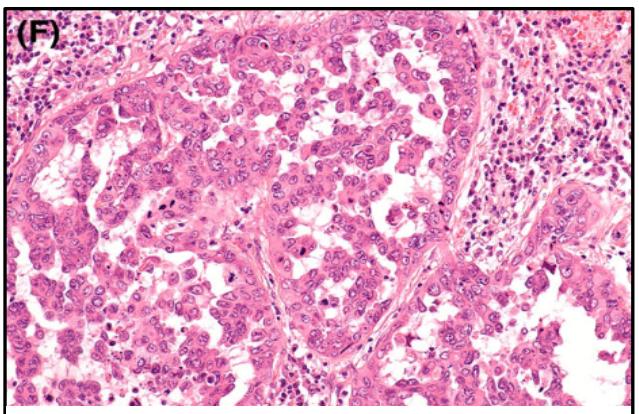


Image issue de la cohorte Hôpital Bichat

- Le pattern **MICROPAPILLAIRE** : caractérisé par des structures papillaires sans axes fibrovasculaires.



Pattern Micropapillaire (2)

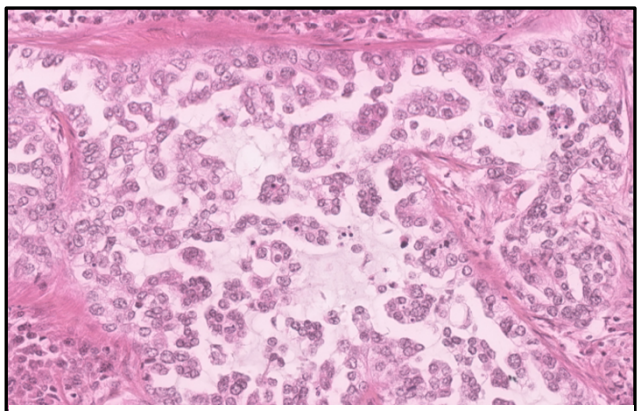
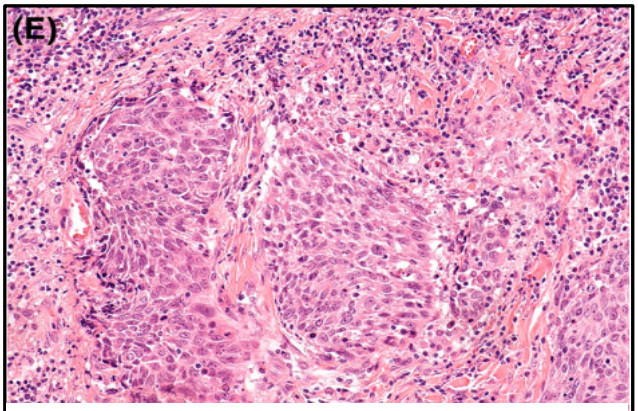


Image issue de la cohorte Hôpital Bichat

- Le pattern **SOLIDE** : caractérisé par des massifs tumoraux sans formation glandulaire évidente.



Pattern Solide (2)

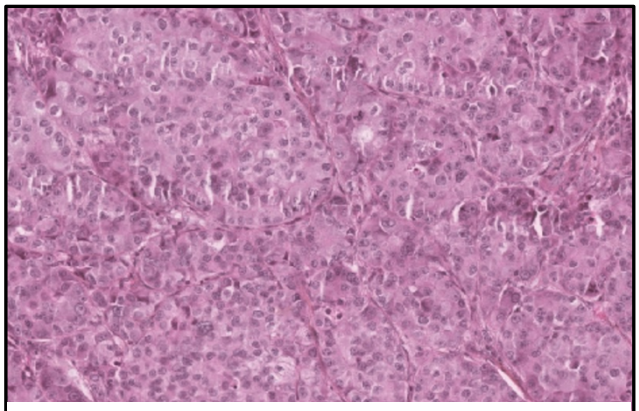
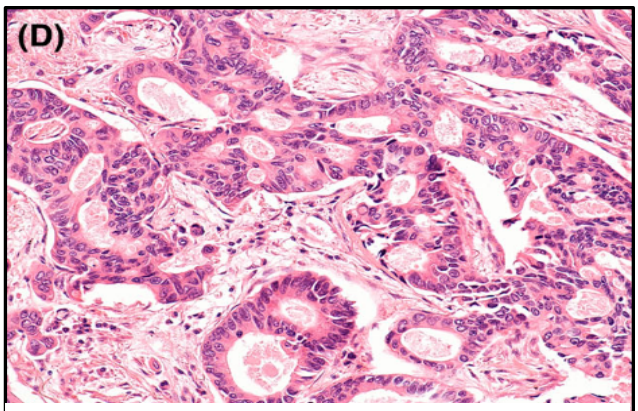


Image issue de la cohorte Hôpital Bichat

- Le pattern **CRIBRIFORME** : caractérisé par des massifs creusés de cavités glandulaires, sans stroma intermédiaire.



Pattern Cribriforme (2)

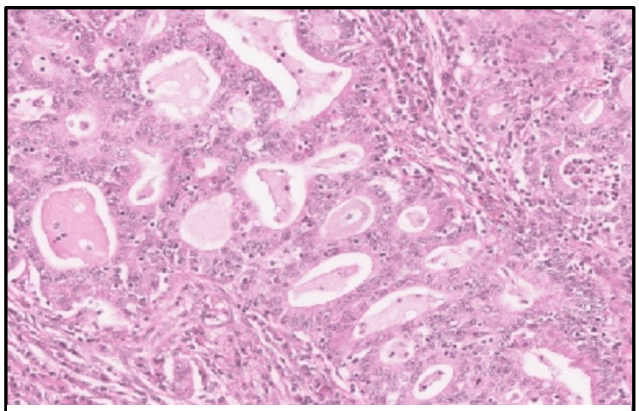
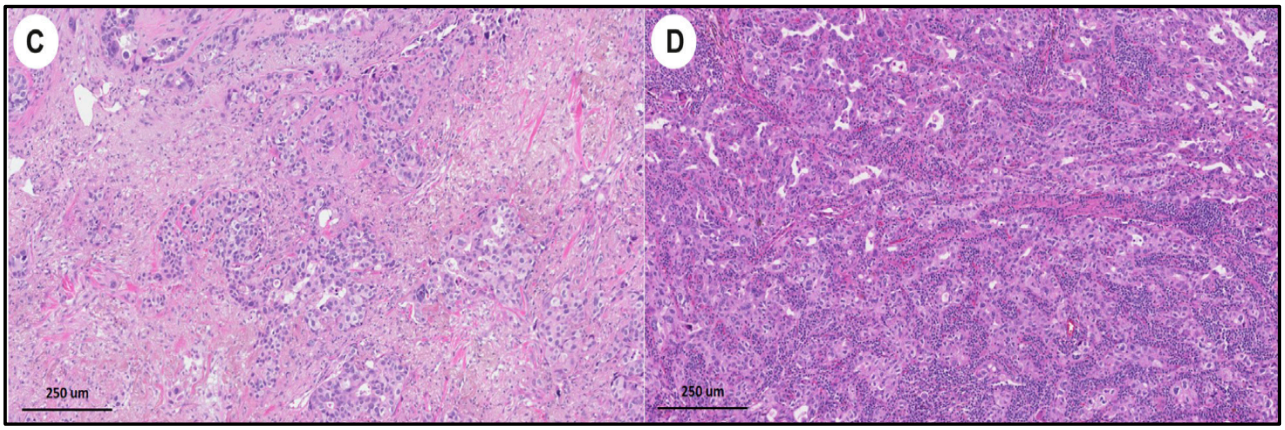
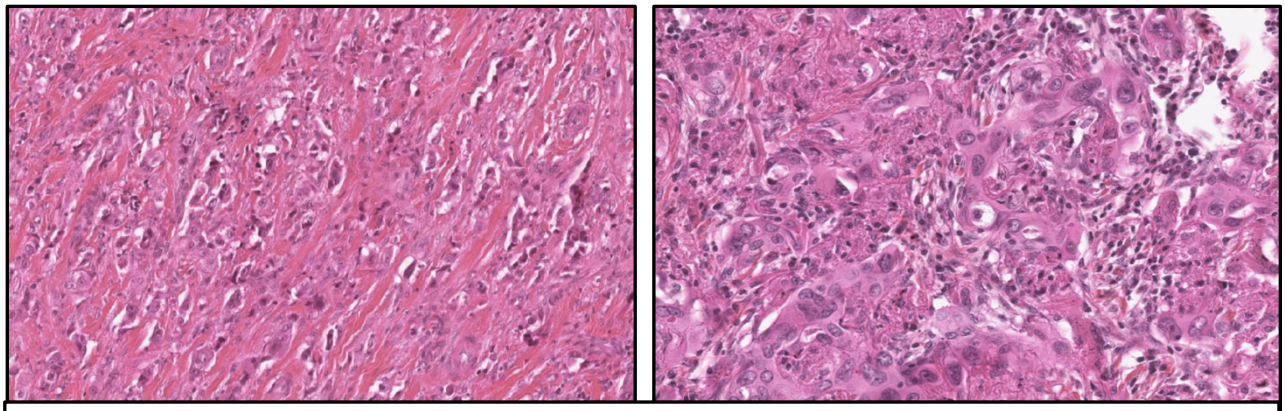


Image issue de la cohorte Hôpital Bichat

- Le pattern **GLANDULAIRE COMPLEXE** : caractérisé par des glandes fusionnées ou cellules uniques infiltrant un stroma desmoplastique (4)(7).



Exemples représentatifs de pattern glandulaires complexes. (C) glandes fusionnées dans un stroma desmoplastique avec des cellules isolées dispersées. (D) amas de cellules tumorales anastomosées en forme de ruban et de travées irrégulières. (7)



Pattern glandulaire complexe : Images issues de la cohorte Hôpital Bichat

- **PRONOSTIC**

De façon intéressante, chacun de ces patterns possède une valeur pronostique distincte chez les patients opérés (4). Par conséquent, l'identification précise de ces patterns est essentielle afin d'établir un grading tumoral fiable et d'évaluer au mieux le pronostic post-opératoire.

Le système de grading des ADK pulmonaires a initialement reposé sur le pattern histologique prédominant, reparti en trois grades (7–10):

- **Grade 1** : tumeurs à prédominance lépidique
- **Grade 2** : tumeurs à prédominance acinaire ou papillaire
- **Grade 3** : tumeurs à prédominance solide ou micropapillaire

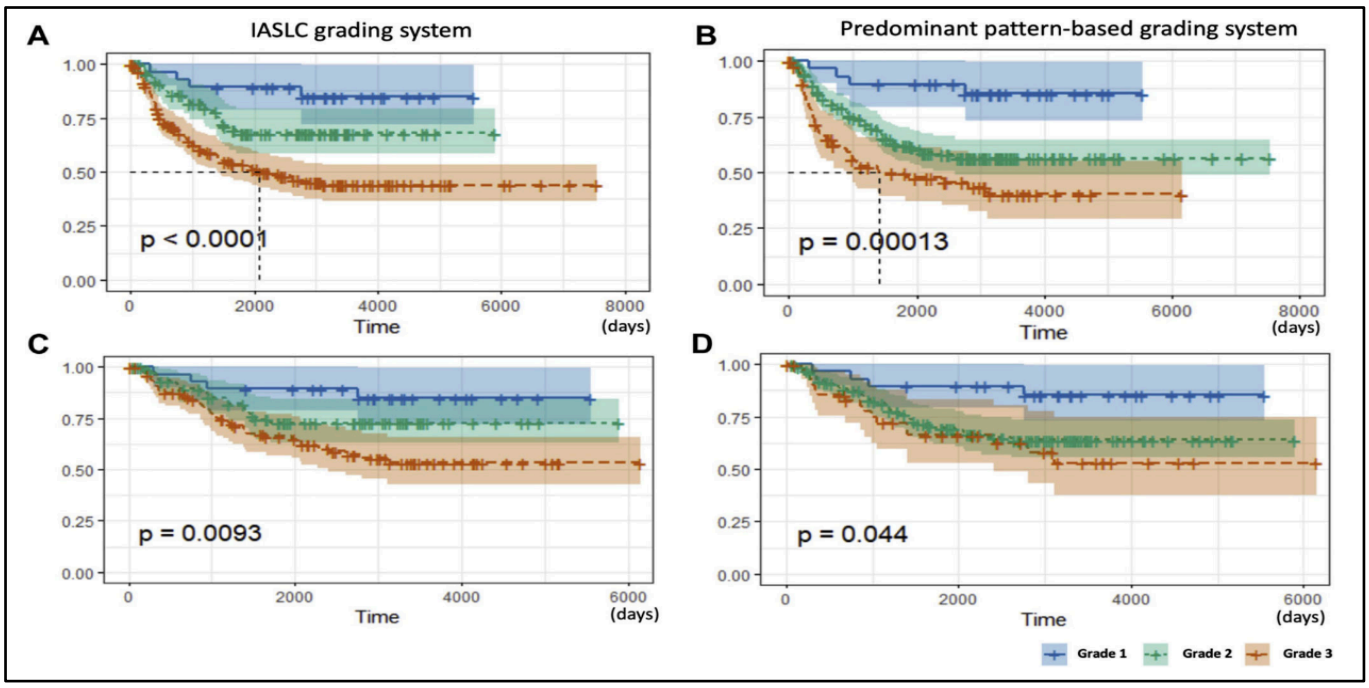


Figure 1 : **Courbes de survie sans récidive selon la classification internationale IASLC et le système de classification basé sur le pattern prédominant. Colonnes A et C :** IASLC grading system (proposé par le comité IASLC, basé sur le pattern prédominant + la proportion de patterns de haut grade ≥ 20%). **Colonnes B et D :**

predominant pattern-based grading system (système plus ancien, basé uniquement sur le pattern architectural prédominant). **Grades : Grade 1 (bleu) :** survie nettement meilleure, peu de récidives **; Grade 2 (vert) :** survie intermédiaire **; Grade 3 (orange) :** survie significativement moins bonne, avec plus de récidives précoces(7).

La **figure 1** compare la survie sans récidive selon le système de grading IASLC et celui basé uniquement sur le pattern prédominant. Dans les deux cas, les grades 1, 2 et 3 montrent des pronostics distincts, mais la discrimination est nettement meilleure avec le système IASLC, qui intègre la proportion de patterns de haut grade ≥20 %. Cette approche permet une stratification pronostique plus robuste, y compris chez les patients de stade I, et confirme la supériorité du système IASLC pour prédire le risque de récidive.

Même si ce système présente une certaine validité, il a été reconnu que le pattern prédominant peut être très hétérogène, et que la variabilité intratumorale est fréquente dans les adénocarcinomes. De plus, certains patterns comme le pattern micropapillaire, bien que non prédominants, peuvent avoir une valeur pronostique importante même lorsqu'ils ne constituent qu'un composant minoritaire.

En tenant compte de ces observations, un nouveau système de grading des ADK pulmonaires a récemment été proposé, combinant deux critères (1) :

1. Le pattern prédominant (en %).
2. Toute tumeur présentant plus de 20 % de pattern de haut grade (micropapillaire, solide, cribriforme ou glandulaire complexe) est classée en Grade 3 (peu différencié).

Avec ce nouveau système de grading, il a été démontré que la stratification pronostique, notamment la survie sans récidive et la survie globale, était améliorée par rapport à l'usage du seul pattern prédominant (7,8).

D'autre part, il est désormais bien établi que la proportion de patterns agressifs est corrélée à une augmentation du risque de récidive et à une diminution de la survie globale, même lorsqu'ils sont présents en faible quantité (≥5 %) (2,4,5).

Dans une vaste étude récente, entreprise par le comité de pathologie de l'Association internationale pour l'étude du cancer du poumon (IASLC), la combinaison du pattern histologique prédominant et du pattern de haut grade (>20%) était associée de façon significative au pronostic (survie sans récidive, survie globale) dans une large cohorte d'ADK pulmonaires réséqués à un stade précoce, et s'est avérée supérieure à d'autres modèles incorporant le nombre de mitoses, le grade nucléaire, le grade cytologique, le STAS ou la nécrose (4).

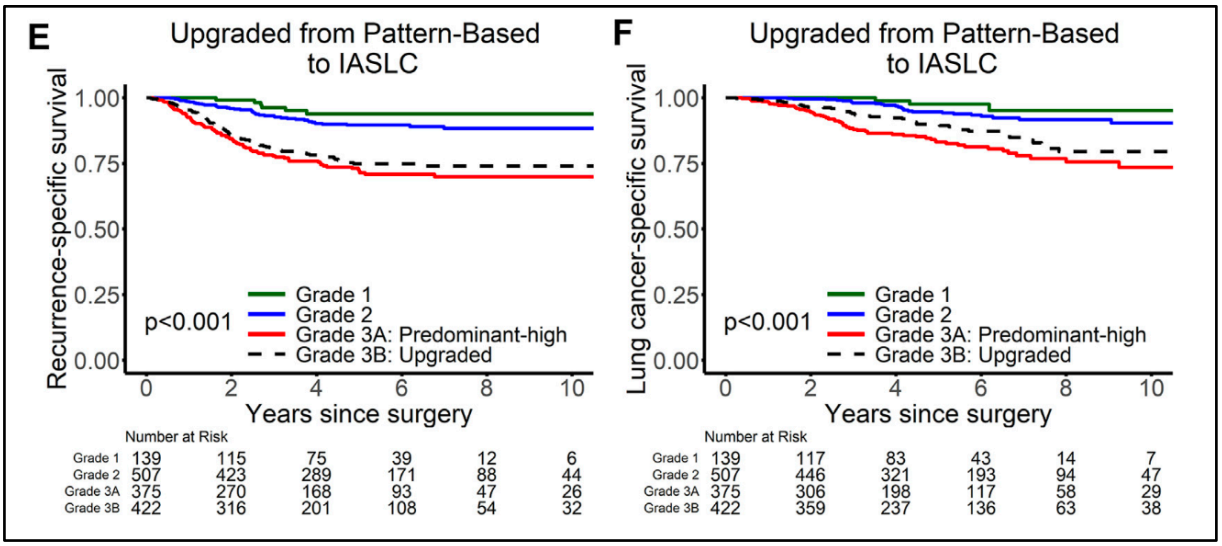


Figure 2 : **Survie sans récidive (E) et survie globale (F) des patients après chirurgie selon le grade histologique de la tumeur.** Courbes spécifiques des tumeurs « upgradées » en grade 3 par présence ≥20% de patterns micropapillaire, solide, cribriforme (4).

La **figure 2** montre que les patterns non prédominants de haut grade modifient significativement le pronostic, même sans être dominants.

Cela a conduit à la recommandation OMS d'un système de classification à trois grades pour les ADK pulmonaires réséqués à un stade précoce **(Table 1.08)(1))** :

Grade	Différenciation	Patterns
1	Bien différencié	Prédominance lépidique, sans pattern de haut grade ou avec < 20 % de pattern de haut grade
2	Modérément différencié	Prédominance acinaire ou papillaire, sans pattern de haut grade ou avec < 20 % de pattern de haut grade
3	Peu différencié	Toute tumeur avec ≥ 20 % de pattern de haut grade (solide, micropapillaire, cribriforme ou glandes complexes[1])

[1] *Glandes fusionnées ou cellules isolées infiltrant dans un stroma desmoplastique.*

Table 1.08 (1) : Système de grading OMS des adénocarcinomes pulmonaires (selon la prédominance architecturale et la proportion de patterns de haut grade)

Les patterns histologiques observés dans l'adénocarcinome pulmonaire possèdent une valeur pronostique significative, car ils reflètent des mécanismes biologiques complexes liés au comportement tumoral. Ces motifs traduisent notamment des différences dans la prolifération cellulaire, la survie, ainsi que dans les propriétés d'adhésion des cellules tumorales entre elles et avec leur environnement stromal.

Par exemple, les tumeurs présentant un pattern solide se caractérisent par une cellularité élevée, pouvant résulter soit d'une activité proliférative accrue, soit d'une une résistance accrue à l'apoptose. De plus, ces tumeurs doivent adapter leurs mécanismes d'adhésion : les cellules situées au centre interagissent entre elles, tandis que celles à la périphérie établissent des contacts avec la matrice extracellulaire, favorisant potentiellement l'invasion locale.

Le pattern micropapillaire, quant à lui, illustre des formes de croissance plus désorganisées, avec des projections cellulaires en doigt de gant, témoignant d'une altération de la cohésion cellulaire, d'une motilité accrue, et probablement d'une capacité d'invasion renforcée.

En revanche, le pattern lépidique apparaît comme un motif de croissance beaucoup plus organisé, avec des cellules formant une couche unique, bien adhérentes aux parois alvéolaires, et sans signe de remaniement stromal : cette architecture suggère un comportement tumoral plus indolent.

Bien que ces observations permettent d'établir des corrélations cliniques robustes entre pattern et pronostic, elles reposent encore largement sur des données descriptives et observationnelles. Les fondements biologiques précis de ces relations, bien qu'intrigants, demeurent insuffisamment élucidés à ce jour (8,10,11).

1.2. Challenges dans l'évaluation des patterns et du grading des adénocarcinomes pulmonaires

Le grading des adénocarcinomes pulmonaires, basé sur la reconnaissance des patterns histologiques et l'évaluation de leurs proportions respectives, est une tâche complexe et ardue pour le pathologiste. L'évaluation des patterns repose sur des critères purement morphologiques évalués au microscope optique, mais ceci soulève un certain nombre de difficultés et limites présentées ci-dessous.

1. Reproductibilité inter-observateur dans l'évaluation des patterns.

La reconnaissance des différents patterns tumoraux repose sur l'expertise du pathologiste, mais cette tâche est complexe et sujette à une variabilité inter-observateur. Des études ont montré que la concordance inter-observateur est particulièrement faible pour certains patterns comme le pattern micropapillaire, avec des taux de concordance allant de 62 % à 85 % (7,8,12).

Plusieurs études ont évalué la reproductibilité de l'approche basée sur l'identification des patterns histologiques. Il en ressort que :

- Le pattern prédominant est généralement associé à une bonne reproductibilité inter observateur.
- Les patterns solide et lépidique obtiennent une meilleure concordance inter-observateurs que les patterns papillaire ou micropapillaire (8,13).

Cependant, la concordance diminue significativement lorsqu'il s'agit de patterns difficiles à interpréter, souvent intriqués. L'introduction de la notion de pattern secondaire, en plus du pattern prédominant, tend également à réduire la reproductibilité.

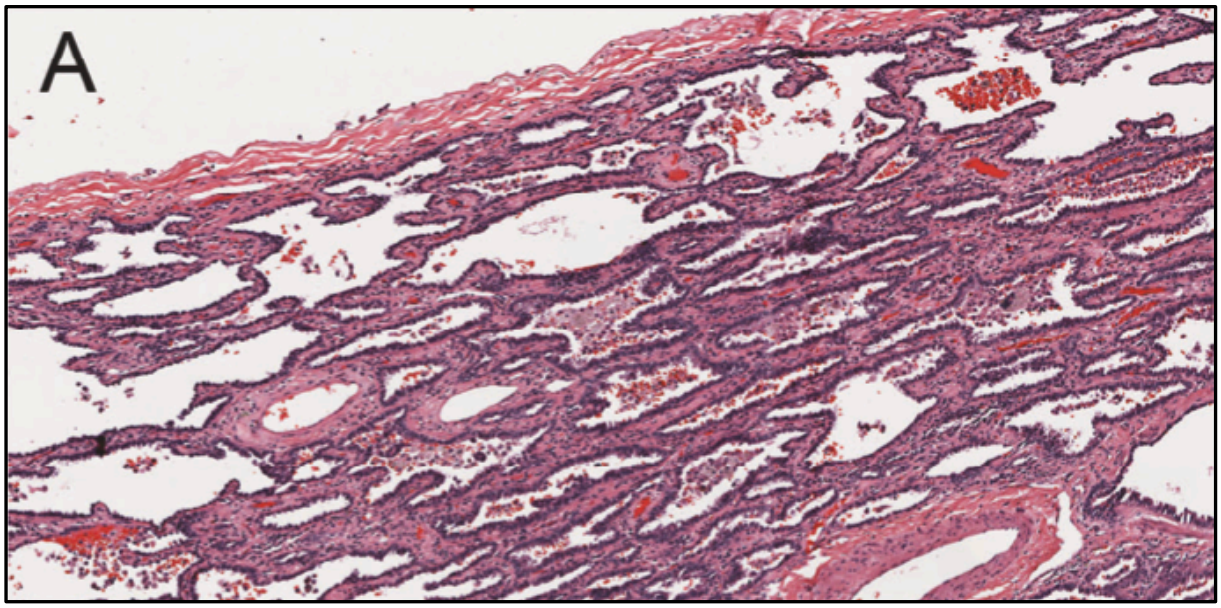
En définitif, la reproductibilité du système de grading proposé dépend de deux choses : la capacité à identifier correctement le pattern prédominant et la reconnaissance fiable des patterns de haut grade, notamment si la proportion de ces derniers est supérieure à 20 %.

L'évolution de l'adoption du système de grading OMS dépendra donc de l'amélioration de ces deux points critiques (8,14).

D'autre part, l'évaluation de l'invasion tumorale dans l'ADK pulmonaire se heurte à plusieurs difficultés liées à la complexité des patterns architecturaux et aux nombreux artefacts pouvant perturber l'interprétation.

- Pattern lépidique versus pattern acinaire

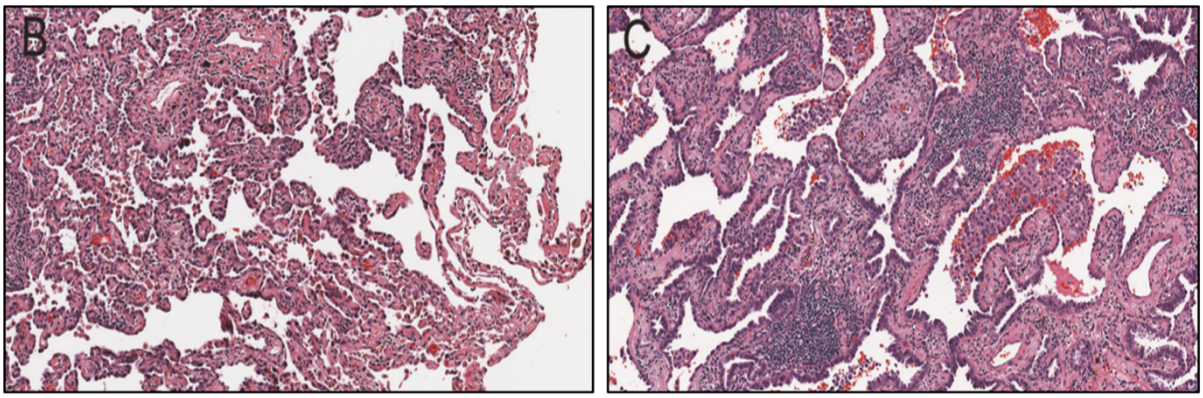
L'un des premiers écueils concerne la distinction entre le pattern lépidique et le pattern acinaire. Lorsque les cellules tumorales bordent des espaces alvéolaires dont l'architecture est globalement préservée, le diagnostic de pattern lépidique est relativement simple et reproductible. Même en cas de collapsus alvéolaire, certains indices morphologiques comme l'alignement parallèle des structures, la présence de macrophages intra-alvéolaires, ou encore la conservation d'un revêtement épithélial monocouche, permettent de conforter cette interprétation.



Exemple de pattern lépidique avec collapsus alvéolaire. Organisation parallèle des espaces alvéolaires, orientés selon un grand axe. Les changements de type collapsus peuvent déformer le schéma lépidique. (8)

Cependant, dans certaines situations, les parois alvéolaires peuvent être épaissies du fait d'une fibrose ou d'une inflammation. Ces altérations rendent plus difficile l'appréciation de l'architecture originelle et peuvent faire apparaître des structures faussement complexes. La distinction devient encore plus difficile lorsque le collapsus alvéolaire produit des profils pseudo-acinaires, avec des contours angulés ou convexes.

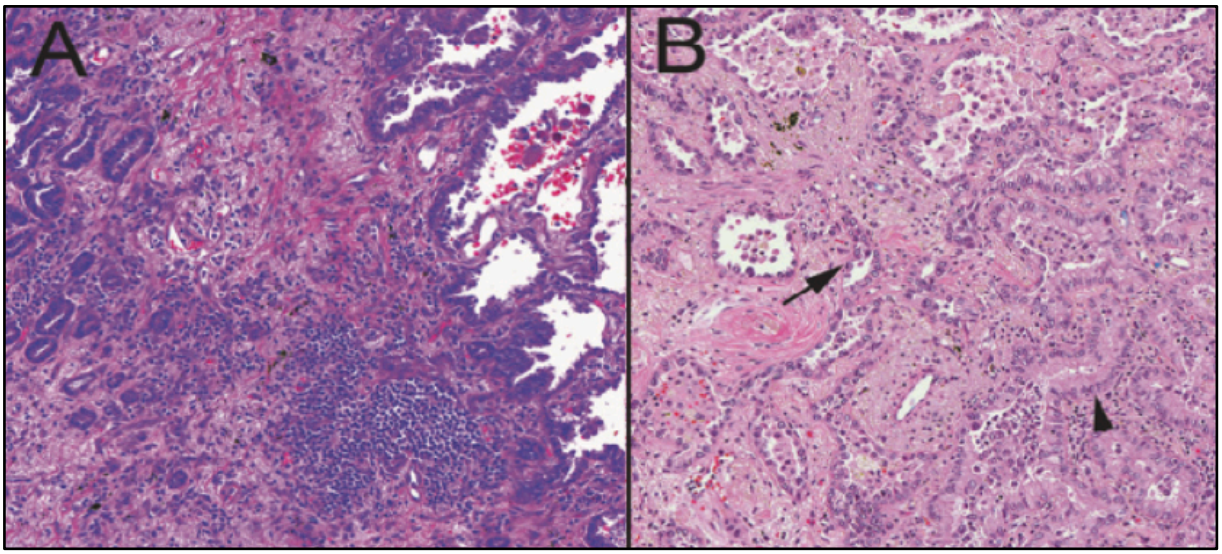
Dans certains cas, la stratification est diffuse, sans structure bien définie, rendant la différenciation d'un collapse lépidique ou d'un pattern acinaire particulièrement difficile.

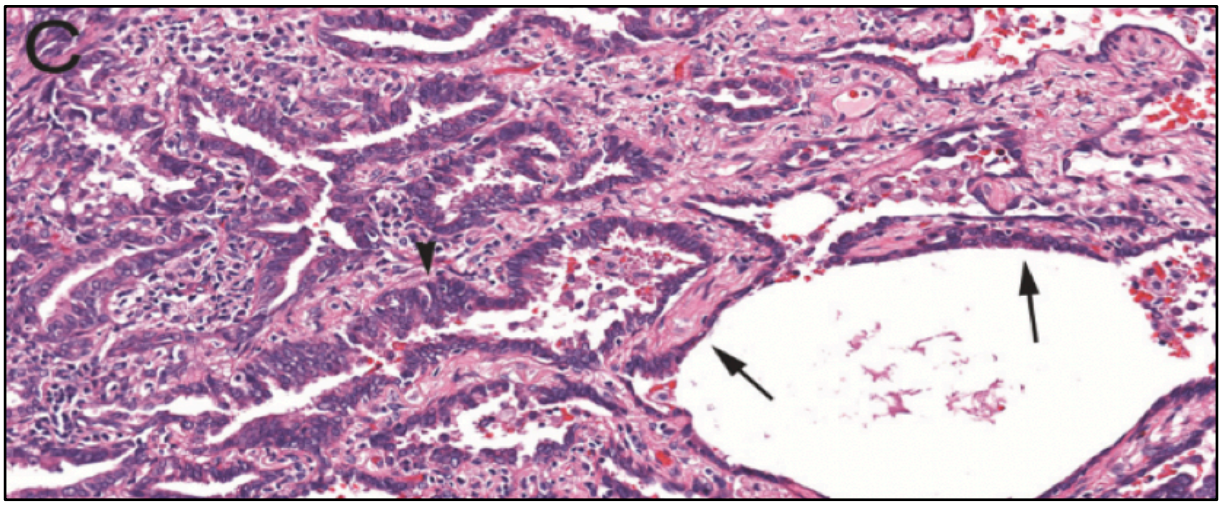


Exemple d'images complexes de collapsus alvéolaire. B. Le collapsus peut produire des images complexes, imitant des architectures complexes (**C**) avec des réactions stromales fibreuses et inflammatoires. (8)

Par ailleurs, dans certains ADK lépidiques, des cellules tumorales peuvent se détacher de la paroi alvéolaire, perdant ainsi leur disposition classique, ce qui complique davantage l'interprétation (15).

Le pattern acinaire, caractérisé par des structures glandulaires bien définies, souvent plus petites que les alvéoles normales, est parfois associé à une stratification cellulaire marquée ou à une réponse stromale sous forme de desmoplasie. Toutefois, même en l'absence de cette réaction stromale, certaines glandes peuvent être reconnues comme invasives, notamment lorsqu'elles infiltrent des structures anatomiques comme la plèvre.



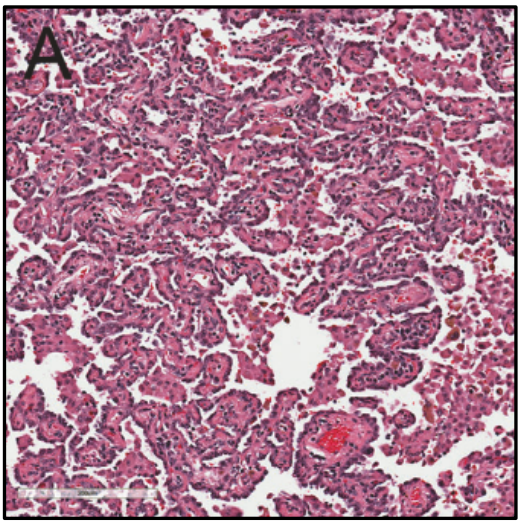


Exemples de distinction difficile entre pattern acinaire et pattern lépidique. A. Petites formations glandulaires à gauche, à côté d'espaces lépidiques soutenant le changement de pattern vers l'invasion. **B.** Bien que les deux intéressent un collapsus alvéolaire et un changement stromal, l'augmentation du grade nucléaire, de la forme et de la taille des cellules peut orienter vers un pattern acinaire par rapport à un pattern lépidique. **C.** En plus de l'atypie, les zones de transition vers l'invasion présentent souvent une stratification plus importante (➤) que le modèle lépidique (➔).(8)

La présence d'une transition morphologique entre un pattern lépidique classique et un agencement plus complexe, avec augmentation de la taille nucléaire, perte de polarité ou stratification en plusieurs couches, constitue également un indice fort en faveur d'un caractère invasif.

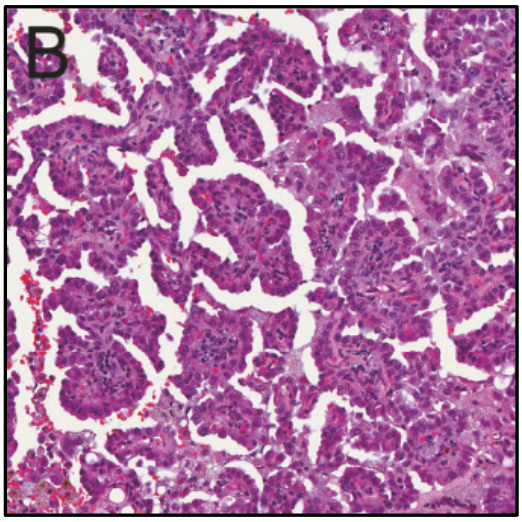
- <u>Pattern papillaire versus pattern lépidique</u>

Une autre situation difficile concerne la distinction entre les patterns papillaire et lépidique. Dans les zones de collapsus alvéolaire ou en présence d'emphysème, la croissance tumorale peut simuler des structures papillaires. Dans ce contexte, il est essentiel de distinguer les vraies structures arborescentes dotées d'un axe fibrovasculaire, caractéristiques du pattern papillaire, des replis alvéolaires épaissis sans réelle stratification.



Pattern lépidique. Le repli alvéolaire, en situation de collapsus, peut imiter un modèle papillaire, mais le revêtement est cuboïdal et monocouche, ce qui peut être comparé aux zones lépidiques sans collapsus. (8)

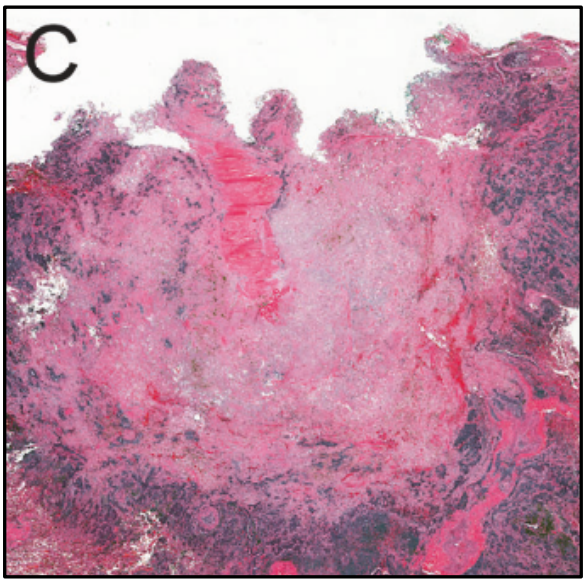
Les cellules tapissant les papilles sont généralement plus hautes, cuboïdes à cylindriques, et peuvent présenter une stratification, contrairement aux cellules « hobnail » (en clous de tapissier) typiques du pattern lépidique. Une observation attentive du contexte architectural, ainsi qu'une comparaison avec le parenchyme pulmonaire adjacent non tumoral, sont souvent nécessaires pour établir un diagnostic fiable.

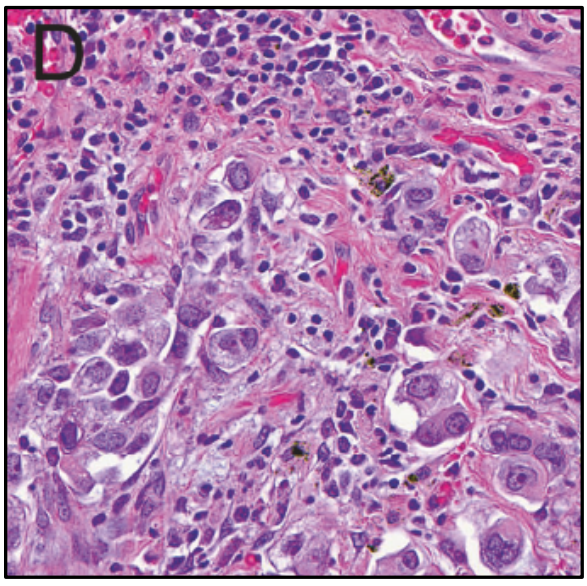


Pattern papillaire. Les structures papillaires ont une arborisation plus complexe et des cellules cuboïdales à cylindriques. (10)

L'évaluation des foyers de fibrose ou de cicatrice centrale constitue une autre source de variabilité diagnostique. Certaines zones fibroélastotiques, pauvres en cellules tumorales, mais entourées d'un pattern lépidique, soulèvent la question de leur caractère invasif (16).

Lorsque des glandes tumorales sont identifiées au sein de ces foyers, surtout en présence d'une desmoplasie, l'interprétation penche clairement en faveur d'une invasion. À l'inverse, en l'absence de cellules tumorales identifiables, ces zones ne doivent pas être comptabilisées dans la mesure de l'invasion. Parfois, une analyse attentive à fort grossissement ou l'utilisation d'une immunohistochimie cytokératine peut révéler des cellules isolées ou en petits amas, justifiant alors une reclassification en foyer invasif.





Exemple d'images complexes de foyer cicatriciel. **C.** Foyer cicatriciel pouvant apparaître paucicellulaires à faible grossissement, mais des cellules tumorales invasives comme dans (**D.**) peuvent être présentes et nécessiter une observation à plus fort grossissement pour les détecter. L'immunohistochimie cytokératine peut également être utile dans ce contexte, si nécessaire. (8)

- Pattern acinaire versus pattern micropapillaire

Enfin, la distinction entre les patterns acinaire et micropapillaire, bien qu'elle n'influe pas directement sur la mesure linéaire de l'invasion, peut avoir un impact sur le grading.

Le pattern micropapillaire est défini par la présence de projections cellulaires sans axe fibrovasculaire, souvent formées de cellules empilées en au moins trois couches, comprendre également l'organisation en « doigts ou en dentelle » (filigree pattern) (17,18).

Il est donc essentiel de définir avec précision ces zones de transition pour parvenir à une interprétation plus homogène et reproductible entre observateurs.

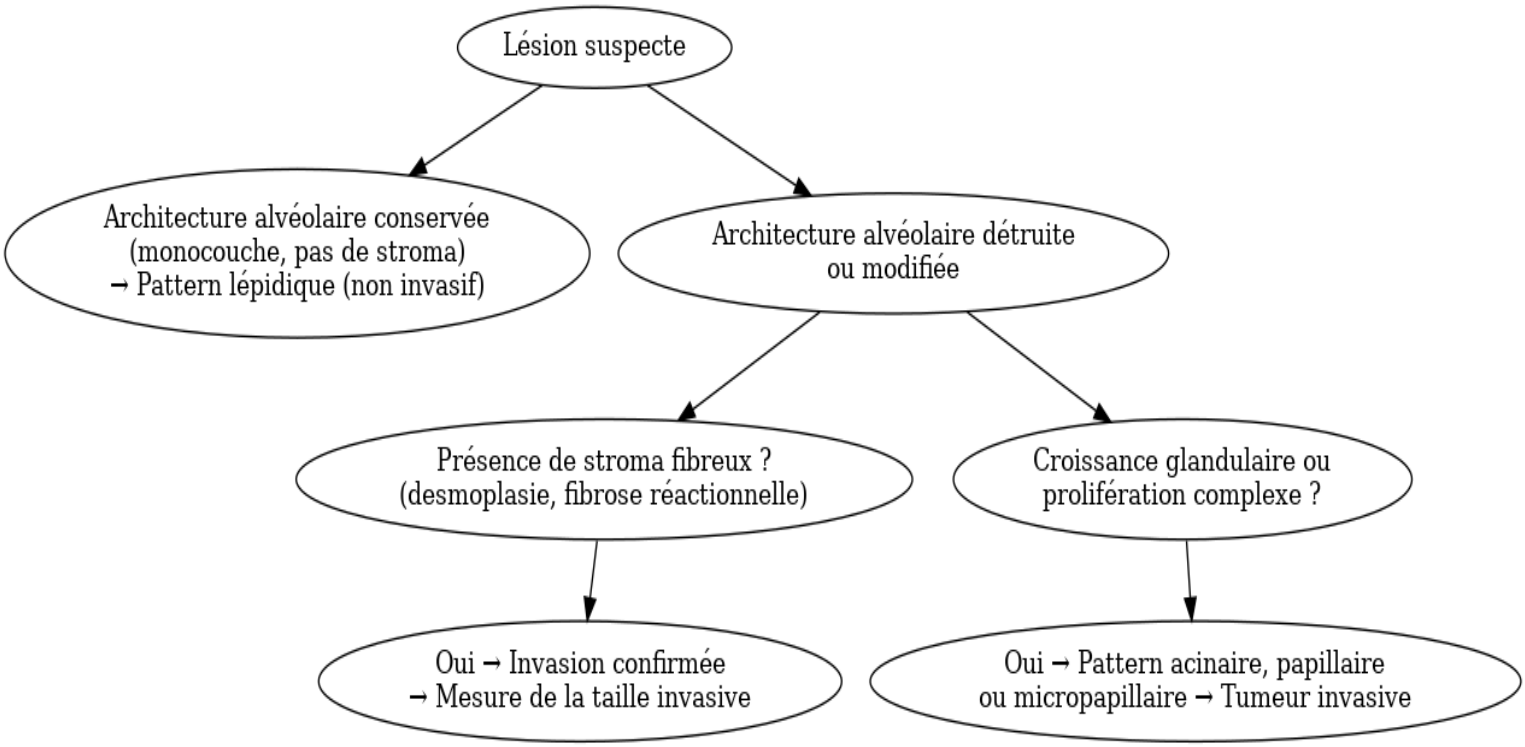


Figure 3 : **Proposition de raisonnement devant une lésion suspecte d'adénocarcinome pulmonaire.**

2. Difficultés du grading des ADK pulmonaires

Le grading d'un ADK pulmonaire est une tâche chronophage et complexe pour le pathologiste, nécessitant une évaluation minutieuse de la proportion des patterns présents sur l'ensemble des lames de la tumeur. Cette difficulté est particulièrement élevée pour les tumeurs de grande taille et présentant une importante hétérogénéité intra-tumorale. Plusieurs études ont montré que la variabilité inter-observateur dans l'évaluation du grading des adénocarcinomes pulmonaires restait importante. Ainsi, l'étude Warth et al. (J Thorac Oncol, 2012) (19) a rapporté un accord seulement modéré entre pathologistes (kappa ≈ 0,48) dans la classification histologique, tandis que dans et (7), les auteurs ont obtenu un kappa global de 0,43, avec des discordances pouvant concerner jusqu'à près d'un tiers des cas, notamment pour les patterns complexes ou moins fréquents. Ces résultats illustrent la difficulté de parvenir à une reproductibilité optimale, même entre pathologistes experts.

Face à ces défis, l'application à la Pathologie d'approches d'intelligence artificielle (IA) offre de nombreuses perspectives. L'IA pourrait, en effet, permettre « d'automatiser » l'identification et le grading des patterns tumoraux, améliorant ainsi la reproductibilité et réduisant le temps d'analyse.

L'intégration de l'IA à l'évaluation du grading des ADK pulmonaires pourrait ainsi contribuer à une meilleure stratification pronostique des patients et à une optimisation des décisions thérapeutiques.

2. L'intelligence articielle en médecine

2.1. Définition et historique

L'IA est une discipline informatique visant à créer des systèmes capables de simuler certaines fonctions cognitives humaines, telles que l'apprentissage, la reconnaissance de motifs et la prise de décision autonome. Son utilisation en médecine a connu une expansion exponentielle au cours des dernières décennies, révolutionnant les pratiques cliniques grâce à des algorithmes capables de traiter et d'analyser des volumes massifs de données médicales (20).

L'évolution de l'IA en médecine peut être retracée à travers plusieurs étapes :
- Années 1950-1980 : Développement des premiers systèmes experts basés sur des règles logiques prédéfinies.
- Années 1980-2000 : Apparition des premières méthodes d'apprentissage automatique permettant aux modèles de s'améliorer en fonction des données analysées.
- Depuis 2010 : Émergence du Deep Learning, révolutionnant l'analyse d'images médicales grâce aux réseaux neuronaux convolutifs (CNN). Le Deep Learning repose notamment sur l'utilisation de réseaux de neurones profonds, comme les « convolutional neural networks » (CNN), particulièrement efficaces dans l'analyse d'images médicales (21,22).

L'IA est aujourd'hui utilisée dans divers domaines médicaux, révolutionnant la manière dont les maladies sont diagnostiquées, traitées et surveillées.

2.2. Principes généraux

- **Machine learning et IA**

L'IA dite "apprenante" repose sur un concept fondamental : le machine learning (ou apprentissage automatique). Contrairement aux logiciels traditionnels, les algorithmes d'apprentissage automatique ne se contentent pas d'exécuter des règles préprogrammées par l'humain ; ils apprennent à partir de données. Ce processus d'apprentissage permet aux machines d'améliorer leurs performances au fil du temps. L'une de ses applications les plus performantes repose sur les réseaux de neurones artificiels, conçus sur le modèle du fonctionnement biologique du cerveau humain.

- **Du neurone biologique au réseau neuronal artificiel**

Le neurone biologique fonctionne en recevant des signaux via ses dendrites, les traite, puis transmet une réponse via son axone. En IA, un neurone artificiel simule ce processus par une fonction mathématique qui pondère les données d'entrée et produit une sortie. En combinant des milliers, voire des millions de ces neurones, on forme un réseau de neurones artificiels. Lorsqu'ils comportent plusieurs couches intermédiaires entre l'entrée et la sortie, on parle de *deep learning* ou apprentissage profond. Ces architectures permettent un traitement hiérarchisé de l'information, en passant progressivement de données simples (pixels, couleurs) à des représentations complexes et abstraites (forme d'un noyau, présence d'une glande, type de tumeur) (23).

- **L'analyse d'images médicales par deep learning**

Comme présenté dans la **figure 4**, pour analyser une image, un algorithme de deep learning extrait progressivement les informations : formes simples (bords, courbes), motifs (textures, contours), puis objets ou structures plus complexes (cellules, noyaux, tissus tumoraux). Ces étapes successives miment l'architecture du cortex visuel humain.

L'algorithme le plus utilisé pour l'analyse d'images est le réseau de neurones convolutif (CNN). Il applique des filtres (appelés convolutions) pour extraire des *features* caractéristiques d'une image (par exemple, des mitoses, des glandes, ou des zones nécrotiques dans une tumeur). Ensuite, l'information est synthétisée (via pooling), puis classifiée dans une couche finale de neurones dite *fully connected*.

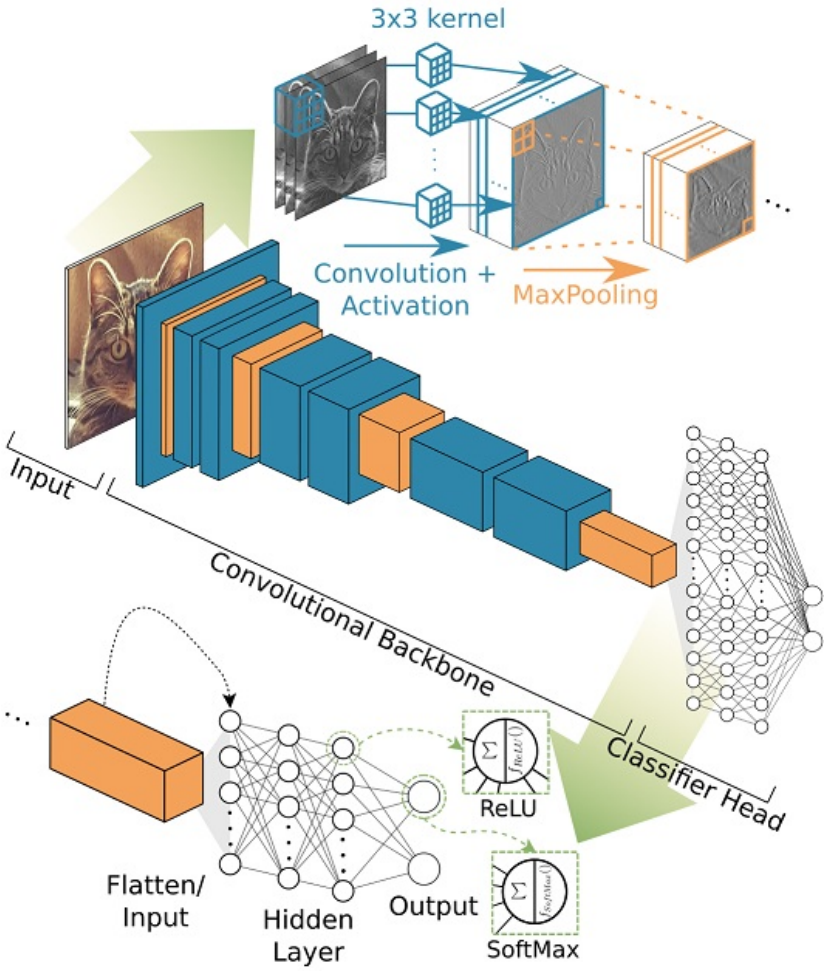


Figure 4 : **Vue d'ensemble et détaillée d'un réseau neuronal convolutif (CNN).** (24)

- **Extraction et analyse des données : comment fonctionne un CNN ?**

Les réseaux de neurones convolutifs (CNN) sont des modèles d'intelligence artificielle conçus pour analyser des images. Leur fonctionnement repose sur plusieurs étapes successives :

1. Découpage et convolution :

L'image est découpée en petites zones appelées « tuiles ». Chacune est analysée à l'aide de filtres qui détectent des motifs simples (comme des bords ou des textures). Cette étape attribue à chaque zone une valeur d'activation selon la présence ou l'intensité des motifs.

2. Pooling (réduction de dimension) **(figure 5)** :

Pour simplifier les données tout en conservant les informations importantes, on applique une opération de réduction sur chaque zone de l'image.

- Max pooling : conserve la valeur la plus élevée d'une zone.
- Average pooling : calcule la moyenne des valeurs.

Cela permet de diminuer la taille des données tout en gardant les éléments les plus significatifs.

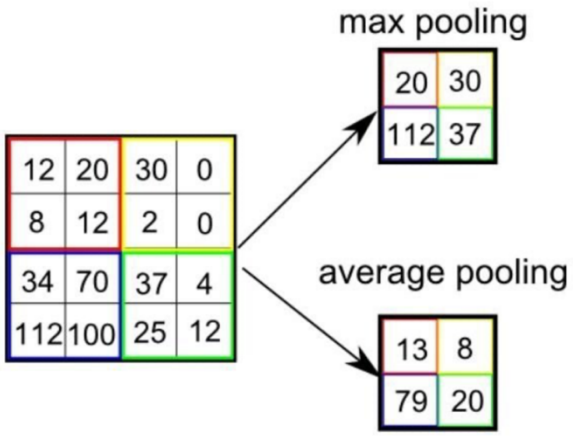


Figure 5 : **Illustration de l'étape de « pooling » d'un CNN.**

3. Apprentissage des représentations :

En enchaînant plusieurs couches de convolution et de pooling, le réseau apprend à détecter des motifs de plus en plus complexes : d'abord des lignes, ensuite des formes, puis des objets entiers (visage, cellule, etc.). Ce processus est appelé « representation learning » (24).

4. Classification finale :

Les informations extraites passent ensuite dans des couches de neurones entièrement connectées. Chaque neurone évalue les caractéristiques d'une zone. En fin de parcours, un dernier ensemble de neurones permet de classer l'image dans une catégorie (ex : tissu sain, lésion, cancer).
La classe prédite correspond à celle dont le neurone final est le plus activé (25).

- **L'entraînement : comment l'IA apprend-elle ?**

L'apprentissage d'un CNN repose sur une base d'images annotées par des experts (par exemple : « cette lame contient un adénocarcinome »). Le modèle confronte ses prédictions à la vérité de terrain (*ground truth*), calcule l'erreur (*loss*), et ajuste progressivement ses paramètres (les *poids synaptiques*) pour améliorer sa précision. Ce processus est appelé rétropropagation (*backpropagation*).

À chaque itération, l'algorithme "apprend" à donner plus d'importance aux features pertinentes et moins aux autres. L'apprentissage est considéré comme optimal lorsque l'erreur globale est minimale sur un jeu de validation indépendant (26).

- **Le problème de la "boîte noire" et les efforts d'explicabilité**

Cependant, cette profondeur algorithmique a un revers : la complexité des calculs réalisés dans les couches intermédiaires rend leur interprétation difficile, voire impossible. Les caractéristiques discriminantes utilisées par le modèle ne sont pas toujours identifiables ou compréhensibles pour l'humain. Ce phénomène, désigné sous le nom de "boîte noire" (*black box*), constitue un défi majeur pour l'intégration clinique de l'IA en Pathologie.

Cette problématique a été clairement décrite par Nagpal et al.(25). Dans ce travail, les auteurs proposent une méthode originale basée sur des vecteurs de concepts interprétables, permettant de visualiser les régions morphologiquement responsables de la prédiction algorithmique, apportant ainsi un début de réponse au problème de la boîte noire.

2.3. Applications actuelles de l'IA en médecine et en cancérologie

L'intelligence artificielle s'est imposée comme un outil puissant dans plusieurs domaines de la médecine, notamment en cancérologie, où elle permet d'améliorer le dépistage, le diagnostic et la personnalisation des traitements. Parmi ses applications majeures, on retrouve :

- Dépistage et imagerie médicale : Analyse automatisée des images radiologiques (scanner, IRM, PET-scan) permettant une détection précoce des anomalies suspectes avec une précision comparable à celle des radiologues humains (27).

- Diagnostic assisté : Systèmes basés sur l'IA intégrés aux plateformes de pathologie numérique permettant l'identification rapide des structures anormales sur les lames histologiques (20,22,27,28).

- Médecine prédictive et thérapies ciblées : Modèles d'IA capables de prédire l'évolution tumorale et d'adapter les stratégies thérapeutiques en fonction du profil moléculaire et histopathologique du patient (29).

Ces avancées montrent que l'IA est devenue un allié incontournable dans la médecine moderne, notamment en oncologie, où elle optimise le diagnostic et le traitement des cancers.

3. L'intelligence artificielle en pathologie médicale

3.1. Différents modèles d'IA

L'intégration de l'intelligence artificielle en pathologie repose principalement sur des techniques d'apprentissage automatique appliquées à des lames histologiques numérisées en haute résolution (*whole slide images*, WSI). Plusieurs approches complémentaires d'IA sont aujourd'hui utilisées, avec des niveaux de supervision, de complexité algorithmique et de généralisation variables.

- **Le Deep Learning**

Cette technique utilise des réseaux de neurones profonds capables d'analyser des images histopathologiques et d'extraire des caractéristiques complexes.

Elle s'appuie notamment sur les réseaux neuronaux convolutifs (CNN), en particulier largement utilisés pour la classification des tissus et la détection des anomalies (3,30). Ces réseaux permettent d'identifier et de classer des patterns morphologiques tumoraux à partir de petites régions d'une lame (patches) ou à l'échelle globale (WSI-

level). Les premiers succès ont porté sur la détection automatique de métastases ganglionnaires mammaires (31,32). Ces réseaux ont montré leur efficacité dans l'évaluation du grading tumoral, notamment dans les cancers du sein et du poumon (30–32).

Depuis, de nombreux CNN ont été entraînés sur des cohortes publiques (TCGA, CAMELYON, NCT-CRC-HE-100K…), avec des performances comparables à celles des pathologistes pour des tâches de diagnostic (classficiation adénocarcinome vs carcinome épidermoïde), de grading (cancer prostate, poumon) comme dans *Pan et al. (Nature 2024)* (33) **(figure 6)**, et de prédiction de biomarqueurs moléculaires comme *EGFR, KRAS, MSI, PD-L1* ou *BRAF* à partir de lames HES (31,34,35).

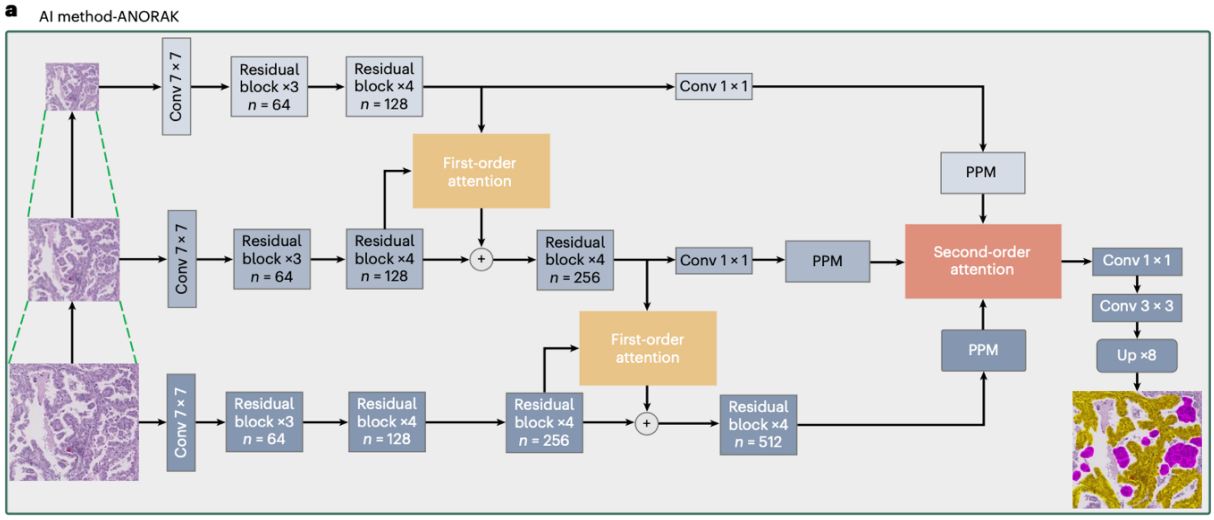


Figure 6 : **Modèle informatique proposé pour la cartographie de précision et les analyses d'hétérogénéité spatiale. a.** L'architecture du réseau d'apprentissage profond pour la segmentation des modèles de croissance intégrant les entrées sur plusieurs résolutions spatiales et fournissant des délimitations au niveau du pixel (33).

- **<u>Le Multiple Instance Learning (MIL), apprentissage faiblement supervisé et non supervisé</u>**

Le Multiple Instance Learning (MIL) permet un apprentissage avec supervision faible : la lame est annotée globalement (ex. "ADC pulmonaire"), sans annotation. Ce type de modèle est utilisé principalement dans la prédiction de la survie ou des altérations moléculaires à partir d'images entières (36–39).

Les modèles comme *HEAL* ou *REMEDIS* utilisent également l'apprentissage auto-supervisé à grande échelle pour extraire des représentations morphologiques robustes sans supervision directe (40).

- **<u>Les modèles de fondation (« foundation models »)</u>**

Les modèles de fondation représentent une nouvelle génération d'architectures d'IA capables de réaliser plusieurs tâches médicales à partir d'un seul et même entraînement. Contrairement aux approches classiques, où chaque algorithme est conçu pour une tâche spécifique (comme classer une tumeur ou détecter une mutation), ces modèles sont pré-entraînés sur des volumes massifs de données très

variées, ce qui leur permet ensuite d'être appliqués à de nouvelles tâches avec peu ou pas de réajustement.

Leur fonctionnement repose souvent sur des « transformers » (encodeurs d'image, de textes…), des architectures initialement conçues pour traiter des séquences de mots, mais adaptées ici à l'analyse d'images histologiques. Concrètement, l'image est d'abord découpée en petites zones (ou "patchs"), comme si l'on découpait une lame au microscope en milliers de vignettes. Chacun de ces patchs est analysé séparément pour en extraire des caractéristiques visuelles (features). Ensuite, le modèle réunit toutes ces informations tout en tenant compte de la position spatiale de chaque patch, ce qui lui permet de prendre en compte à la fois l'information contenue dans chaque patch et sa relation avec les patchs voisins. Cette capacité à intégrer le contexte spatial est essentielle en pathologie, où la relation entre les cellules et leur environnement immédiat est essentiel.

Ces modèles ont montré leur efficacité dans des domaines variés : classification tumorale, prédiction de biomarqueurs moléculaires (comme MSI, EGFR, HER2), segmentation de structures tissulaires, et même prédiction du pronostic (29,40–42).

Pour l'histopathologie, les « *foundation models* » sont pré-entraînés sur des données massives, incluant une importante hétérogénéité dans les données d'entrainement (qualités de coloration, scanners de lames, tissus étudiés…), avec l'objectif d'être génériques et transférables à de nombreuses tâches sans reconfiguration lourde. Parmi les modèles les plus aboutis dans le domaine de l'anatomie pathologique, il existe le modèle UNI, développé par l'équipe du Pr Faisal Mahmood, présenté dans l'article "Towards a general-purpose foundation model for computational pathology", *Nature Medicine*, 2024 (41).

Le modèle UNI **(figure 7)** est organisé en plusieurs parties qui travaillent ensemble.

- D'abord, une première partie (appelée "encodeur visuel") analyse l'image et détecte des motifs importants, un peu comme un œil numérique entraîné à repérer des formes.

- Ensuite, l'image est découpée en petits morceaux appelés *patchs* (comme des vignettes), que le modèle traite séparément ; c'est ce qu'on appelle une approche "ViT" (*Vision Transformer*).

- Le modèle rassemble toutes ces informations et tient compte de leur emplacement dans l'image (haut, bas, centre…) pour produire une analyse globale, en gardant le contexte spatial (reconstruction loss).

- Enfin, le système apprend à aligner les représentations du modèle avec celle du modèle enseignant (alignment loss).

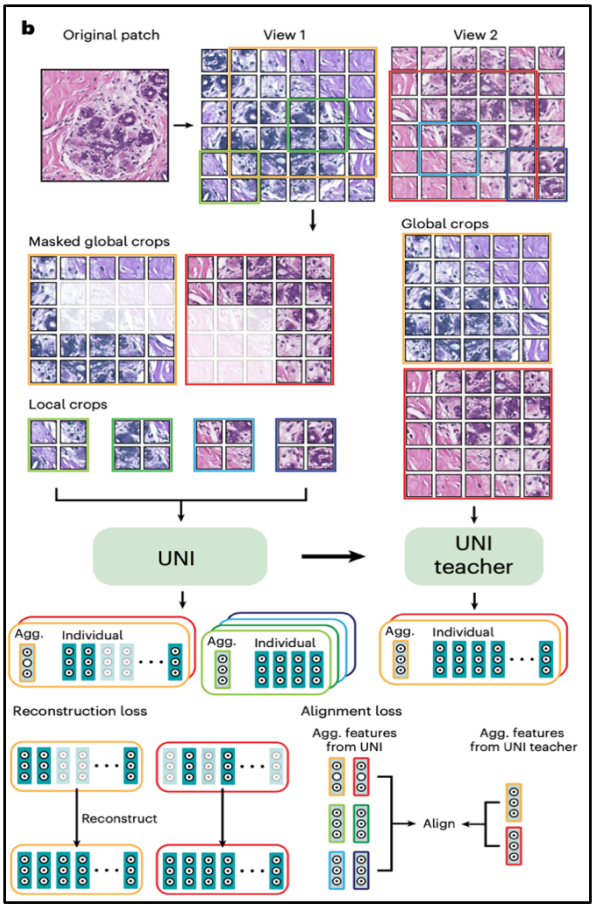


Figure 7 : **Architecture du modèle UNI et couverture des tâches cliniques** (41).

Cela permet à UNI d'apprendre à comprendre l'image sans besoin d'annotations par un Pathologiste.

Le modèle est pré-entraîné sur un gigantesque corpus de plus de 100 000 lames numérisées contenant plus de 100 millions d'images HES, intégrées dans le dataset *Mass-100K,* provenant de plus de 30 types tumoraux, incluant aussi bien des diagnostics que des annotations cliniques et moléculaires. Les images étaient prétraitées pour garantir une harmonisation inter-hôpital, puis divisées en patchs multi résolution (5x à 40x).

Il a été entraîné à l'aide d'un modèle multitâche, combinant apprentissage faiblement supervisé (labels globaux d'échantillons), et prédiction guidée de paramètres biologiques (MSI, statut ER, mutation BRAF…). L'optimisation s'est faite via l'algorithme AdamW, qui ajuste progressivement les paramètres internes du modèle pour le rendre plus performant à chaque étape, avec un apprentissage progressif par curriculum, c'est-à-dire par tâche de difficultés croissante. Les performances ont été validées par une validation croisée sur 16 centres indépendants, incluant des cohortes publiques et privées, démontrant une robustesse intersites remarquable (41).

Les modèles de fondation appliqués à l'anatomopathologie ont démontré une polyvalence remarquable dans un large éventail de tâches cliniques, dépassant souvent les approches traditionnelles supervisées. Grâce à leur capacité à encoder une représentation riche et généralisable des images histologiques, ces modèles sont mobilisables dans plusieurs domaines majeurs :

- Classification histologique : ils permettent d'identifier avec précision les types tumoraux, de différencier des sous-types (ex. carcinomes, adénocarcinomes, sarcomes) et de prédire des grades histologiques (ex. score de Gleason, grades de l'IASLC pour l'adénocarcinome pulmonaire) à partir de lames entières.

- Analyse de biomarqueurs : ils permettent d'inférer des altérations moléculaires (MSI, mutations KRAS, BRAF, EGFR, HER2, etc.) directement à partir d'images HES, sans recours à l'immunohistochimie ou au séquençage, comme démontré dans plusieurs études fondatrices (cf. "Pan-cancer image-based detection of clinically actionable genetic alterations", *Nature Cancer*, 2020).

- Évaluation du microenvironnement tumoral : ils peuvent identifier et quantifier les lymphocytes infiltrants tumoraux (TILs), les foyers inflammatoires ou la fibrose, et en extraire des indices pour des scores pronostiques, comme illustré dans l'approche développée par Ding et al. (*"Deep learning-based classification and spatial prognosis risk score on whole-slide images"*, *Histopathology*, 2023).

- Analyse multi-organes : leur architecture permet de transférer les représentations apprises d'un organe à un autre, ce qui les rend capables de s'adapter à de nouvelles localisations tumorales (sein, poumon, côlon, foie…) sans nécessiter un réentraînement complet.

- Tâches transversales : certains modèles ont été utilisés pour la segmentation de régions d'intérêt (ROIs), la recherche d'images similaires (*image retrieval*), la génération de comptes-rendus automatiques, ou encore pour des outils interactifs d'aide au diagnostic.

En résumé, les modèles de fondation jouent un rôle central dans l'évolution vers une pathologie augmentée par l'IA, en alliant performance, adaptabilité inter-tissu, et réduction des besoins en annotations. Ils constituent aujourd'hui une plateforme puissante pour le développement d'outils cliniques robustes, interprétables et déployables à large échelle.

Par exemple **figure 8**, le modèle UNI surpasse ou égale les approches supervisées spécialisées dans la plupart des benchmarks standards, tout en étant plus économe en annotations expertes (41).

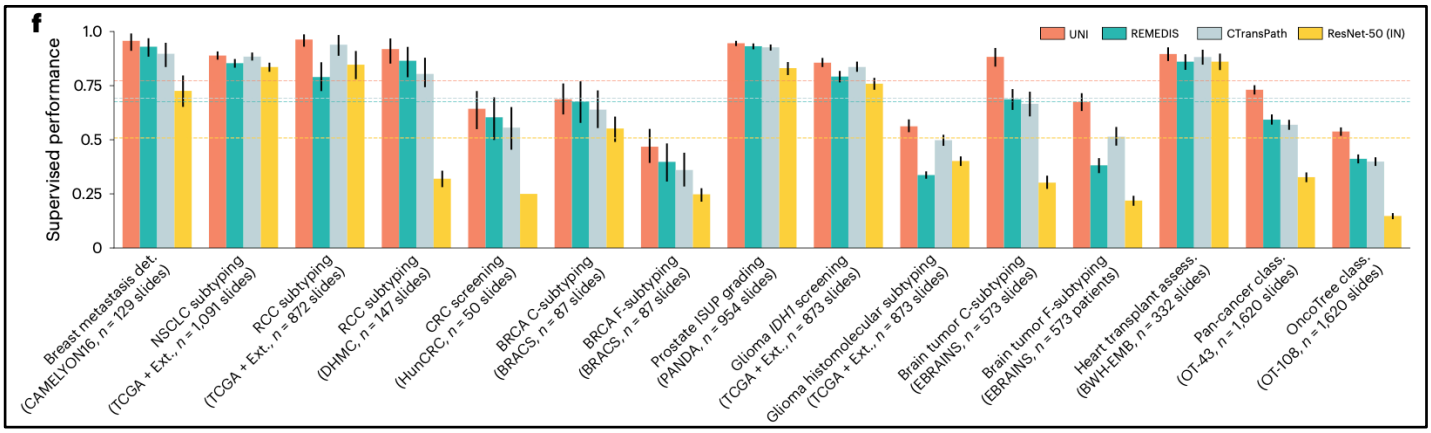


Figure 8 : **f. Performances de différents modèles de fondation (UNI, ResNet, REMEDIS, CTransPath) pour 15 tâches** (41).

3.2. IA et grading des adénocarcinomes pulmonaires

Les progrès récents de l'IA ont transformé la capacité à analyser et quantifier automatiquement les patterns histologiques dans les adénocarcinomes pulmonaires. Des modèles récents de deep learning, y compris ceux reposant sur des approches supervisées, faiblement supervisées ou auto-supervisées, ont montré qu'ils pouvaient reconnaître avec précision non seulement les patterns classiques (acinaire, papillaire, solide, lépidique, micropapillaire), mais aussi des sous-types émergents à valeur pronostique tels que cribriforme, complex glandular pattern, filigree micropapillaire et discohésif.

Ainsi, quelques travaux ont proposé des modèles d'IA, principalement fondés sur des réseaux de neurones convolutifs (CNN), afin d'automatiser et de standardiser cette évaluation. Certains de ces modèles se limitent à une classification globale de la tumeur, tandis que d'autres, plus avancés, réalisent une segmentation sémantique pixel-par-pixel permettant d'identifier la distribution spatiale des différents patterns (lépidique, acinaire, micropapillaire, etc.). Le tableau **(figure 9)** suivant présente une synthèse comparative des principaux modèles récents, en détaillant leur architecture,

leur tâche principale, leurs performances quantitatives, ainsi que leur capacité à détecter et caractériser les patterns tumoraux.

Article (année)	Modèle IA	Type de modèle IA	Tâche principale	Performance principale
Lami et al. (Am J Pathol 2023)(43)	CNN AI-1 / CNN AI-2	CNNs standards (ResNet-like)	Classification + Grading	F1-score > 0.90
Pan et al. (Nature Cancer 2024)(33)	CNN ANORAK	Pyramid Pooling Cross-stream Attention Network	Segmentation sémantique + Grading IASLC	Dice = 0.6034 ; Accord 60–65 % ; κ = 0.46
Miura et al. (Jpn J Clin Oncol 2024)(44)	CNN Keio	CNN classique sur tuiles représentatives	Subtyping + analyse spatiale	Accuracy = 90.4 % ; Accord 75.5 %
Ding et al. (Histopathology 2023)(37)	LSDLM (Lung Spatial Deep Learning Model)	CNN ResNet34 + classifieur dense	Classification + score spatial pronostique	AUC = 0.93–0.96
Zhao et al. (BMJ Open 2023)(45)	CNN DeepLab v3	Segmentation sémantique CNN (DeepLab v3 + ResNet-50)	Segmentation + classification de patterns	AUC > 0.8 (sauf acinar) ; NDCG > 75 %
Wei et al. (Scientific Reports 2019)(46)	CNN + vote	CNN avec agrégation sur WSI	Classification du pattern prédominant	Kappa = 0.525 ; Accord = 66.6 %

Figure 9 : **Tableau comparatif des performances de modèles IA récents dans la caractérisation des patterns tumoraux des ADK pulmonaires.**

PROBLEMATIQUE

L'essor de l'IA en pathologie ouvre de nouvelles perspectives pour la reconnaissance morphologique automatisée. Toutefois, malgré les performances remarquables des modèles d'apprentissage profond dans des tâches spécifiques, plusieurs limites structurelles, techniques et cliniques entravent encore leur adoption généralisée en pratique courante. Ces obstacles concernent la généralisation des modèles, la transparence décisionnelle, la dépendance aux annotations humaines, et l'intégrabilité clinique.

Premièrement, les modèles de Deep Learning, notamment les réseaux neuronaux convolutifs (CNN), reposent généralement sur un entraînement supervisé massif à partir d'images annotées manuellement par des pathologistes. Or, plusieurs limitations majeures émergent. Tout d'abord ces modèles sont dépendants au processus d'annotations manuelles exhaustives, qui est extrêmement chronophage, coûteux, et source de variabilité inter-observateur. Il constitue un goulot d'étranglement majeur, en particulier pour des tâches nécessitant une annotation pixel par pixel (segmentation des patterns, invasion, marges) (47).

Il faut aussi relever que la généralisation de ces modèles est une tâche difficile. Les modèles supervisés entraînés sur des bases internes montrent souvent une chute de performance sur des données issues d'autres centres, en raison des variations dans la préparation des lames, les scanners utilisés, ou la représentation ethnique et histologique des cohortes (29).

Un autre reproche fréquemment adressé aux modèles de Deep Learning est leur manque de transparence décisionnelle. Les algorithmes peuvent atteindre une performance équivalente à celle d'un expert, mais leur mode de décision reste difficilement compréhensible pour l'humain, notamment avec le concept de boîte noire.

En effet, les modèles CNN classiques ne donnent que rarement accès à une justification visuelle ou textuelle de leur prédiction. Cela limite la confiance que les pathologistes peuvent accorder à ces systèmes (25). C'est pour cela que certaines études se sont attachées à se détourner de ce modèle prédictif rapide, mais opaque de chercher à développer un modèle prédictif plus visuel et qui s'appuie sur des concepts morphologiques et sémantique de la décision (25).

De nos jours, malgré les avancées techniques, peu de modèles sont réellement utilisés dans les services hospitaliers de pathologie. Les raisons en sont multiples : difficulté de validation réglementaire, infrastructure informatique inadaptée, ou manque de confiance des cliniciens (48).

Aujourd'hui, il est donc intéressant de réfléchir aux perspectives d'avenir concernant les modèles IA qui pourront potentiellement être généralisés et appliqués à notre activité courante de pathologiste. Les modèles de fondation peuvent apporter une réponse à certaines de ces limites. Ils représentent une avancée majeure par rapport aux architectures convolutionnelles classiques (CNN) dans le champ de la pathologie numérique. Contrairement aux CNN, qui sont généralement entraînés sur des tâches

spécifiques avec des jeux de données limités et fortement annotés, les modèles de fondation reposent sur un pré-entraînement auto-supervisé ou multimodal à grande échelle, souvent sur des centaines de milliers de lames histologiques. Cette approche leur confère une capacité de généralisation supérieure, leur permettant d'être réutilisés pour de multiples tâches en apprentissage par transfert ou en inférence directe (*zero-shot learning*).

Par exemple, le **modèle UNI** présenté dans l'article de Chen et al. (Nature, 2024)(41) est pré-entraîné de manière auto-supervisée sur plus de 100 millions de patches issus de 100 000 lames, ce qui lui permet de surpasser les CNN classiques dans des tâches telles que le sous-typage tumoral ou la prédiction moléculaire sans avoir recours à une annotation manuelle coûteuse.

De même, le **modèle CHIEF**, décrit dans Wang et al. (*Nature*, 2024), est capable d'extraire des représentations histologiques robustes à partir de 60 000 lames, entraîné selon un paradigme dual mêlant pré-entraînement auto-supervisé et supervision faible. Ce modèle démontre une capacité à détecter des caractéristiques morphologiques pertinentes dans des données bruitées ou incomplètes, ce qui est souvent un point faible des CNN entraînés de manière supervisée.

Le **modèle Virchow**, détaillé dans Vorontsov et al. (Nature, 2024) (40), montre en outre que les modèles de fondation permettent d'atteindre une performance clinique dans des contextes rares ou à faible volume de données, grâce à leur robustesse aux variations de distribution et aux artefacts techniques.

Enfin, l'intégration de modalités multiples, comme dans le **modèle CONCH** de Lu et al. (*Nature*, 2024)(42), ouvre la voie à une interprétation plus riche en combinant images histologiques et rapports textuels, ce qui demeure hors de portée des CNN traditionnels.

Ces travaux illustrent que les modèles de fondation constituent aujourd'hui une brique technologique stratégique pour l'avenir de la pathologie computationnelle, en offrant une scalabilité, une transférabilité et une efficacité accrues dans des contextes cliniques variés.

Critères de comparaison	CNN classiques	Modèles de fondation
Méthode d'apprentissage	Apprentissage supervisé spécifique à une tâche	Pré-entraînement auto-supervisé ou multimodal à grande échelle
Taille du jeu de données requis	Jeux de données de taille moyenne	Jeux de données massifs (> 50 000 lames)
Besoin en annotation manuelle	Annotation fine requise (ROI, pixel)	Peu ou pas d'annotation requise pour le pré-entraînement
Capacité de généralisation	Faible à modérée (fortement dépendant de la tâche)	Excellente, même sur des tâches non vues
Performance sur tâches variées	Limitées à la tâche d'entraînement	Haute performance sur de nombreuses tâches (zero-shot, few-shot)

Adaptabilité au transfert d'apprentissage	Moins adapté (fine-tuning requis)	Très bien adapté (représentations réutilisables)
Robustesse aux variations (artefacts, scanners, laboratoires)	Sensible aux variations techniques	Haute robustesse démontrée
Capacité à intégrer plusieurs modalités (image + texte)	Non (image seule)	Oui (ex : modèle CONCH)
Utilisation en contexte de données rares ou déséquilibrées	Moins efficace	Très efficace (ex : modèle Virchow sur cancers rares)

Figure 10 : **Tableau comparatif des convolutional neural network (CNN) classiques et des modèles de fondation**

La **figure 10** met en évidence que les modèles de fondation représentent donc une nouvelle approche d'IA particulièrement prometteuse pour la reconnaissance d'images histopathologiques complexes. Cependant à ce jour, cette approche n'a jamais été appliquée à la reconnaissance des patterns et le grading des ADK pulmonaires.

OBJECTIFS DE THÈSE

Les objectifs principaux de ce travail sont d'évaluer la performance d'un modèle de fondation « *H-optimus-0* » pour la classification des patterns et le grading des adénocarcinomes pulmonaires.

MÉTHODOLOGIE

1. Patients et échantillons

Cette étude, présentée **figure 11**, a été menée au sein du Département de Pathologie du CHU Bichat-Claude Bernard (AP-HP, Paris). Elle repose sur une cohorte monocentrique de 400 patients opérés d'un ADK pulmonaire (stade localisé ou localement avancé), entre le 1er janvier 2015 et le 31 décembre 2018. Chaque patient inclus a bénéficié d'une résection chirurgicale dont la pièce opératoire a été traitée selon les protocoles macroscopiques standard. Les prélèvements tumoraux ont été inclus en paraffine, coupés, colorés à l'hématoxyline-éosine-safran (HES), puis montés sur lames de verre.

Un échantillon représentatif de 122 patients a été sélectionné de manière aléatoire à partir de cette cohorte, en veillant à une répartition équilibrée entre hommes et femmes. Les critères d'inclusion comprenaient :

- un diagnostic confirmé d'adénocarcinome pulmonaire non mucineux ;
- l'absence de traitement néoadjuvant;
- l'existence d'un suivi clinique documenté ;

A partir de ces 122 patients, deux cohortes ont été définies :

- une cohorte d'entrainement de 36 patients
- une cohorte de validation de 86 patients

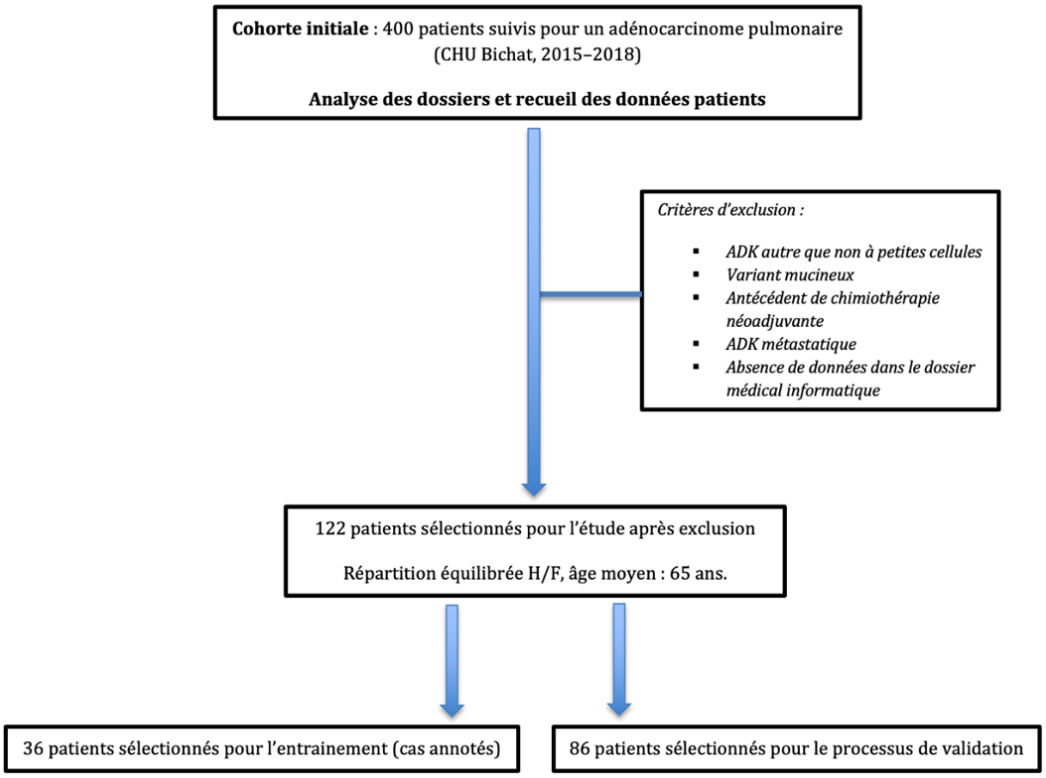


Figure 11 : **Diagramme de flux présentant la sélection des patients.**

Pour chaque patient, une lame représentative de la tumeur a été sélectionnée, anonymisée via un système d'identifiants codés (ADK-1, ADK-2, etc.), puis numérisée à l'aide d'un scanner Aperio GT 450 (Leica) (Dpt de Pathologie, Bichat).
Chaque lame numérisée, en format SVS, a ensuite été analysée indépendamment par deux observateurs : un pathologiste sénior expert en pathologie thoracique (Dr N Poté) et un pathologiste junior interne en 4e année de spécialité en anatomie pathologique (Dr M Gassier).
L'analyse a porté sur la morphologie tumorale, avec pour objectif principal l'établissement du grading histologique selon les critères actualisés de l'IASLC 2025 (International Association for the Study of Lung Cancer).

2. Annotation des lames (cohorte d'entrainement)

Pour 36 patients, une lame histologique représentative de la tumeur a été sélectionnée. Sur chacune de ces lames, deux niveaux d'annotation distincts ont été réalisés conjointement par les deux Pathologistes, à l'aide de l'outil numérique d'annotation « QuPath » :

1. **Délimitation de la zone tumorale** : chaque tumeur a été annotée précisément afin d'exclure les artefacts, les zones non tumorales ou les tissus annexes (parenchyme adjacent, stroma extratumoral).
2. **Annotation des patterns tumoraux** : dans la zone tumorale définie, les différents patterns architecturaux identifiables ont été annotés (acinaires, papillaires, micropapillaires, solides, cribriformes, etc.). Pour chaque pattern identifié, au moins cinq zones distinctes ont été annotées, assurant une représentativité spatiale et morphologique de la distribution tumorale sur la lame.
3. **Annotation des structures non tumorales d'intérêt** :
    - les zones de fibrose tumorale,
    - les foyers inflammatoires intra ou péritumoraux,
    - la présence de nécrose,
    - les structures bronchiques (cartilage),
    - ainsi que les vaisseaux intratumoraux ou extratumoraux

Les annotations réalisées sur les lames histologiques ont été effectuées de manière rigoureuse, en suivant un code couleur prédéfini permettant de distinguer la zone tumorale et les différents patterns tumoraux **(Figure 12)**.

Ainsi, le cerclage de la tumeur était systématiquement annoté en rouge, tandis que les patterns histologiques étaient distingués comme suit : acinaire en violet, lépidique en vert, cribriforme en jaune, micropapillaire en cian, solide en bleu foncé, et papillaire en orange. Les zones de nécrose étaient annotées en noir, le stroma inflammatoire en magenta, le stroma fibreux en vert d'eau, les tissus non tumoraux en marron, et enfin les zones d'absence de tissu (liées à des artefacts de lame) étaient annotées en pourpre.

Ce système d'annotation standardisé visait à assurer une homogénéité et une reproductibilité dans l'analyse morphologique des cas. Les résultats des annotations sur « QuPath », à partir des fichiers SVS, ont ensuite été agrégés et enregistrés sous forme de fichiers GeoJSON.

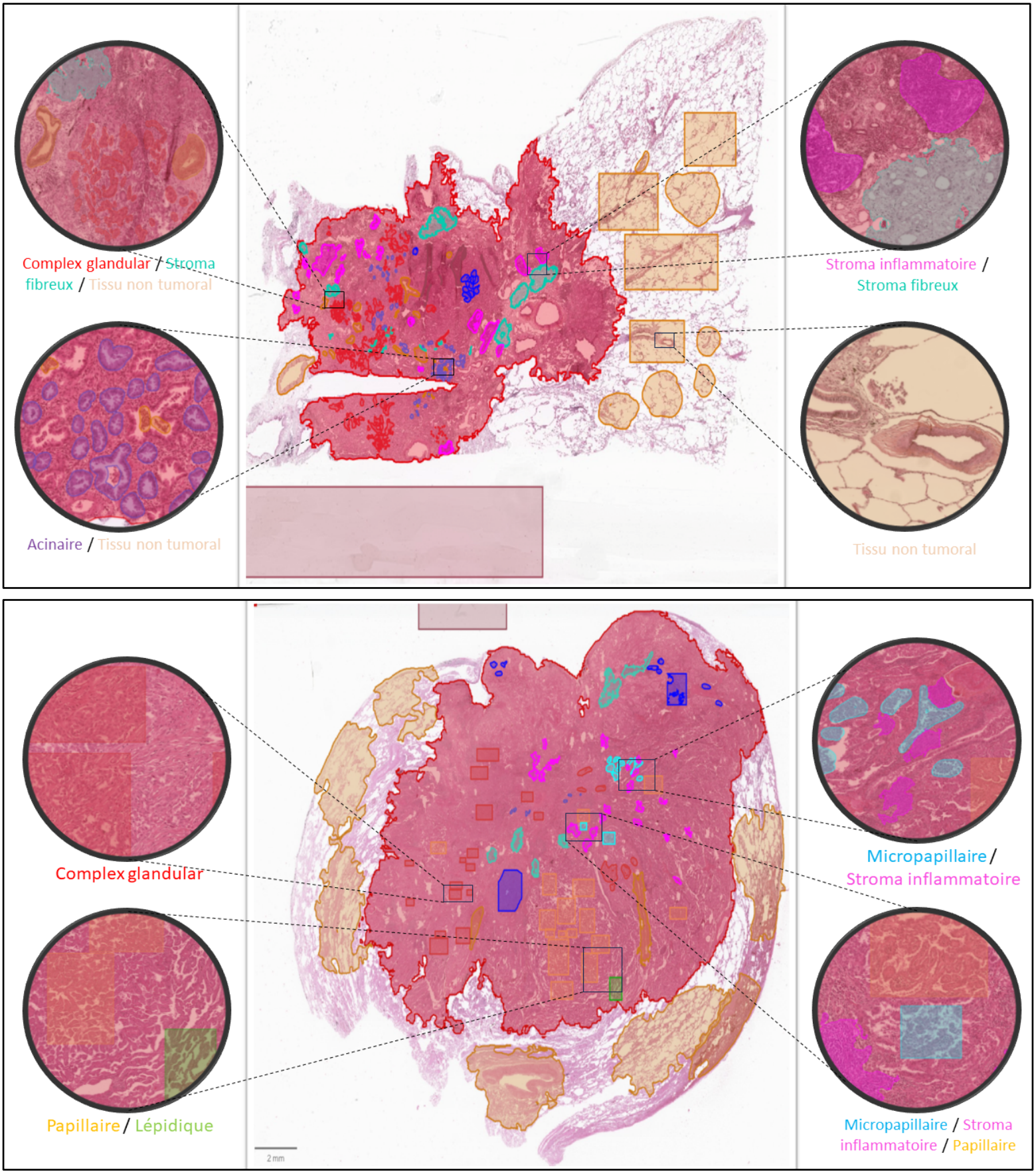


Figure 12 : **Exemples d'annotations réalisées sur lames numérisées (logiciel QuPATH).**

3. Grading des tumeurs

Sur l'ensemble des 122 patients inclus dans l'étude, les deux pathologistes ont réalisé, indépendamment, le grading tumoral selon les recommandations de l'IASLC 2025 sur une lame représentative de la tumeur. Chaque lame tumorale a été évaluée de manière systématique, en quantifiant le pourcentage de chaque pattern histologique observé. Ce travail a permis de consolider la fiabilité des annotations en confrontant les évaluations inter-observateurs et d'établir un référentiel de grading robuste pour l'analyse ultérieure par le modèle de fondation.

4. Modèle de classification : Foundation model *« H-Optimus-0 »*

4.1 Description du modèle *H-Opimus-0*

H-Optimus-0, comme UNI (modèle IA abordé précédemment dans la section 1.1.3.1 Les différents modèles d'IA), repose sur une architecture de type transformer, optimisée pour analyser directement des images de lames entières (WSI) via une représentation multi-échelle et hiérarchique. Le modèle est structuré autour d'un encodeur visuel en patchs (façon ViT) et d'un module d'agrégation globale intégrant la position spatiale des éléments histologiques. Cette architecture permet aux modèles de fondation de traiter de multiples résolutions simultanément, et de produire une représentation morphologique unifiée utilisable pour plusieurs tâches (classification, prédiction moléculaire, segmentation).

H-Optimus-0 a été testé de manière comparative dans plusieurs études internes et publications tierces portant sur l'évaluation de différents modèles de fondation, dans lesquelles il a démontré une supériorité de performance pour de nombreuses tâches (classification, segmentation, détection) (49) **(Figure 13)**. Cette convergence d'arguments techniques, pratiques et scientifiques nous a motivé à sélectionner ce modèle pour notre protocole expérimental.

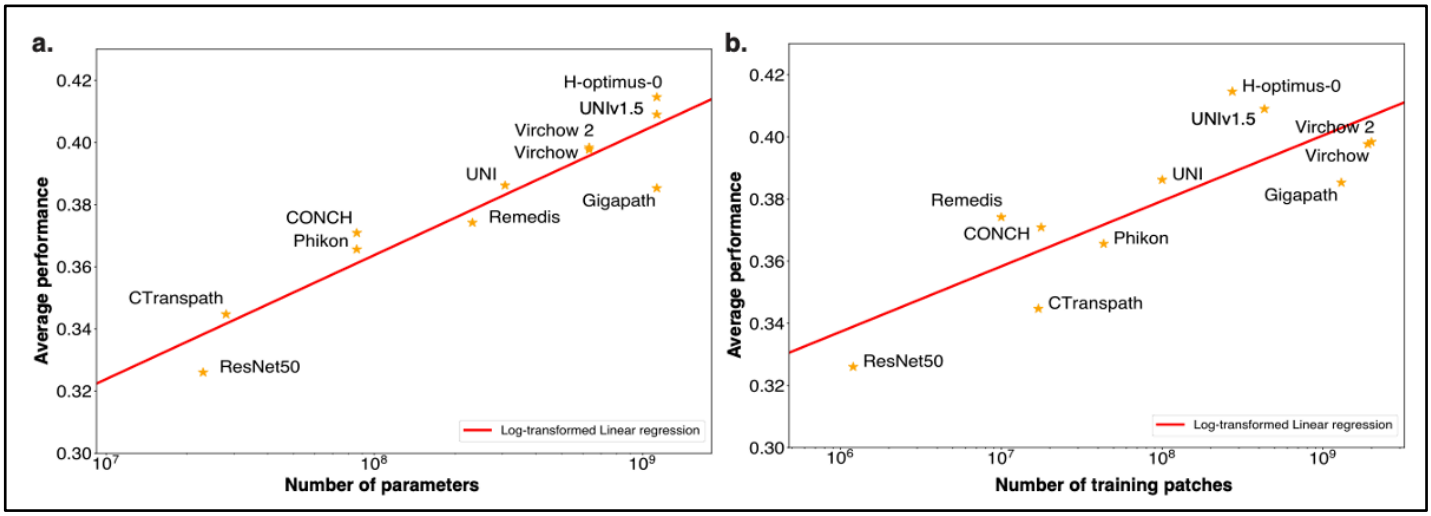


Figure 13 : **Tableau comparatif des performances des modèles de fondation les plus récents** (49). **a.** relation entre la performance moyenne et le nombre de paramètres. **b.** rapport de la performance moyenne et du nombre de patches utilisés lors du pré-entraînement.

La **Figure 13** présente une comparaison des performances moyennes des modèles de fondation les plus récents en pathologie numérique. Le **graphique a.** illustre la relation entre la performance moyenne et le nombre de paramètres des différents modèles. On observe une tendance globale à l'amélioration des performances avec l'augmentation de la taille des modèles, allant des architectures plus modestes comme **ResNet50** et **CTransPath**, jusqu'aux modèles de nouvelle génération tels que **H-optimus-0**, **UNI v1.5** et **Virchow 2**, qui se situent parmi les plus performants.
Le **graphique b.** met en parallèle la performance moyenne et le nombre de patches utilisés lors du pré-entraînement. Là encore, la tendance générale montre que les modèles ayant bénéficié d'un apprentissage sur des bases massives (par exemple **UNI**, **Gigapath** ou **Virchow**) atteignent des performances supérieures, confirmant l'importance de la taille des données dans la qualité des modèles de fondation. Ces deux représentations soulignent que la scalabilité, à la fois en termes de paramètres et de volume de données, constitue un facteur déterminant dans la supériorité des modèles de fondation récents par rapport aux approches plus classiques.

5. Protocole du modèle de classification

Le modèle de classification a été réalisé en plusieurs étapes **(figure 15)** :

1. Dans un premier temps, nous avons appliqué le modèle « H-Optimus-0 », dont l'objectif initial est de **distinguer les zones tumorales des zones non tumorales** sur des lames scannées (WSI). Une fois la zone tumorale identifiée et isolée, le modèle a été appliqué uniquement à cette région dans le but de caractériser les différents patterns histologiques présents.

2. Le modèle H-Optimus-0, basé sur l'architecture Vision Transformer, traite chaque image en la découpant en petits patchs carrés, chacun étant transformé en une représentation numérique exploitable par le modèle. Un « vecteur mathématique de classification » est ajouté à la séquence de patchs afin de résumer l'information globale de l'image, et un encodage positionnel est intégré pour indiquer la localisation de chaque patch dans l'espace.

3. Ces représentations sont ensuite traitées par une série d'unités de mesure contextuelle, où chaque patch peut interagir avec tous les autres grâce à un mécanisme d'attention globale. Cela permet au modèle de capter des relations complexes, même à longue distance, entre les structures tissulaires. La sortie finale est un ensemble d'environ 1500 valeurs (features) qui vont être ensuite transmise à une couche de classification (classifier).

Chaque patch tumoral identifié est alors associé à un pattern (par exemple : lépidique, acinaire, solide…), sur la base des caractéristiques apprises par le modèle. Une fois tous les patchs catégorisés, la proportion globale de chaque pattern sur la zone tumorale est calculée **(figure 14)**. Le label final attribué à la lame correspond au pattern le plus représenté au sein de cette zone. Pour l'établissement du grading, ce sont les recommandations de l'IASLC 2025 qui sont appliquées.

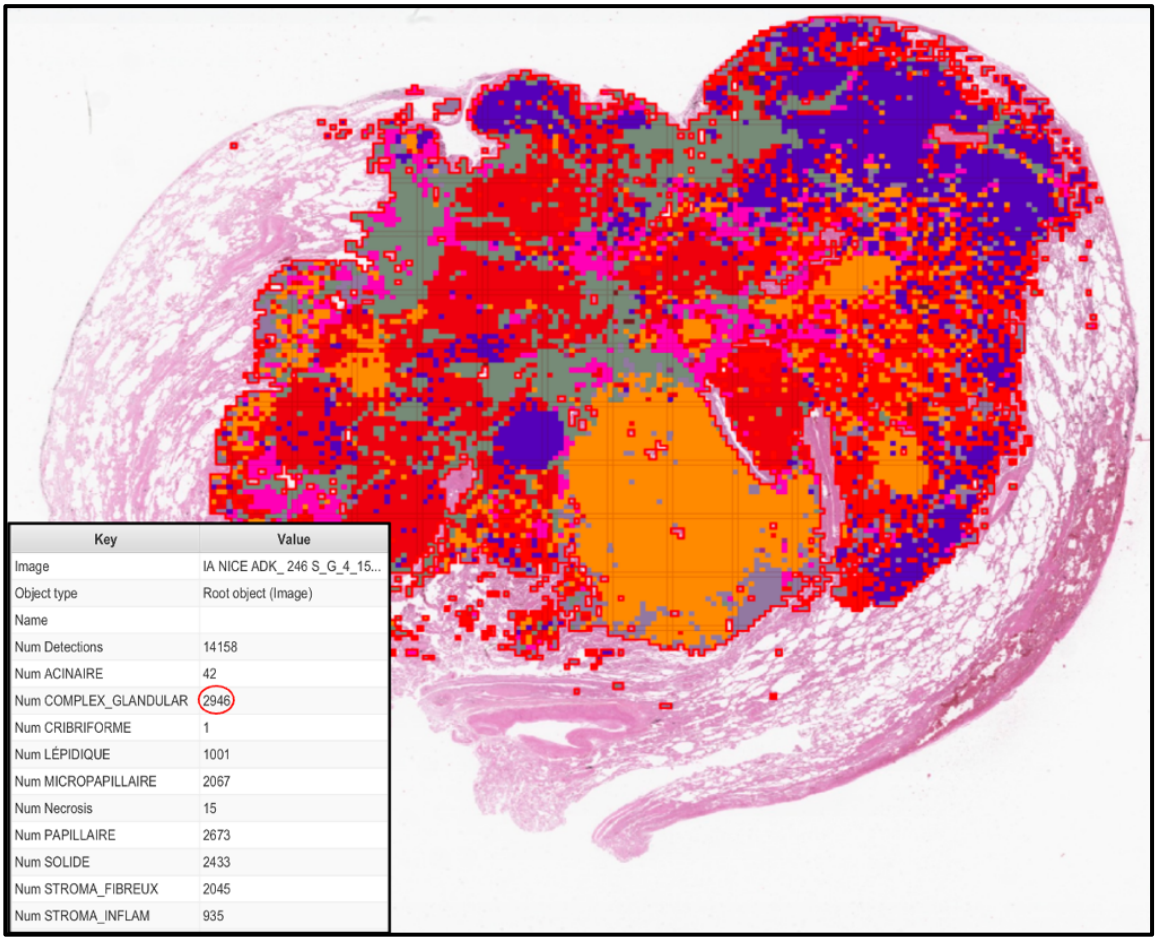


Figure 14 : **Evaluation par le modèle de fondation «** H-Optimus-0 » **de la proportion des patterns sur une lame numérisée d'adénocarcinome pulmonaire.**

4. Afin d'optimiser le modèle, une **étape de lissage** a été réalisée, visant à atténuer les artefacts liés à la détection de patterns isolés. Un seuil minimal de 9 patchs a été défini : ainsi, lorsqu'un patch est attribué à un type de pattern et que les huit autres autour de lui sont attribués à un autre et même pattern, alors ce dernier est « lissé » au groupe de patchs environnements. Cette approche reproduit la manière dont un pathologiste interprète une lame, en raisonnant par régions continues plutôt que patch par patch.

5. Enfin, une **étape de pondération des valeurs brutes** a été appliquée, afin de minimiser le poids des patterns de haut grade: un facteur de pondération est appliqué à chaque pattern détecté, en fonction de son pronostic (pattern de bas grade: pondération x1 ; pattern de haut grade : pondérationx0,8).

L'association de ces deux dernières étapes permet d'obtenir une classification plus robuste et cohérente, en renforçant la lisibilité morphologique des patterns dominants, et en assurant un grading plus fidèle aux critères histologiques standardisés.

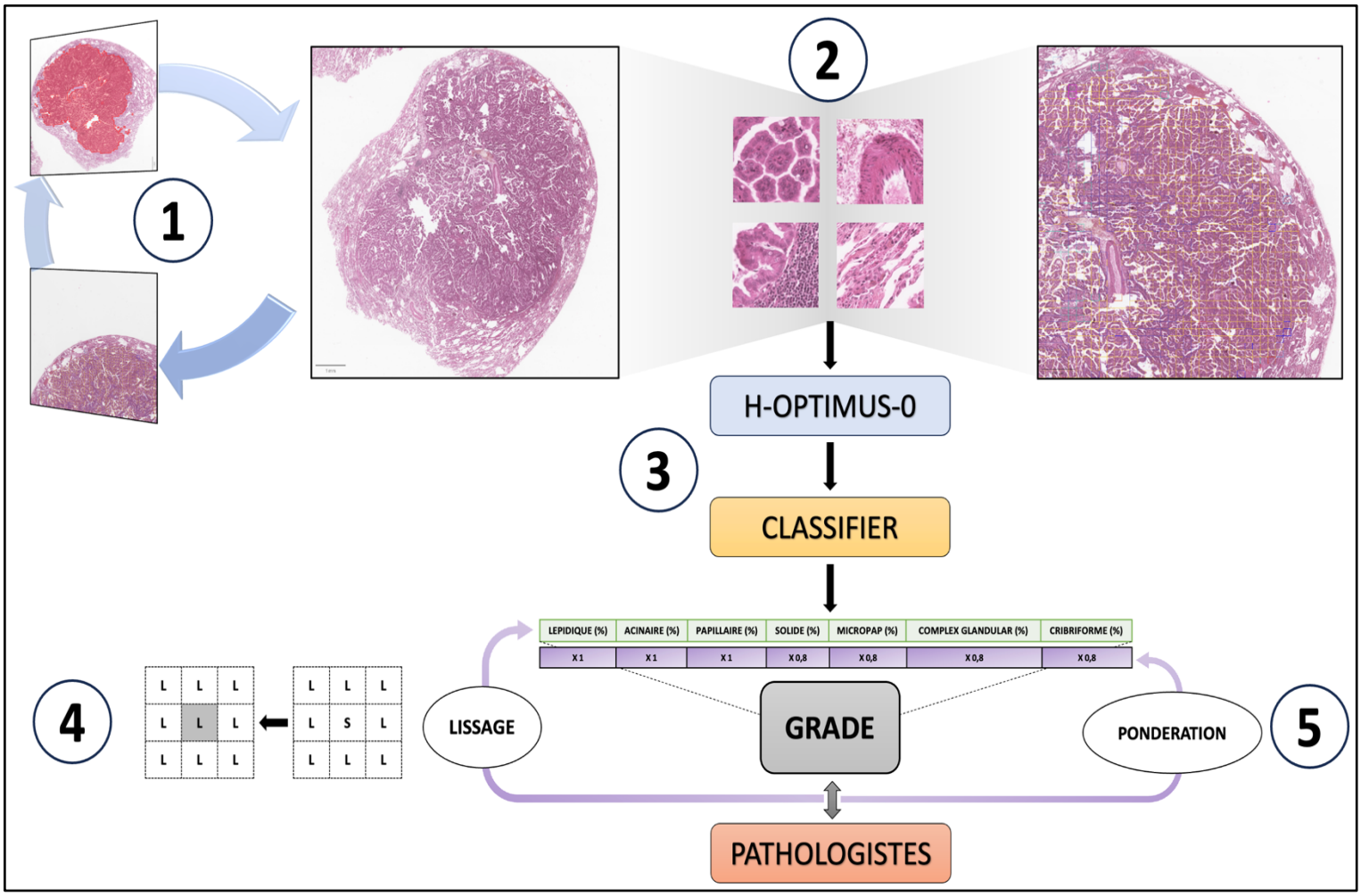


Figure 15 : **Protocole de classification du grading des adénocarcinomes pulmonaires, basé sur le modèle de fondation « H-Optimus-0 ».**

6. Analyses statistiques

Concernant les données clinico-pathologiques, les variables quantitatives ont été comparées à l'aide du test de Mann Withney et les variables qualitatives à l'aide des tests de Ficher et du Chi². Une *p*-value <0.05 était considérée comme statistiquement significative. Les analyses statistiques ont été réalisées à l'aide du logiciel PRISM.

Concernant le modèle de fondation, l'ensemble des analyses statistiques a été réalisé avec le langage de programmation Python (version 3.10.13), en utilisant notamment les bibliothèques pandas pour la manipulation des données, scipy pour les tests statistiques, ainsi que seaborn et matplotlib pour la visualisation.

- Données qualitatives et concordance inter-annotateurs vs IA

Pour chaque lame, le pattern histologique dominant a été défini à partir des annotations du pathologiste junior, du pathologiste senior et des différentes variantes du modèle d'intelligence artificielle (IA brute, lissée, pondérée, etc.). La classe majoritaire a été déterminée à partir des probabilités associées à chaque pattern, puis les prédictions ont été comparées deux à deux. Des matrices de confusion normalisées ont été établies pour chaque comparaison (IA vs junior, IA vs senior, junior vs senior), offrant

une visualisation claire des accords interclasses. Ces matrices ont été représentées sous forme de heatmaps, avec une échelle de couleurs normalisée par colonne. Les comparaisons entre sous-groupes ont ensuite été évaluées à l'aide du test du Chi², de l'indice de Kappa simple et de l'indice de Kappa pondéré.

- Comparaisons appariées des proportions

Une analyse appariée a été conduite sur les proportions attribuées à chaque pattern histologique selon les différentes méthodes (junior, senior, IA), en utilisant le test de Wilcoxon pour données appariées. Les p-valeurs correspondantes ont été indiquées directement sur les boxplots (*$p < 0,05$ (*), $p < 0,01$ (**), $p < 0,001$ (***), et $p < 0,0001$ (****)*).

RÉSULTATS

1. Caractéristiques anatomo-cliniques des cohortes

		COHORTE D'ENTRAINEMENT (n=36)	COHORTE DE VALIDATION (n=86)	*p*
AGE moyen (min-max)		66,5 (77-44)	64,6 (82-47)	ns (0,23)
SEXE (h), n(%)		23 (64%)	43 (50%)	ns (0,17)
TABAGISME en paquet-année moy (min-max)		41 (0-80)	35 (0-120)	ns (0,08)
pT, n(%)	pT1	10 (27%)	35 (41%)	ns (0,14)
	pT2	22 (62%)	38 (44%)	
	pT3	4 (11%)	7 (8%)	
	pT4	0 (0%)	6 (7%)	
pN, n(%)	N0	24 (67%)	67 (78%)	ns (0,29)
	N1	7 (19%)	7 (8%)	
	N2	5 (14%)	11 (13%)	
	N3	0%	1 (1%)	
PATTERN prédominant n(%)	Lépidique	6 (17%)	17 (19%)	ns (0,76)
	Acinaire	2 (5%)	12 (14%)	
	Papillaire	2 (5%)	7 (8%)	
	Micropapillaire	5 (14%)	10 (12%)	
	Cribriforme	1 (3%)	4 (5%)	
	Solide	15 (42%)	26 (30%)	
	Glandulaire complexe	5 (14%)	10 (12%)	
GRADE n(%)	1	6 (17%)	17 (19%)	ns (0,13)
	2	2 (5%)	16 (18%)	
	3	28 (78%)	53 (63%)	

Tableau 1 : **Caractéristiques anatomo-cliniques des patients**

Le **Tableau 1** résume les principales caractéristiques cliniques et anatomopathologiques des patients inclus, et montre que les cohortes d'entrainement et de validation sont comparables sur l'ensemble de ces caractéristiques. A noter que les tumeurs de grade 3 sont largement majoritaires dans les deux cohortes, suivies par les grades 1 et 2. Ce profil suggère une prévalence des formes histologiquement agressives dans la population étudiée, cohérente avec l'objectif de modélisation prédictive de patterns tumoraux à visée pronostique.

2. Concordance inter-observateurs

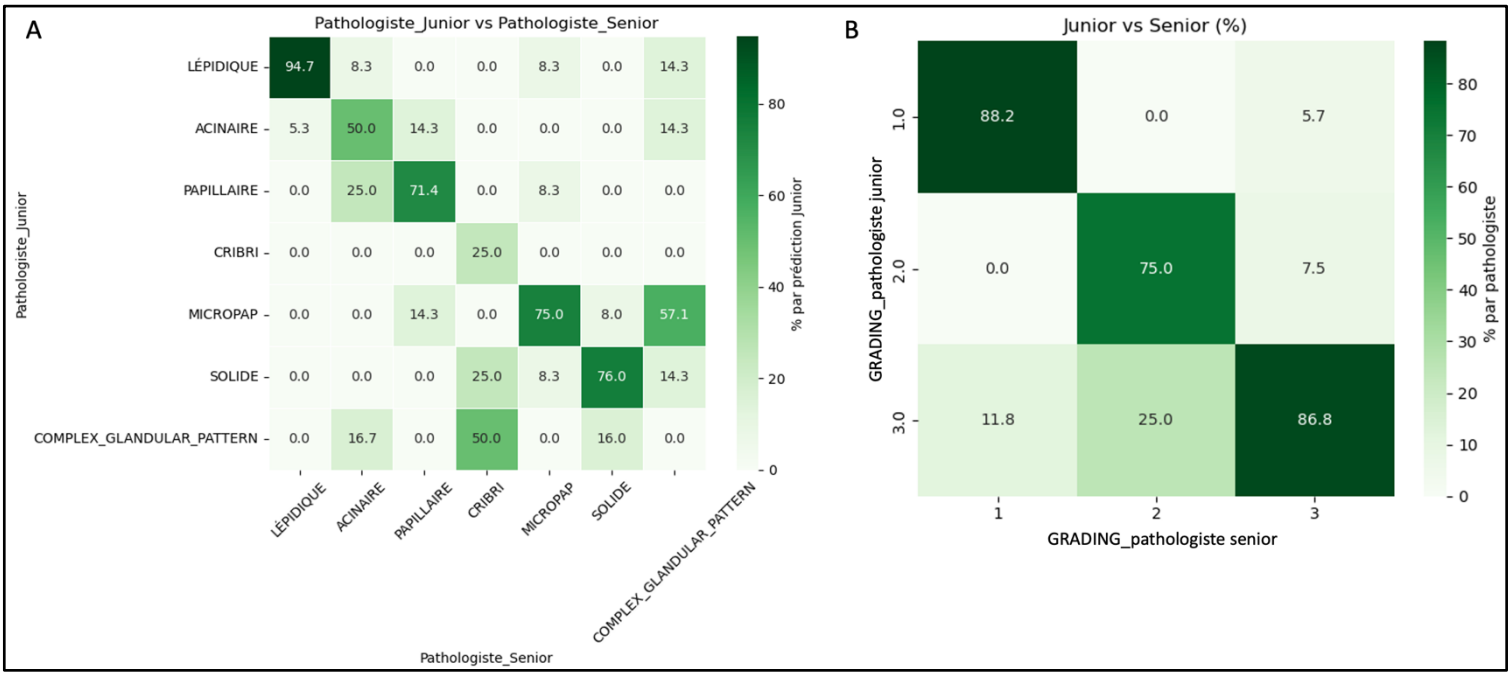


Figure 16 : **Concordance inter-observateurs (pathologiste senior versus pathologiste junior)** sur le pattern prédominant (A) et sur le grading (B).

Une analyse de concordance entre le pathologiste junior et le pathologiste senior sur la classification des patterns histologiques et sur l'attribution du grade a été réalisée **(figure 16)**. La matrice de covariance présentée en **figure A** met en évidence un taux de concordance globalement élevée sur plusieurs patterns tumoraux. On observe par exemple une concordance quasi-parfaite sur le pattern lépidique, avec 94,7 % des cas annotés lépidiques par le junior également reconnus comme tels par le senior. Des concordances élevées sont également observées pour les patterns papillaire (71,4 %) et micropapillaire (75,0 %). En revanche, des divergences notables apparaissent pour des patterns plus complexes ou mixtes, notamment le « complex glandular pattern », qui est identifié de manière hétérogène (avec 50 % des cas classés papillaire par le senior et 16,7 % acinaire).

La **figure B** présente la comparaison des grades attribués par les deux pathologistes. Un taux de concordance très élevée est observé pour les grades 1 (88%) et 3 (87%). Le grade 2 présente quant à lui une concordance plus modérée (75 %), avec certains cas classés différemment par l'un ou l'autre évaluateur, notamment entre les grades 2 et 3. La valeur de l'indice de Kappa pondéré ($\kappa = 0,75$) indique un très bon niveau de cond-concordance inter-observateur pour le grading global, ce qui confirme la robustesse des critères utilisés pour l'évaluation morphologique. Cette concordance est également soutenue par une valeur de Kappa simple de 0,73 et un test du Chi² significatif ($\chi^2 = 96,12$, $p < 10^{-19}$).

Ces résultats soulignent une reproductibilité satisfaisante de l'évaluation histologique entre l'observateur junior et senior, en particulier sur les patterns les plus fréquents et sur les grades extrêmes, tout en mettant en évidence certaines limites dans l'interprétation des patterns mixtes ou moins bien définis.

3. Performance du modèle pour la classification des patterns et le grading

L'interprétation des performances du modèle a été réalisée en prenant le pathologiste senior comme Gold standard pour toutes les analyses ci-dessous.

3.1 Performances du modèle « brut »

Dans un premier temps, les performances du modèle « brut » (non lissé, non pondéré) pour la classification des patterns et le grading ont été évaluées, en se basant sur les données du Pathologiste senior **(figure 17)**.

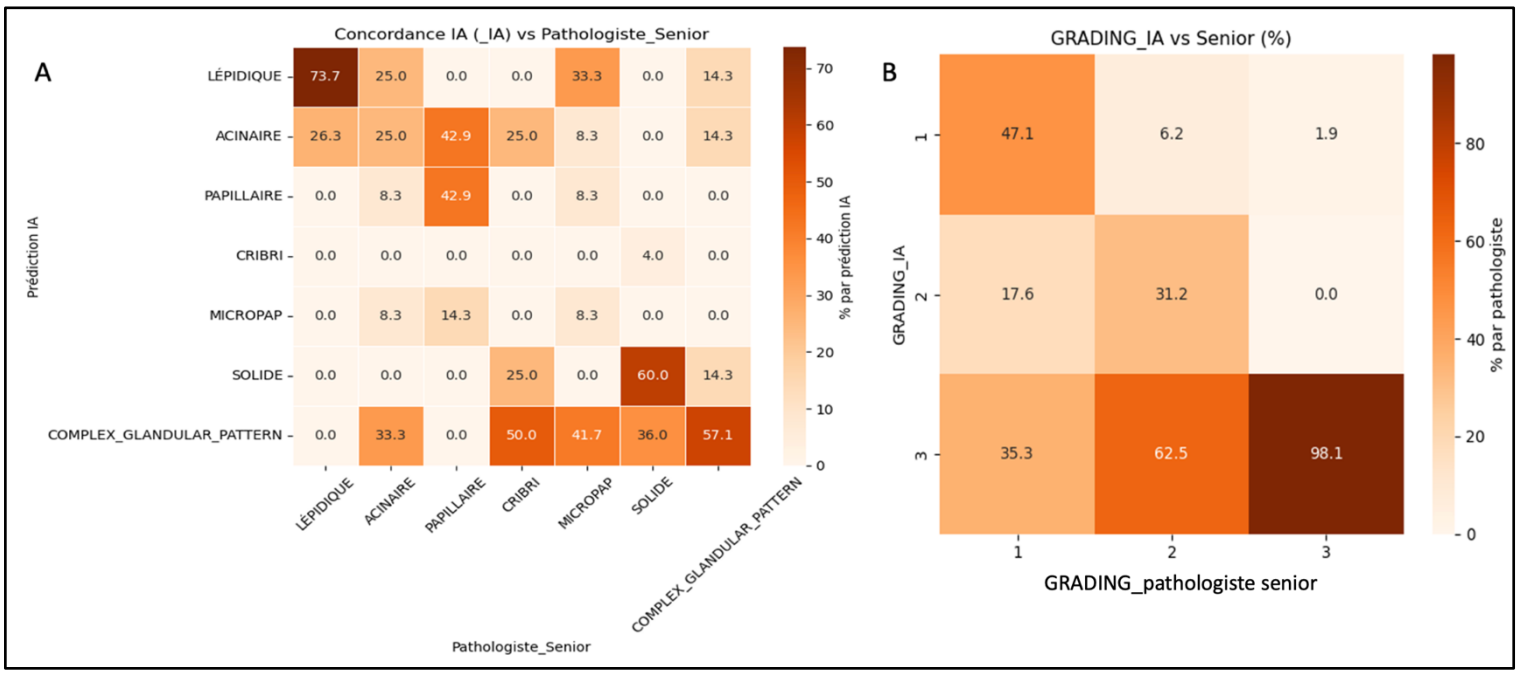


Figure 17 : **Performances du modèle « brut » pour la classification des patterns (A) et le grading (B) des adénocarcinomes pulmonaires.**

La matrice de covariance **(fig A)** montre une concordance partielle entre les patterns histologiques prédits par l'IA et ceux du pathologiste senior. Si les patterns comme lépidique (73,7 %) ou solide (60 %) sont relativement concordants, d'autres comme le cribriforme, le papillaire, ou le micropapillaire sont largement sujets à la confusion, traduisant une difficulté du modèle brut à bien distinguer ces patterns.
Concernant le grading tumoral **(fig B)**, la concordance entre le modèle brut **le** pathologiste senior est faible, avec une valeur de Kappa pondéré de 0,58. Ceci s'explique par une faible concordance sur les grades 1 et 2. En effet, 35,3 % des cas classés grade 1 par le Pathologiste senior sont classés grade 3 par le modèle, et 62,5% des cas classés grades 2 par le pathologiste senior sont classés grade 3 par le modèle. Cela traduit une très probable surestimation du grade par le modèle brut, tiré par une surestimation des patterns de haut grade.

3.2 Performances du modèle « lisse »

Dans un second temps, une technique de lissage a été appliquée au modèle, afin d'améliorer la classification (cf méthodologie). Les performances de ce nouveau modèle « lissé » sont présentées dans la **figure 18**.

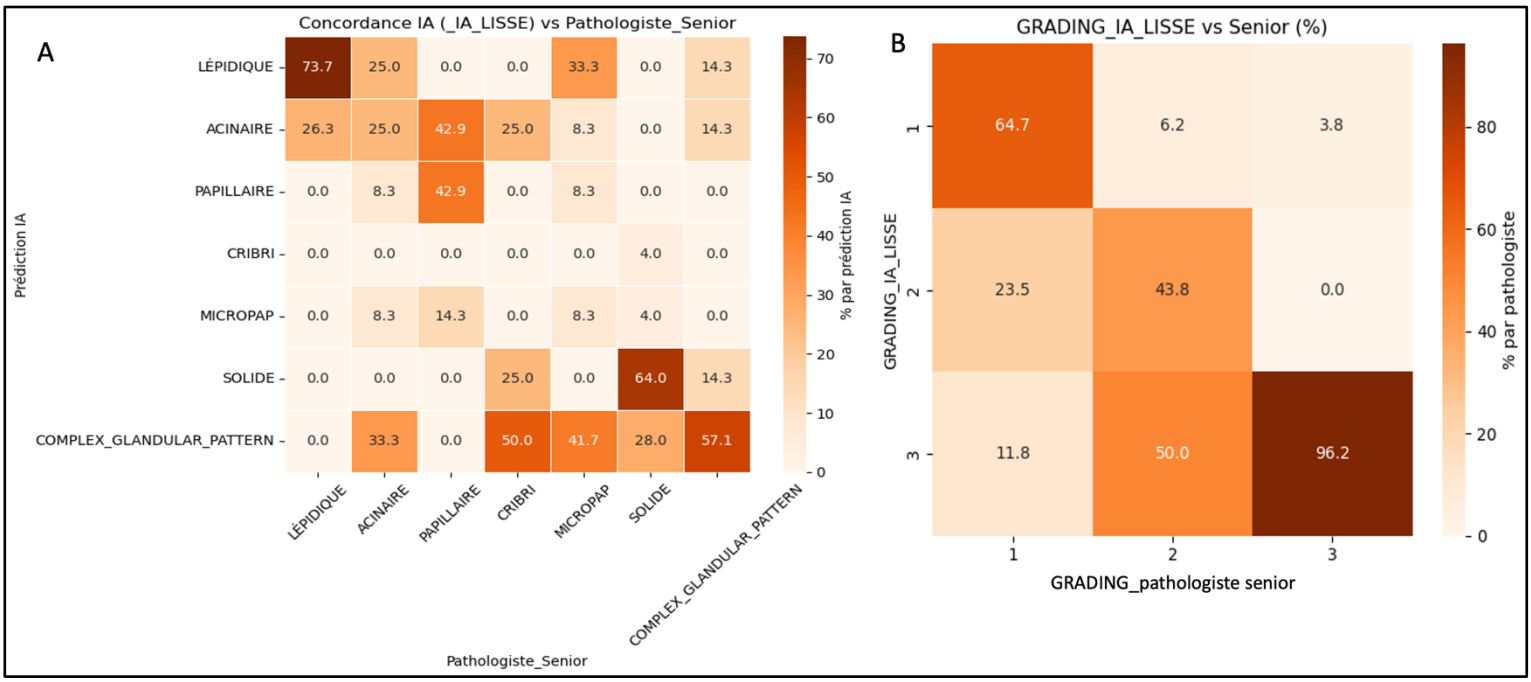


Figure 18 : **Performances du modèle « lissé » pour la classification des patterns (A) et le grading (B) des adénocarcinomes pulmonaires.**

Concernant la classification des patterns **(fig A)**, le lissage améliore légèrement la stabilité des prédictions IA en réduisant les petits groupes de patchs aberrants. On observe seulement une meilleure homogénéité dans la classification du pattern solide, mais les confusions persistent dans les autres patterns (cribriforme, papillaire, miropapillaire...).

Concernant la classification du grade **(fig B)**, le lissage améliore nettement la concordance avec le pathologiste senior, avec un Kappa pondéré de 0,73. On note une réduction marquée des erreurs de classification sur le grade 1, qui est mieux distingué du grade 3 (11,8%).

3.3 Performances du modèle après pondération

Dans un troisième temps, nous avons appliqué une méthode de pondération (cf méthodologie) afin de diminuer le poids des patterns de haut grade, surestimés par le modèle **(figure 19)**.

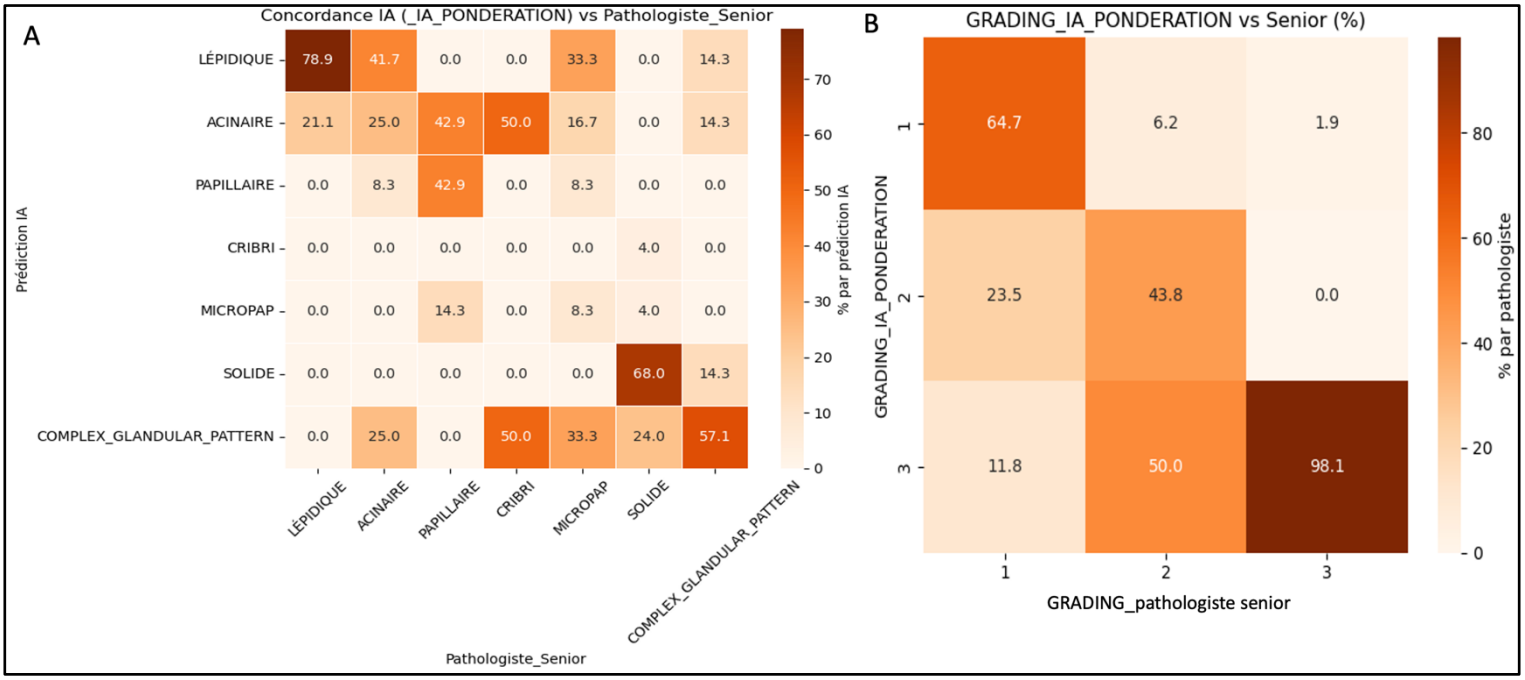


Figure 19 : **Performances du modèle « pondéré » pour la classification des patterns (A) et le grading (B) des adénocarcinomes pulmonaires.**

Concernant la classification des patterns **(fig A)**, la pondération des valeurs de sortie IA améliore aussi la prédiction des patterns, notamment en accentuant les patterns majoritaires dans les zones tumorales. Cela permet par exemple un gain de précision sur le pattern solide (68%) et lépidique (78,9%). Toutefois, des erreurs ponctuelles persistent sur les autres patterns.

Concernant la classification du grade **(fig B)**, la pondération améliore légèrement la concordance avec le Pathologiste senior, avec un Kappa pondéré de 0,76. Les principales erreurs de surestimation (grade 1 classé à tort en grade 3) persistent (11,8 %), mais on observe une amélioration de la concordance pour le grade 3. La pondération agit donc ici comme un rééquilibrage pour la reconnaissance du grade 3.

3.4 Performances du modèle après lissage et pondération

Dans un dernier temps, nous avons combiné les méthodes de lissage et de pondération afin d'évaluer leur impact sur la performance du modèle **(figure 20)**.

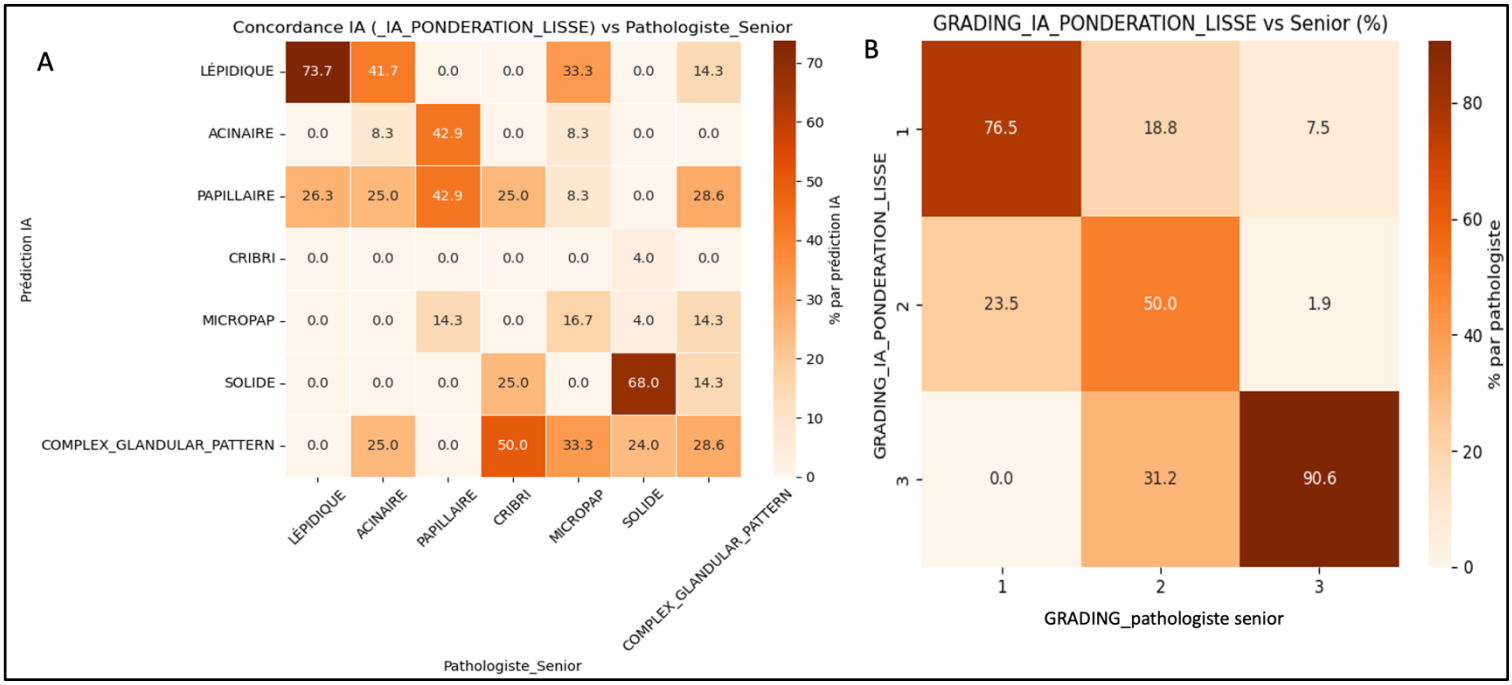


Figure 20 : **Performances du modèle « lissé » et « pondéré » pour la classification des patterns (A) et le grading (B) des adénocarcinomes pulmonaires.**

Concernant la classification des patterns **(fig A)**, l'association des deux techniques (lissage + pondération) produit une classification plus harmonisée des patterns histologiques. Bien que les confusions résiduelles persistent légèrement (notamment entre micropapillaire et complex glandular), on note une réduction des classifications erratiques et une meilleure reconnaissance en valeur des patterns dominants.

Concernant la classification du grade **(fig B)**, cette combinaison offre un compromis optimal : le Kappa pondéré atteint 0,75, et les confusions critiques (ex. grade 1/grade 3) sont pratiquement éliminées. L'accord est excellent pour le grade 3 (90,6 %), tandis que le grade 2 reste partiellement confondu avec les autres. Cette configuration constitue donc le meilleur équilibre entre fiabilité morphologique et stabilité des prédictions du modèle.

3.5 Comparaison appariée des proportions par pattern sur la cohorte test (validation)

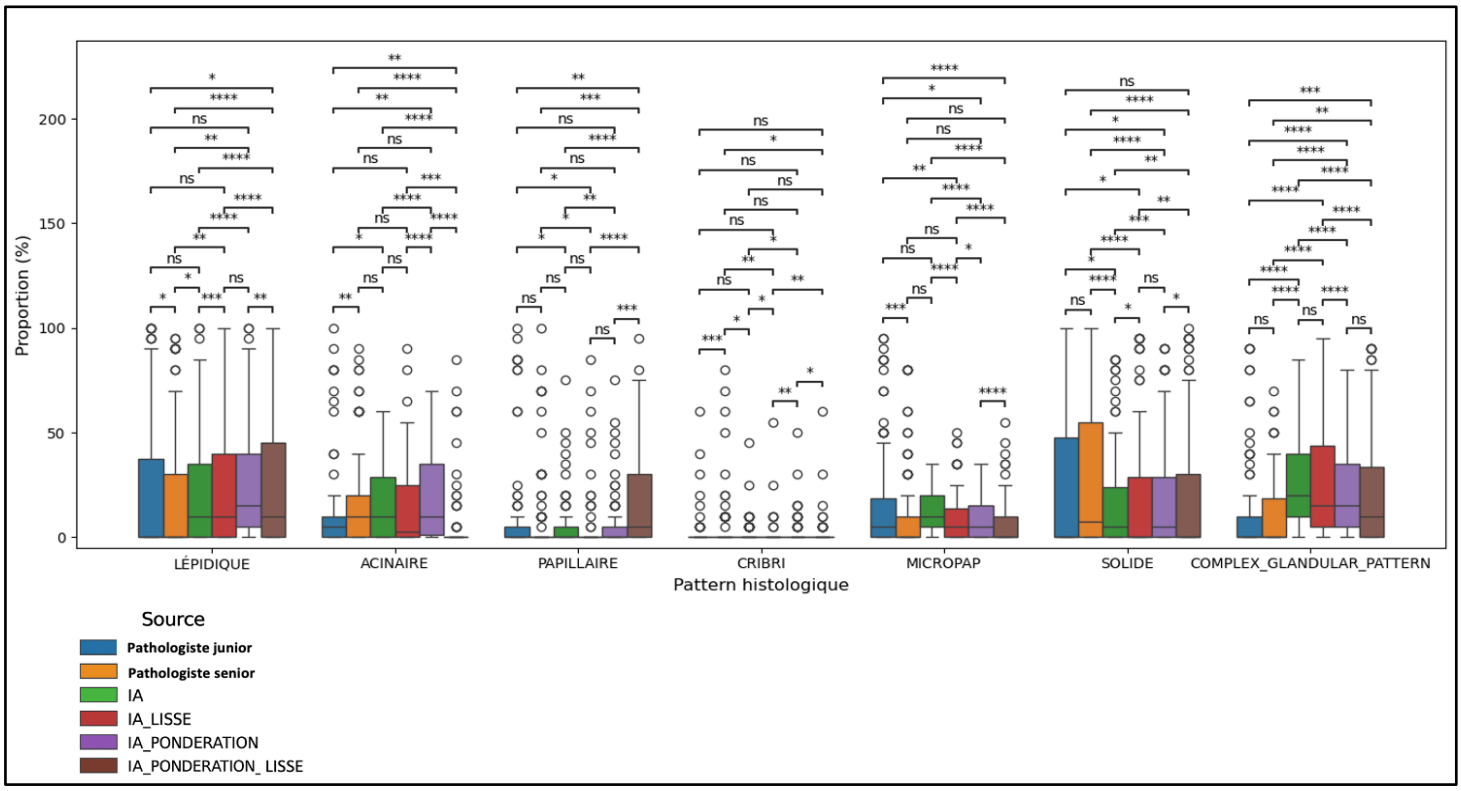


Figure 21 : **Comparaison appariée des proportions par pattern sur la cohorte test (validation).** *Les niveaux de significativité statistique sont indiqués par des astérisques : p < 0,05 (*), p < 0,01 (**), p < 0,001 (***), et p < 0,0001 (****), selon l'intensité de la différence observée. Plus le nombre d'astérisques est élevé, plus la différence est jugée statistiquement robuste. À l'inverse, la mention "ns" signifie "non significatif", indiquant qu'il n'y a pas de différence statistique notable entre les groupes comparés (p > 0,05).*

La **figure 21** ci-dessus présente une comparaison appariée des proportions de chaque pattern histologique estimées par différents observateurs et méthodes, incluant les annotations du pathologiste junior, du pathologiste senior, ainsi que celles issues des différents modèles d'IA (lissage, pondération, ou combinaison des deux). L'objectif était d'évaluer la proximité des répartitions morphologiques générées par le modèle par rapport à l'expertise humaine.

On observe que pour plusieurs patterns, les proportions estimées par le modèle brut (en vert) diffèrent significativement de celle du pathologiste senior (gold standard), notamment pour les patterns solide et complex glandular, avec des p-values souvent inférieures à 0,001 (*** ou ****), traduisant des écarts marqués. En revanche, l'application du lissage (rouge), de la pondération (violet), ou de leur combinaison (marron) permet de réduire ces écarts.

L'application du lissage seul (rouge) permet dans plusieurs cas de réduire les valeurs extrêmes et d'homogénéiser les prédictions : par exemple, les différences diminuent pour les patterns acinaire et solide. De son côté, la pondération seule (violet) agit différemment, en réajustant la pondération globale des patterns en fonction de leur fréquence, ce qui améliore notamment la concordance pour les patterns lépidique et « complex glandular ».

La combinaison du lissage et de la pondération (marron) donne dans plusieurs cas la répartition la plus proche des annotations du pathologiste senior, avec une diminution significative des écarts.

Enfin, la dispersion des valeurs (outliers) reste plus importante pour certains patterns complexes, ce qui reflète leur difficulté de segmentation, mais aussi leur variabilité inter-observateur naturelle.

Ces résultats confirment que lissage et la pondération du modèle IA contribuent à **rapprocher la distribution des patterns prédits de celles observées par des pathologistes expérimentés**, améliorant ainsi la fiabilité du modèle dans l'estimation des proportions morphologiques au sein des adénocarcinomes pulmonaires.

3.6 Comparaison appariée des proportions par pattern sur la cohorte d'entrainement

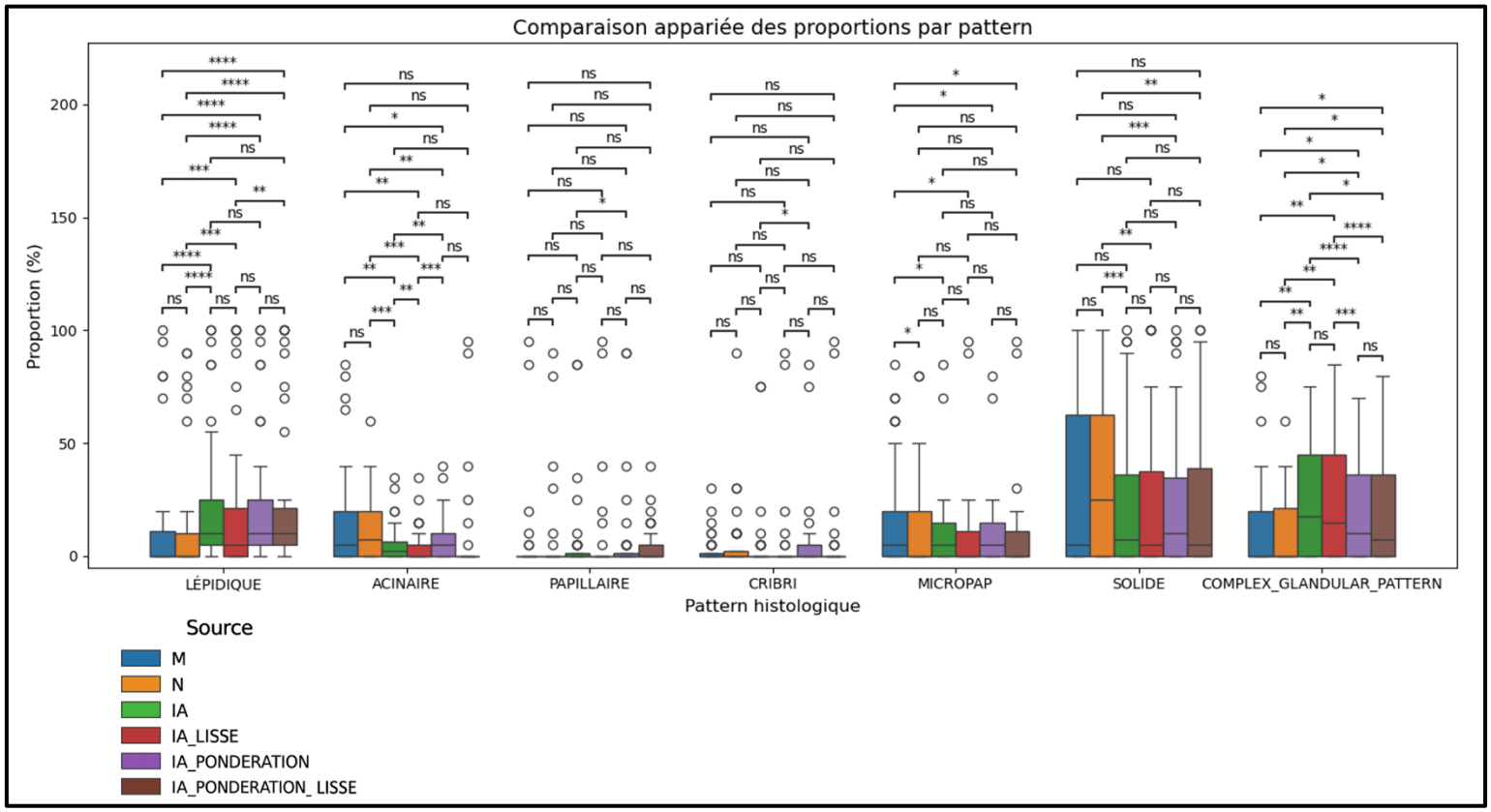


Figure 22 : **Comparaison appariée des proportions par pattern sur la cohorte d'entrainement.** *Les niveaux de significativité statistique sont indiqués par des astérisques : p < 0,05 (*), p < 0,01 (**), p < 0,001 (***), et p < 0,0001 (****), selon l'intensité de la différence observée. Plus le nombre d'astérisques est élevé, plus la différence est jugée statistiquement robuste. À l'inverse, la mention "ns" signifie "non significatif", indiquant qu'il n'y a pas de différence statistique notable entre les groupes comparés (p > 0,05).*

La **figure 22** met en évidence des écarts significatifs entre l'IA brute et le pathologiste senior**,** notamment pour les patterns fréquents tels que le lépidique, l'acinaire et le solide**,** confirmant la difficulté de leur classification automatique. À l'inverse, les patterns plus rares (papillaire, cribriforme) montrent une bonne concordance, sans

différence notable. L'application du lissage et de la pondération améliorent la concordance globale, en particulier pour les patterns micropapillaire et complexe glandular, réduisant les écarts et rendant certaines comparaisons non significatives. Ces résultats suggèrent que, malgré des difficultés persistantes pour certains patterns morphologiquement hétérogènes, **les ajustements apportés au modèle permettent de renforcer sa fiabilité et sa proximité avec l'évaluation experte**.

4. Exemples de classification (heatmap) des patterns avec le modèle « H-Optimus-0 »

Le principe de la heat-map consiste à subdiviser la région tumorale en un ensemble de tuiles. À chacune d'elles, le modèle attribue un score probabiliste, exprimé en pourcentage (0 à 100 %), ainsi qu'un code couleur gradué du noir au jaune. Ce score reflète la probabilité que la tuile analysée corresponde au pattern histologique recherché (fig. 23-29).

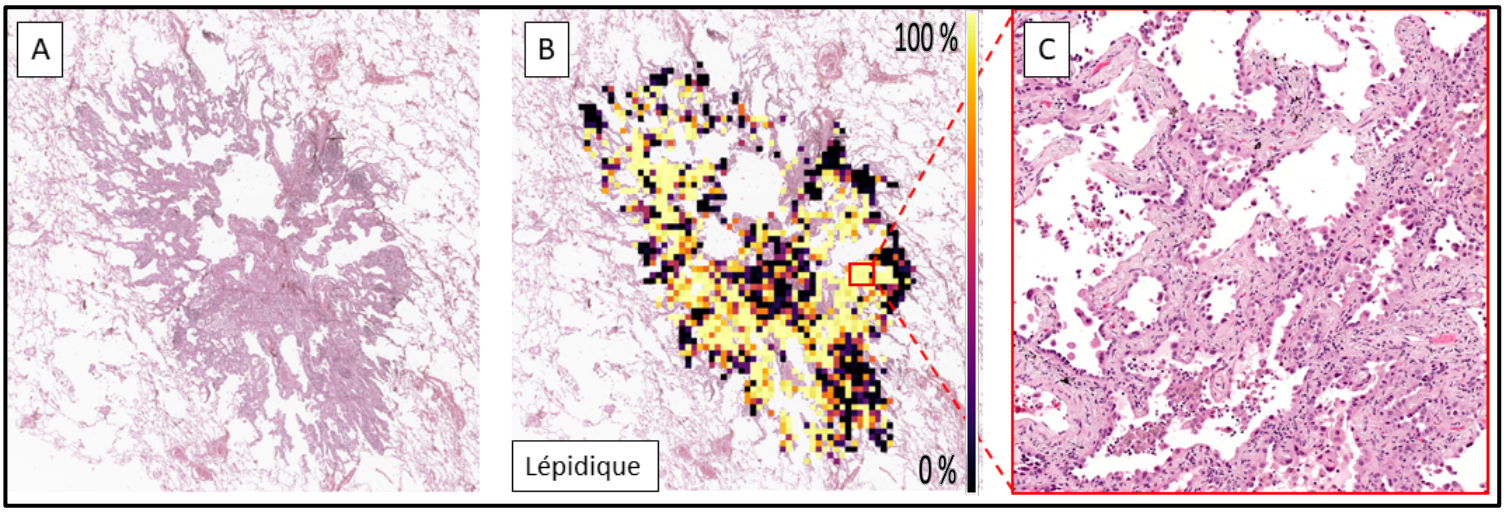


Figure 23 : **Heat-map d'un adénocarcinome comportant un pattern lépidique. A.** Lame HES montrant un adénocarcinome d'architecture majoritairement lépidique (x2) **B.** heat-map montrant la répartition du pattern lépidique **C.** Pattern lépidique à fort grossissement (x10).

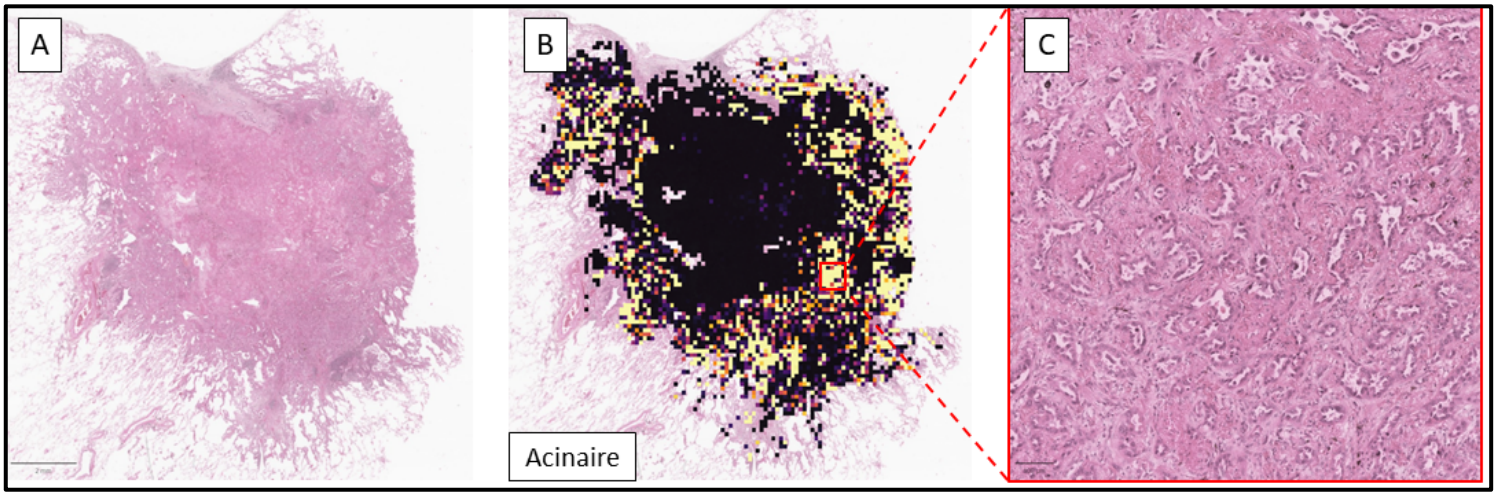


Figure 24 : **Heat-map d'un adénocarcinome comportant un pattern acinaire. A.** Lame HES montrant un adénocarcinome comportant un contingent de pattern acinaire (en périphérie, x2). **B.** heat-map montrant la répartition du pattern acinaire. **C.** Pattern acinaire à fort grossissement (x10).

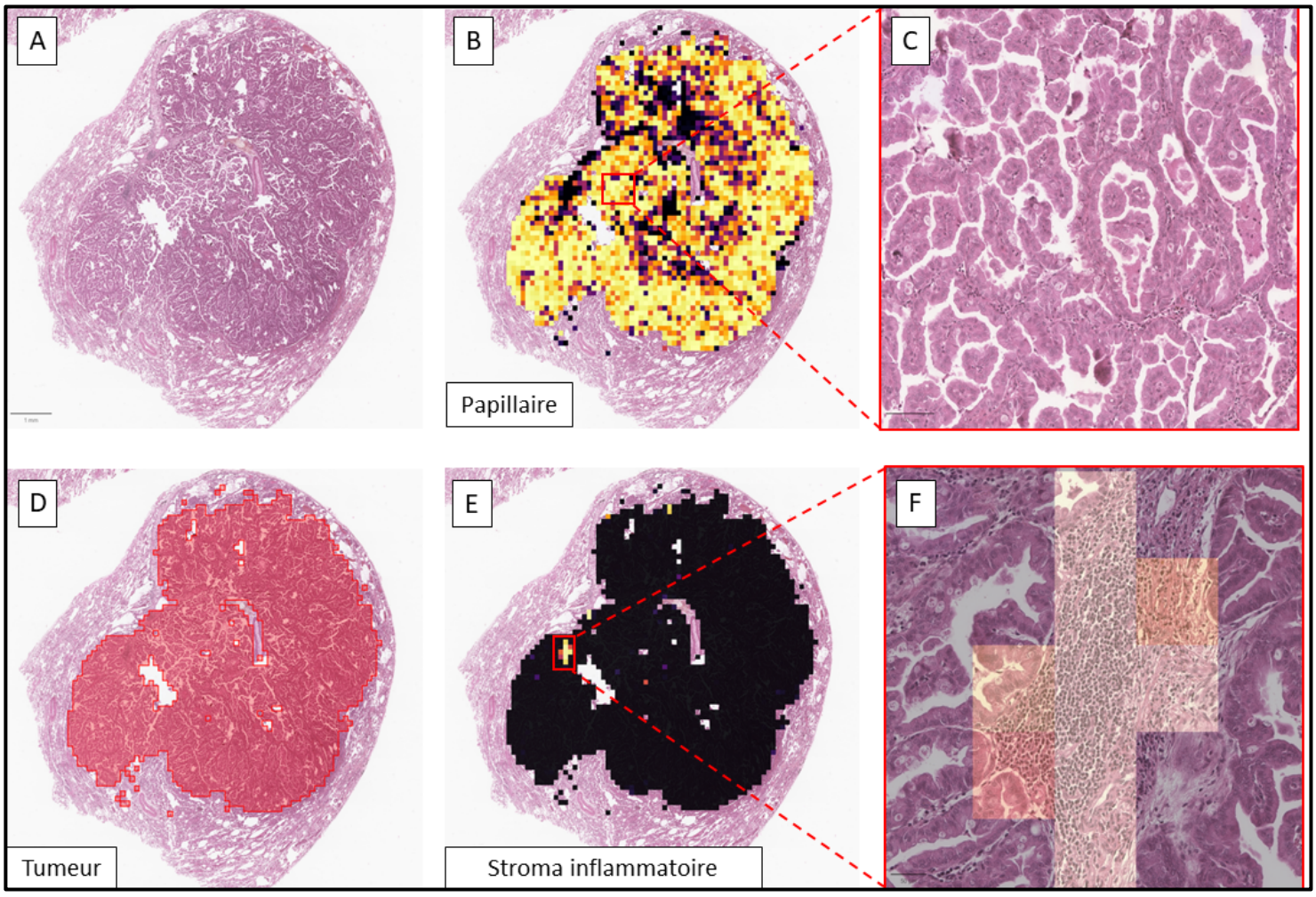


Figure 25 : **Heat-map d'un adénocarcinome comportant un pattern papillaire et un foyer de stroma inflammatoire. A.** Lame HES montrant un adénocarcinome

majoritairement d'architecture papillaire. **B.** heat-map montrant la répartition du pattern papillaire. **C.** Pattern papillaire à fort grossissement (x10). **D.** Délimitation de la tumeur par le modèle. **E-F** Heat-map montrant un foyer de stroma inflammatoire au sein de la tumeur.

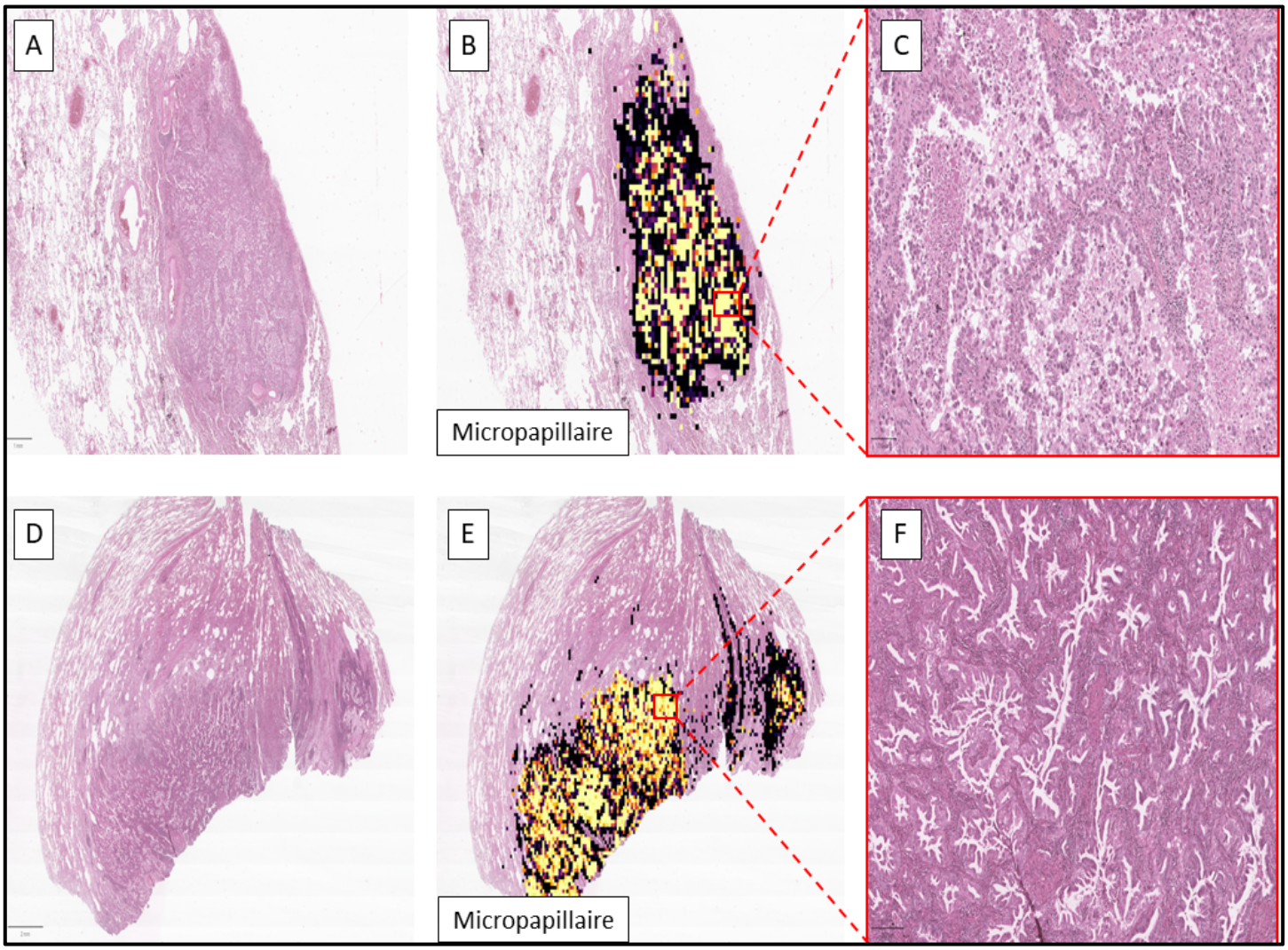


Figure 26 : **Heat-map d'un adénocarcinome comportant un pattern micropapillaire. A.** Lame HES du nodule tumoral n°1. **B.** heat-map montrant la répartition du pattern micropapillaire. **C.** Pattern micropapillaire à fort grossissement. **D.** Lame HES du nodule tumoral n°2. **E.** heat-map du pattern micropapillaire. **F.** Pattern micropapillaire (variante « filigree pattern ») à fort grossissement (x10).

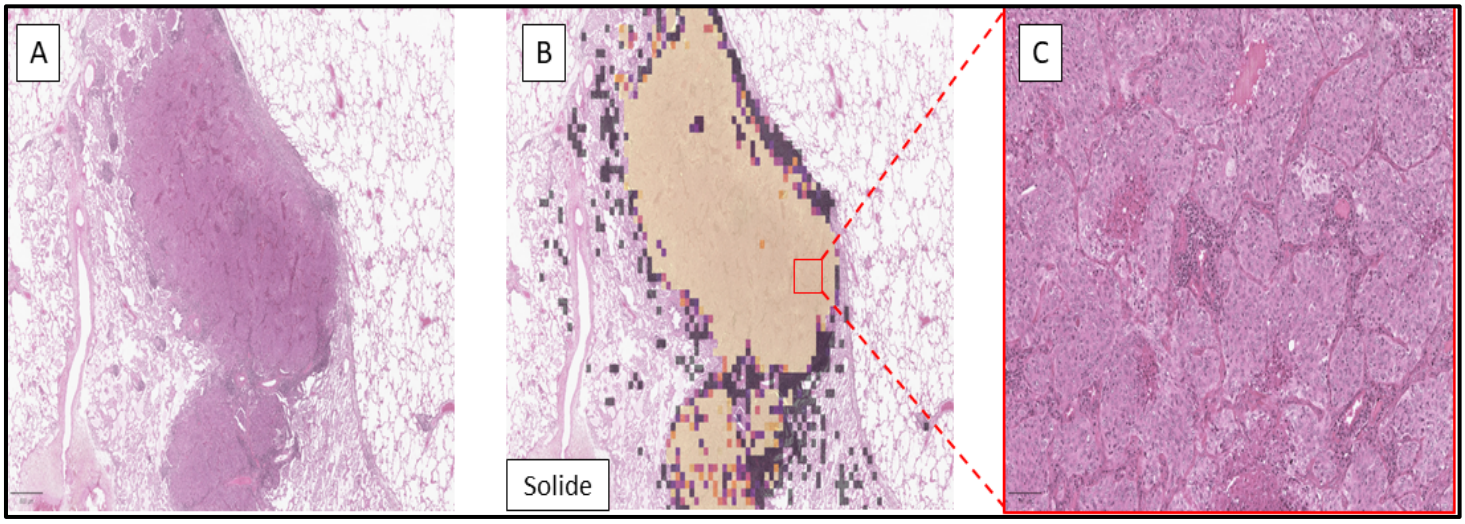


Figure 27 : **Heat-map d'un adénocarcinome comportant un pattern solide. A.** Lame HES montrant un adénocarcinome majoritairement solide. **B.** Heat-map du pattern solide. **C.** Pattern solide à fort grossissement (x10).

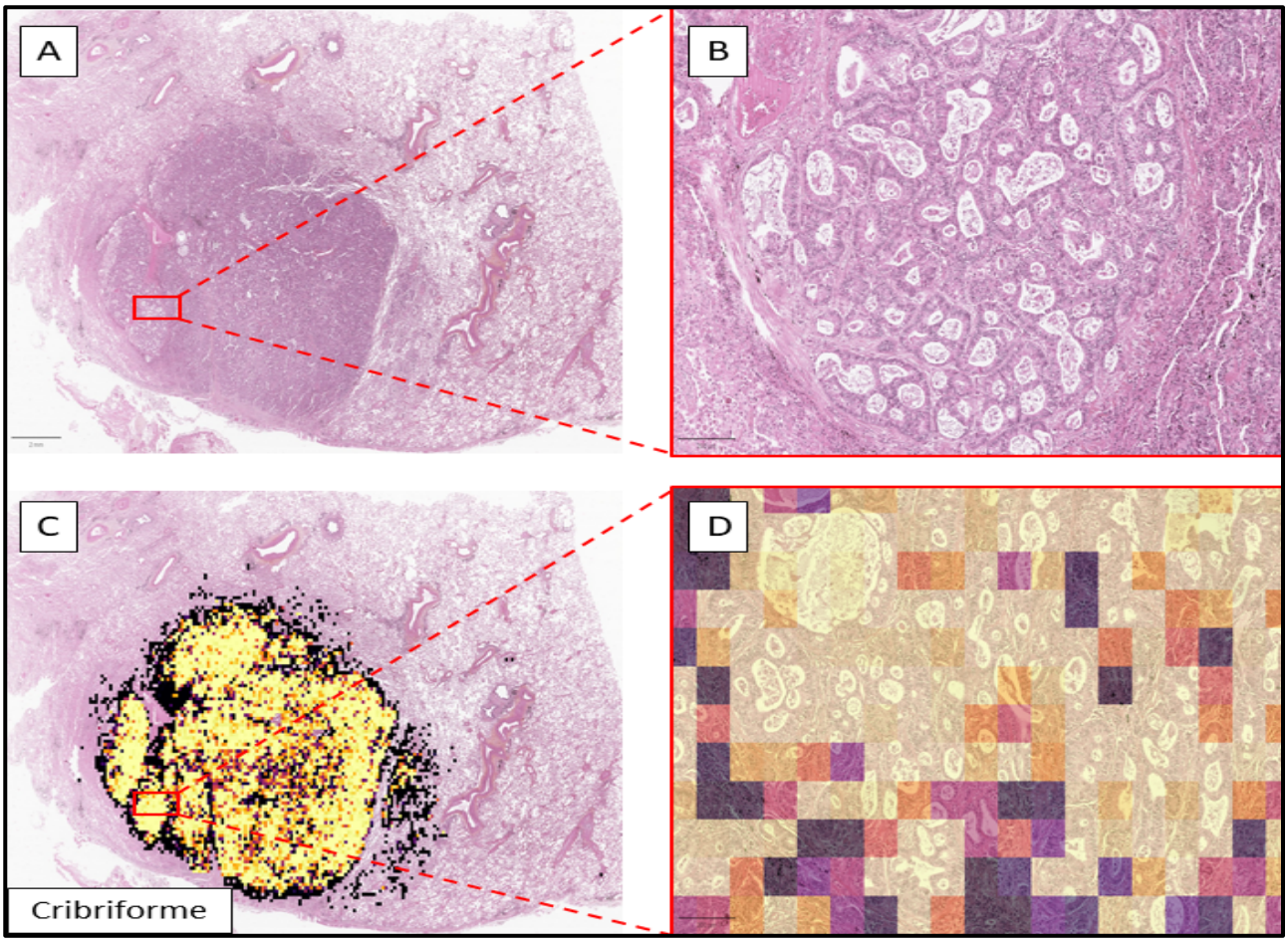


Figure 28 : **Heat-map d'un adénocarcinome comportant un pattern cribriforme. A.** Lame HES du nodule tumoral. **B.** Pattern cribriforme à fort grossissement (x10). **C.** heat-map du pattern cribriforme. **D.** heat-map du pattern cribriforme à fort grossissement (x10).

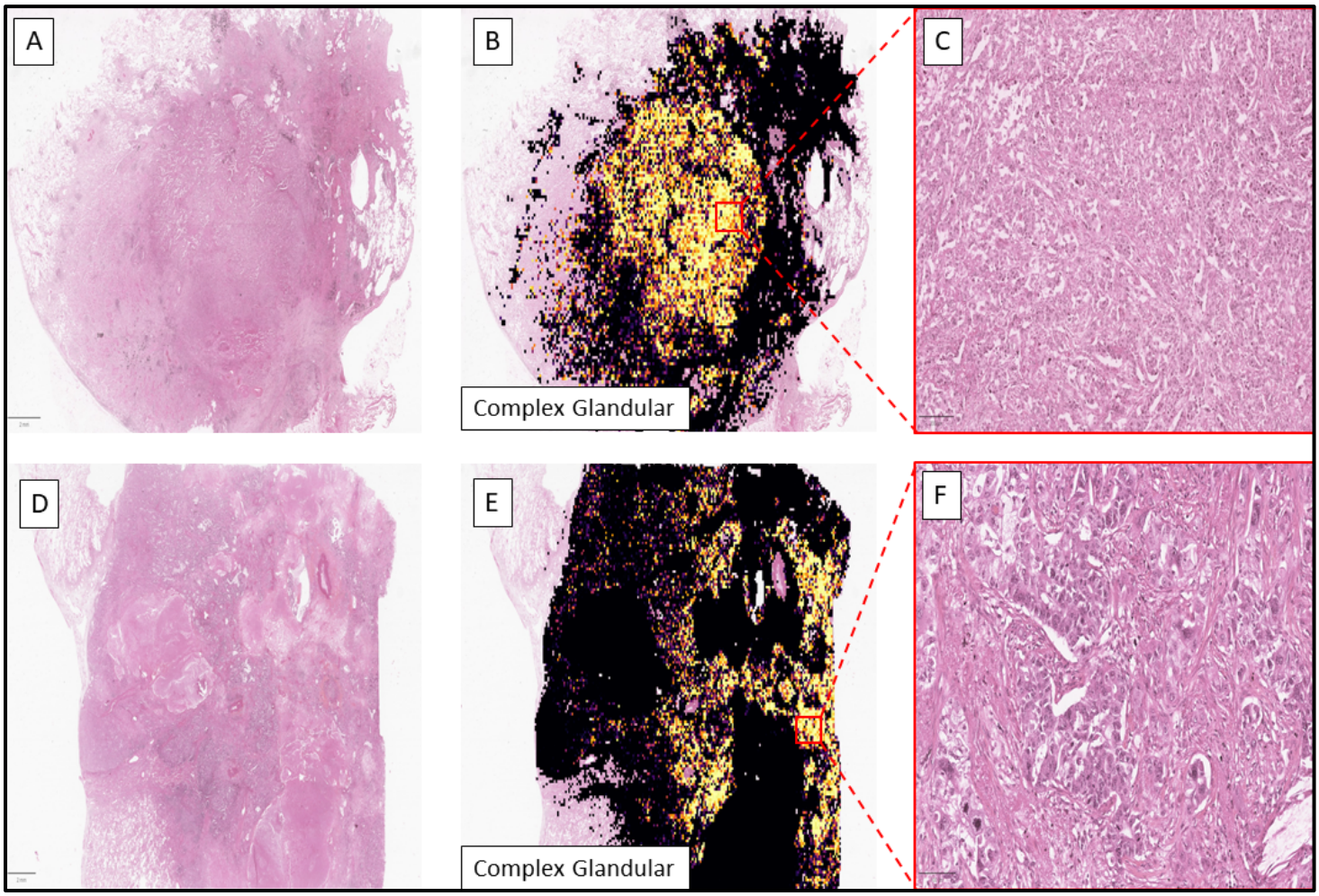


Figure 29 : **Heat-map d'un adénocarcinome comportant un pattern « complex glandular ». A.** Lame HES du nodule tumoral n°1. **B.** heat-map du « complex glandular pattern ». **C.** « complex glandular pattern » à fort grossissement (x10). **D.** Lame HES du nodule tumoral n°2. **E.** heat-map du « complex glandular pattern ». **F.** « complex glandular pattern » à fort grossissement (x10).

5. Exemples de résultats discordants entre le modèle et le pathologiste sénior

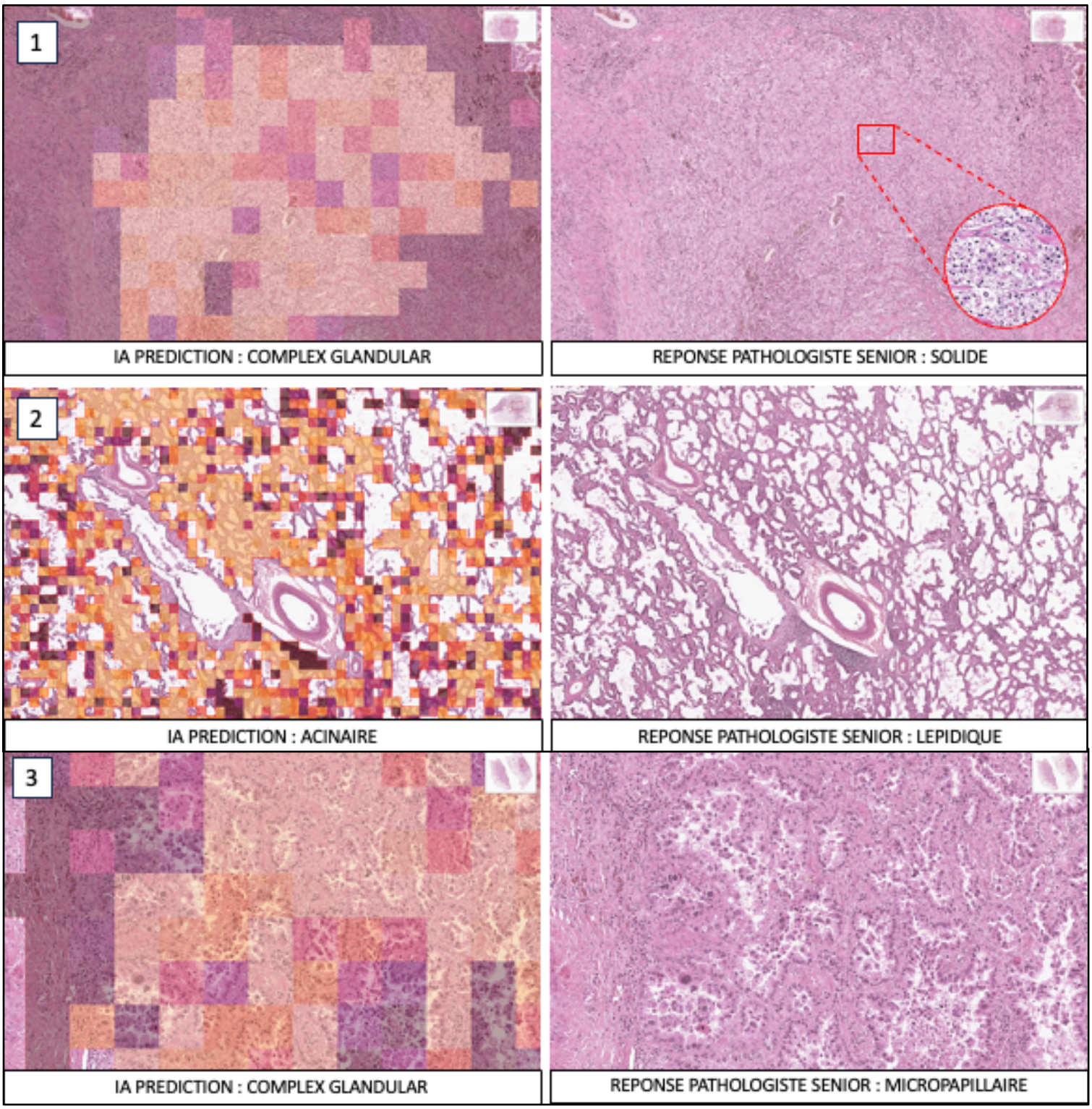


Figure 30 : **Exemples de discordances sur les patterns. 1.** discordance entre pattern glandulaire complexe et pattern solide. **2.** discordance entre pattern acinaire et pattern lépidique. **3.** Discordance entre pattern glandulaire complexe et foyers de pattern micropapillaire.

L'analyse des discordances sur la classification des patterns entre le modèle et le pathologiste sénior a permis d'identifier trois situations dans lesquelles le modèle avait tendance à se tromper. **Premièrement** (situation n°1 **Fig.30**), dans certains adénocarcinomes pulmonaires présentant un pattern solide à différentiation « cellules claires », le modèle avait tendance à les classer en « complex glandular pattern ». **Deuxièmement** (situation n°2 **Fig.30**), lorsque le pattern lépidique présentait un aspect légèrement collabé, soit du fait d'une architecture rigide du nodule tumoral, soit en lien avec un artefact mineur de coupe, le modèle le classait préférentiellement en pattern acinaire. **Troisièmement** (situation n°3 **Fig.30**), dans certaines zones de morphologie complexe où on retrouvait principalement un pattern micropapillaire, ce dernier n'était pas détecté du fait de l'importance des remaniements environnants (fibrose, inflammation, nécrose). Néanmoins, dans l'ensemble de ces situations, le grading final de la tumeur n'était pas modifié.

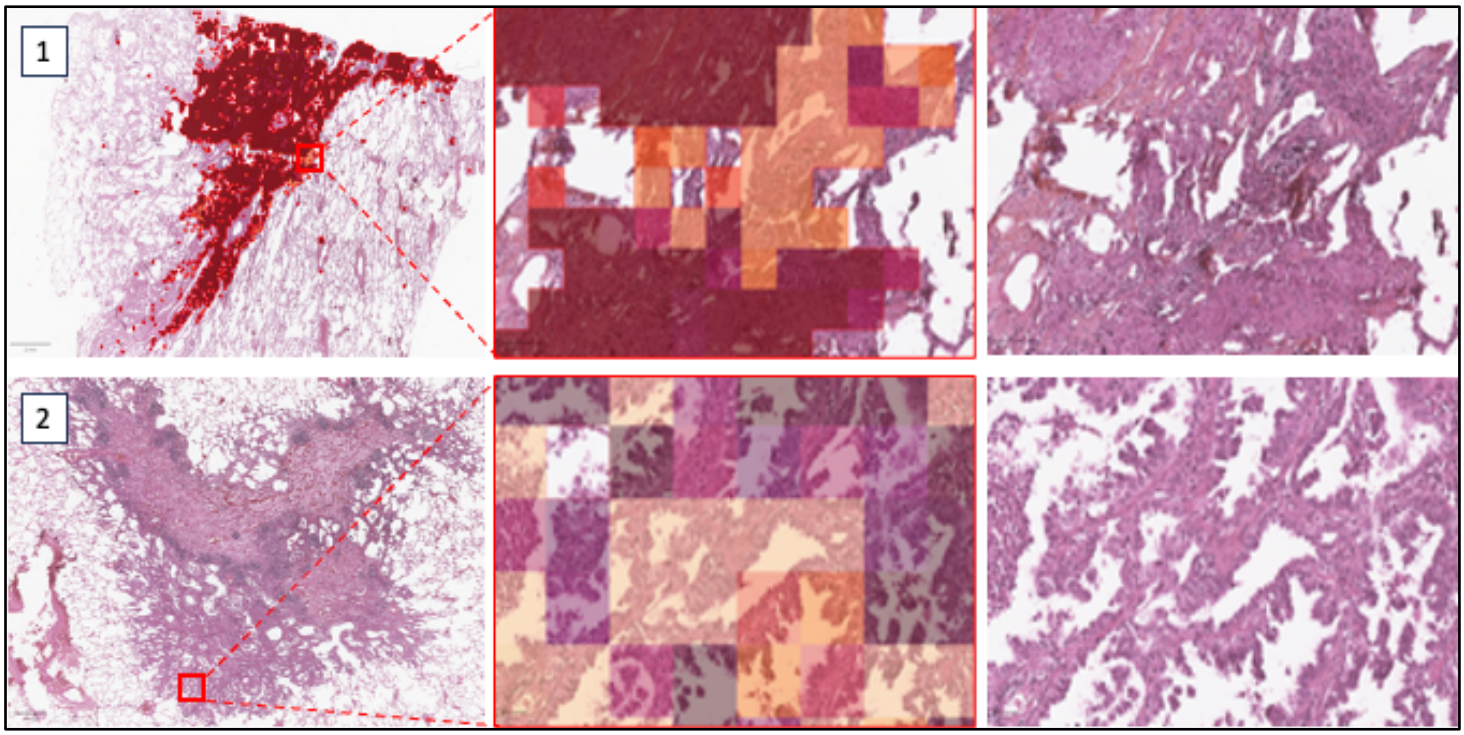


Figure 31 : **Exemples de discordance sur les grades (grade 3 versus grade 1).**
1. Cas discordant (1) : modèle de fondation : pattern lépidique (grade 1) ; Pathologiste : secteurs remaniés par la fibrose (grade 3).
2. Cas discordant (2) : modèle de fondation : pattern lépidique (grade 1) ; Pathologiste : pattern micropapillaire (grade 3).

Après comparaison des résultats de grading entre le modèle et le pathologiste senior, seuls quatre cas présentaient des discordances de deux niveaux de grade (grade 1 versus grade 3) : dans tous ces cas, le pathologiste senior avait attribué un grade 3, tandis que l'IA proposait un grade 1. Aucun cas inverse (grade 1 attribué par le pathologiste senior et grade 3 par l'IA) n'a été recensé. Ces quatre cas concernaient des nodules tumoraux très remaniés, avec une proportion importante de stroma fibreux et inflammatoire, les zones tumorales cellulaires ne représentant qu'une faible part de la surface totale. Dans la **première situation (Figure 31)**, regroupant deux de ces cas, le modèle s'est focalisé sur des secteurs où l'architecture alvéolaire était fortement altérée par des remaniements fibro-inflammatoires, identifiant à tort un

pattern lépidique alors qu'existaient, dans d'autres régions, des foyers de pattern micropapillaire non détectés. Dans la **seconde situation (Figure 31)**, correspondant aux deux autres cas, l'aspect du foyer tumoral à faible grossissement pouvait évoquer un pattern lépidique, interprété comme tel par le modèle, mais l'examen à fort grossissement révélait des modifications cytonucléaires en faveur d'un pattern micropapillaire.

6. Exemples de classification (segmentation) des patterns avec « H-Optimus-0 »

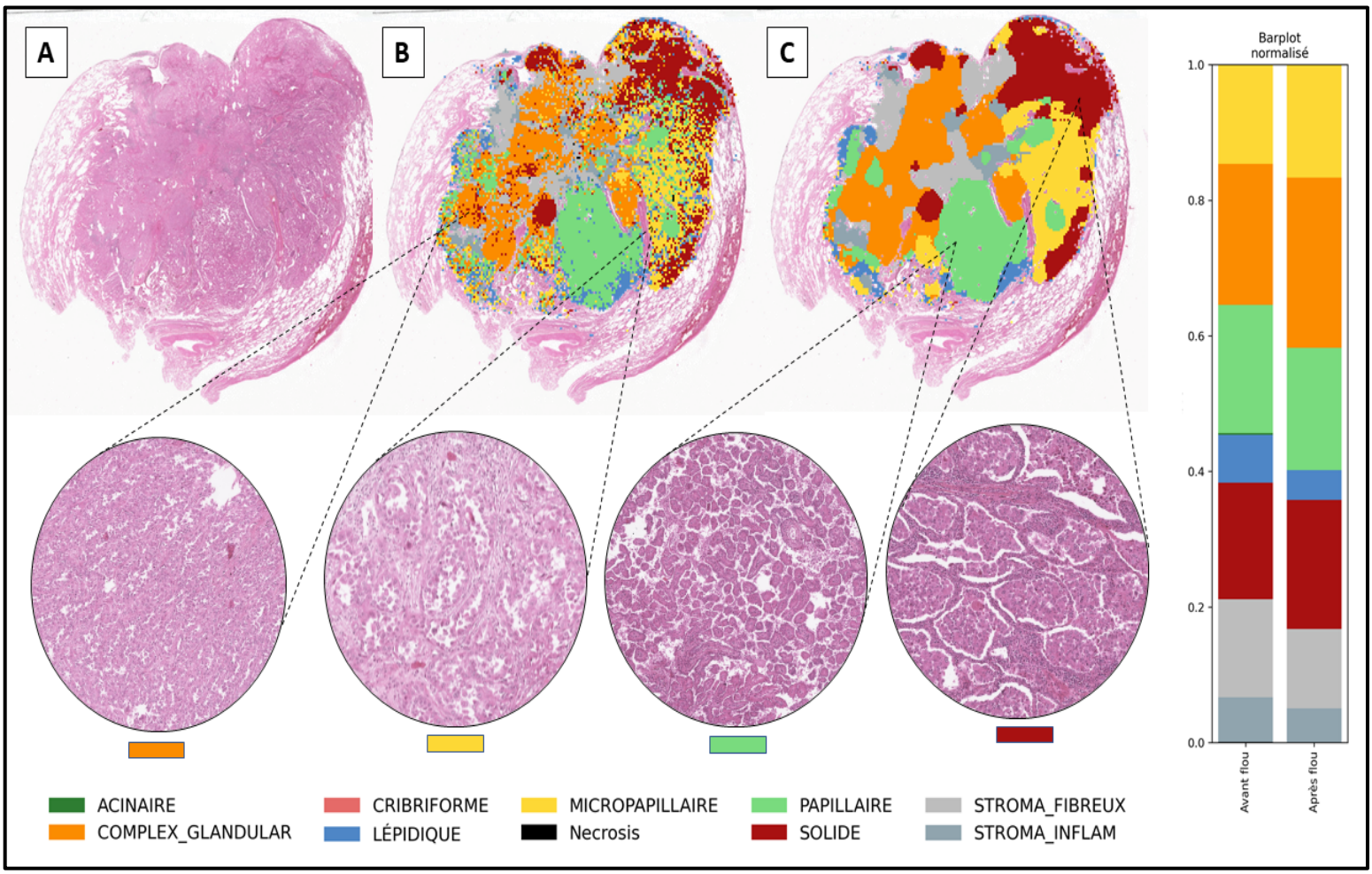


Figure 32 : **Exemple n°1. A.** Lame HES. **B.** Segmentation avec le modèle « brut ». **C.** Segmentation avec le modèle « lissé ».

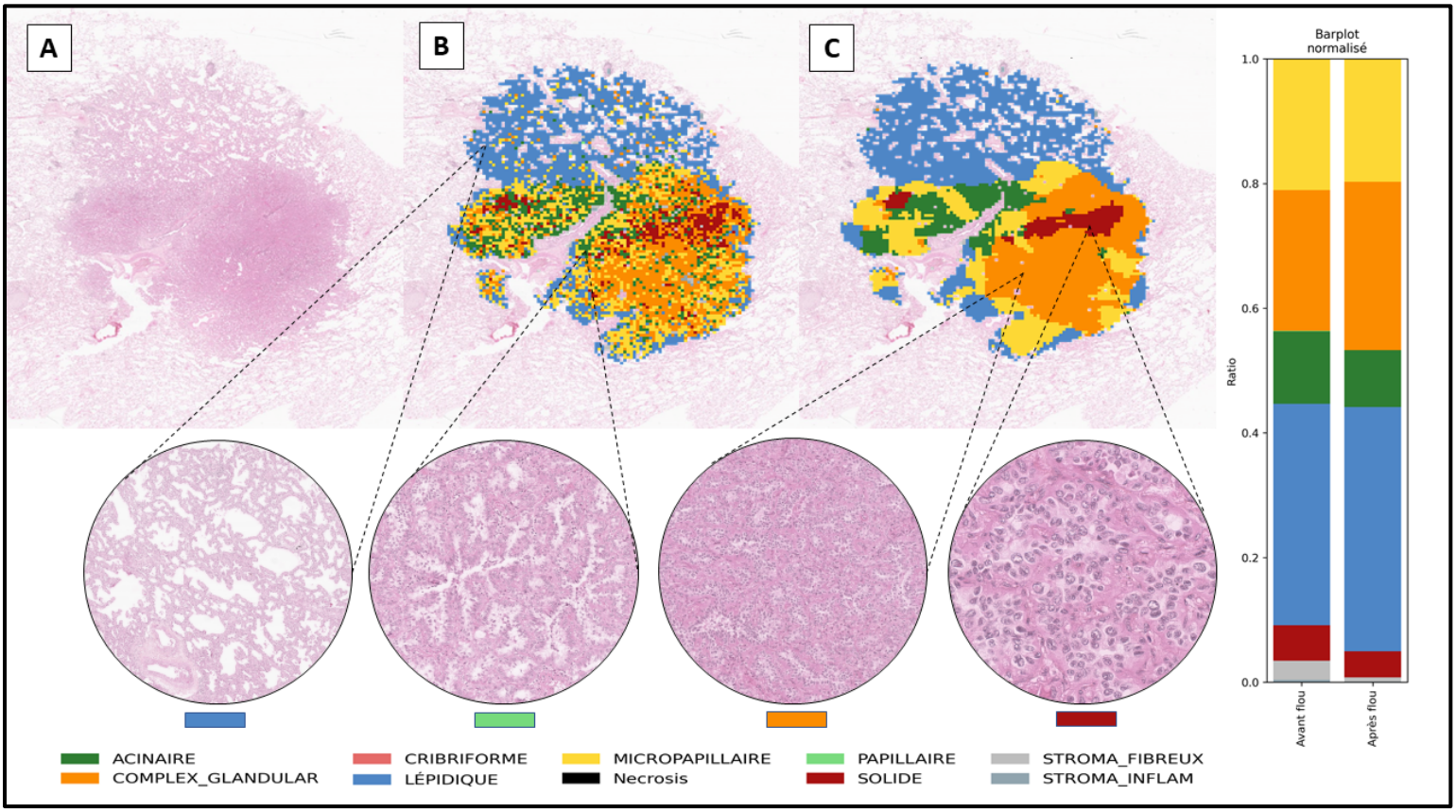


Figure 33 : **Exemple n°2. A.** Lame HES. **B.** Segmentation avec le modèle « brut ». **C.** Segmentation avec le modèle « lissé ».

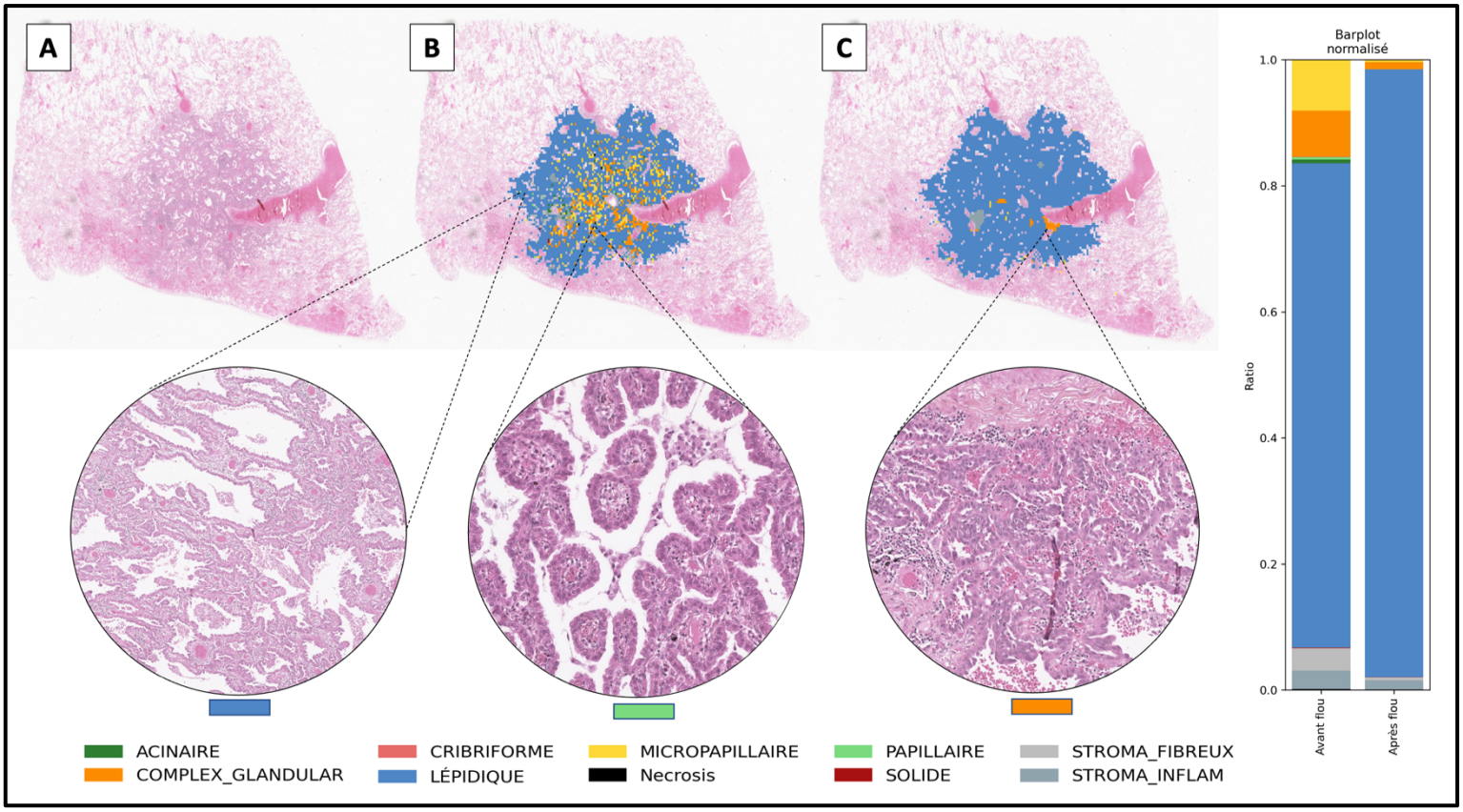


Figure 34 : **Exemple n°3. A.** Lame HES. **B.** Segmentation avec le modèle « brut » **C.** Segmentation avec le modèle « lissé ».

DISCUSSION

Le grade histologique, basé sur la proportion des différents patterns observables dans la tumeur, est un critère histopronostique majeur des ADK pulmonaires opérés. Cependant, l'hétérogénéité intra-tumorale rend le grading complexe et sujet à une variabilité inter-observateur. Dans ce contexte, notre étude avait pour objectif d'évaluer les performances d'un modèle de fondation de type « Vision Transformer » pour la reconnaissance des patterns histologiques et le grading tumoral selon les recommandations de l'IASLC. Les résultats obtenus montrent une bonne concordance globale entre le modèle et le pathologiste, avec des indices Kappa pondérés élevés, notamment après application du lissage et pondération, qui améliorent la performance de la classification. Les erreurs de classification concernaient principalement certains patterns, comme le pattern glandulaire complexe, sans impact sur le grading tumoral dans la très grande majorité des cas.

En résumé, sur les **figures 17 à 20 (tableaux A)**, on observe que la concordance brute du modèle vs pathologiste senior est modérée sur les patterns les plus complexes, en particulier pour les pattern "complex glandular " et cribriforme, souvent mal classés par le modèle. Après application du lissage, puis de la pondération, et enfin de la combinaison des deux (pondération + lissage), la performance du modèle est améliorée. Ces ajustements permettent notamment de réduire les erreurs de classification isolées, comme les confusions entre les patterns micropapillaire et complex glandular, observées dans le modèle brut.
Ces mêmes **figures 17-20 (tableaux B),** présentent également l'évolution de la concordance pour le grading tumoral. Initialement, le modèle présente une valeur de Kappa pondéré de 0,58, traduisant une concordance seulement modérée avec l'évaluation du pathologistesenior. L'application du lissage seul améliore sensiblement la concordance (Kappa pondéré = 0,73), en réduisant notamment les erreurs de classification sur le grade 1. La pondération seule permet également une nette amélioration (Kappa pondéré = 0,76), en corrigeant particulièrement les surestimations de grade (ex. : 11,8 % des cas de grade 1 initialement attribués à tort au grade 3 par l'IA). Enfin, la combinaison des deux approches offre le meilleur compromis avec un Kappa pondéré de 0,75 et une réduction importante des erreurs critiques (notamment l'absence totale de confusion entre grades 1 et 3).

Ces résultats démontrent que les techniques de lissage et de pondération améliorent significativement la fiabilité du grading produit par le modèle, en le rapprochant de celui d'un expert pathologiste senior, tout en atténuant les effets des prédictions aberrantes. Ces ajustements apparaissent donc comme des étapes essentielles pour optimiser l'intégration de ce type de modèle dans la pratique diagnostique.

Cependant, dans quatre cas **(fig 31)**, les erreurs de classification ont entraîné une différence de deux niveaux de grade (grade 1 versus 3).

Notre étude présente plusieurs atouts méthodologiques qui renforcent la pertinence des résultats obtenus. En premier lieu, elle repose sur l'utilisation d'un modèle de fondation de dernière génération, de type Vision Transformer (ViT), pré-

entraîné sur un corpus histologique massif et hétérogène. Ce choix technologique confère au modèle une capacité de généralisation supérieure à celle des approches CNN traditionnelles, comme l'ont montré récemment les travaux *"Towards a general-purpose foundation model for computational pathology"* (*Nature* 2024) (41) et *"A foundation model for clinical-grade computational pathology and rare cancers detection"* (*Nature* 2024) (40). D'autre part, le choix d'implémenter le modèle *« H-Optimus-0 »* repose sur plusieurs considérations stratégiques et méthodologiques. Tout d'abord, *« H-Optimus-0 »* possède des droits d'utilisation moins restrictifs que ceux d'UNI. Ensuite, contrairement à la majorité des modèles entraînés sur des lames HE (hématoxyline-éosine), *« H-Optimus-0 »* a été spécifiquement pré-entraîné sur une large base de données de lames en coloration HE mais également HES (hématoxyline-éosine-safran). Cette caractéristique est avantageuse pour notre étude, puisque l'ensemble des lames analysées dans ce travail ont été préparées selon ce protocole de coloration, le plus courant en France. Il est donc raisonnable de penser que le modèle *« H-Optimus-0 » pourrait* offrir une meilleure sensibilité morphologique et une interprétation plus fine des structures tissulaires colorées au safran (ex. stroma, cartilage). Deuxièmement, l'évaluation a été réalisée à partir d'un jeu de données diversifié, intégrant une large variété de patterns histologiques d'ADK pulmonaire, y compris des formes complexes ou peu fréquentes comme le « complex glandular pattern ». La double annotation par un pathologiste senior et un pathologiste junior a permis d'établir un référentiel robuste et de mesurer la variabilité inter-observateur, point essentiel pour évaluer les performances du modèle. Troisièmement, notre méthodologie inclut un processus d'optimisation post-prédiction innovant, combinant lissage et pondération, visant à corriger les erreurs ponctuelles de classification et à renforcer la cohérence morphologique des résultats. Cette approche a démontré une amélioration nette des indices de concordance, particulièrement pour les grades extrêmes. Enfin, cette étude ne se limite pas à une évaluation brute des performances : elle propose une analyse fine des situations d'erreur et de leur impact potentiel sur le grading tumoral, ce qui apporte un éclairage concret sur les conditions d'intégration clinique d'un tel outil. Cette démarche s'inscrit dans une perspective pragmatique, visant à rapprocher le modèle de la pratique quotidienne du pathologiste.

Notre étude présente cependant certaines limites qui doivent être prises en compte dans l'interprétation des résultats. Tout d'abord, l'effectif de la cohorte de validation est relativement modeste, en particulier pour certains patterns peu représentés (cribriforme, micropapillaire), ce qui peut limiter la robustesse statistique des performances observées. Une validation sur des cohortes plus larges et multicentriques est nécessaire pour l'évaluation de la performance du modèle et sa généralisation. D'autre part, une seule lame représentative de la tumeur a été sélectionnée pour l'évaluation des patterns et du grading. Cependant, le grading doit prendre en compte l'évaluation des patterns sur l'ensemble des lames de la tumeur, selon les recommandations internationales, afin de le corréler au pronostic. Ensuite, bien que notre jeu de données reflète la diversité morphologique des ADK pulmonaires, il s'agit d'une étude monocentrique, ce qui peut introduire un biais de représentativité lié aux spécificités techniques (scanners de lames, protocoles histologiques, colorations). Comme le soulignent *Campanella et al.* (*Nature* 2019) (36) et *Lu et al.* (*Nature* 2024) (42), l'hétérogénéité inter-centres et inter-scanners constitue un défi majeur pour la transposabilité clinique des modèles d'IA en pathologie.De plus, certaines discordances observées entre le modèle et le pathologiste senior se concentrent sur des situations morphologiquement complexes : patterns solides à

cellules claires confondus avec des complex glandular patterns, patterns lépidiques « collabés » reclassés en acinaires, ou encore zones à pattern micropapillaire prédominant mais non détecté par le modèle. Bien que ces erreurs n'aient pas modifié le grading final dans la majorité des cas, elles mettent en évidence des zones de fragilité diagnostique qui pourraient avoir des implications pronostiques dans des contextes particuliers. Enfin, l'utilisation d'un modèle de fondation pré-entraîné apporte une grande capacité de généralisation, mais limite la transparence sur certaines étapes d'apprentissage initial. L'explicabilité et l'interprétabilité des décisions du modèle restent ainsi perfectibles, en particulier pour comprendre pourquoi certains patterns sont privilégiés par le modèle dans des zones morphologiquement hétérogènes.

La comparaison des résultats issus de la cohorte d'entraînement et de la cohorte de test (validation) met en évidence une même particularité : l'IA tend à surestimer la présence du pattern glandulaire complexe, tout en sous-estimant le pattern solide. Cette surestimation ne peut être attribuée aux annotations utilisées pour l'entraînement, mais semble plutôt liée à la difficulté d'isoler le pattern glandulaire complexe. Une solution envisageable pour remédier à ce biais consisterait à enrichir l'entraînement par l'ajout d'annotations supplémentaires de ce pattern.

Plusieurs travaux utilisant des modèles d'IA pour le grading des ADK pulmonaires ont été publiés. Ces modèles ont principalement recours à des architectures de réseaux de neurones convolutifs (CNN), parfois associées à des approches plus récentes intégrant des mécanismes d'attention ou des réseaux hybrides. L'objectif commun est d'automatiser et de standardiser l'évaluation morphologique, afin de réduire la variabilité inter-observateur et d'optimiser la reproductibilité du grading tumoral. Les modèles CNN classiques, tels que ceux utilisés par Miura et al. (*Jpn J Clin Oncol*, 2024) (44) ou Wei et al. (*Scientific Reports*, 2019) (46), ont démontré des performances intéressantes en classification des patterns prédominants, avec des précisions globales élevées (accuracy > 90 % pour Miura) mais des scores kappa plus modestes (0,525 pour Wei). Ces approches se limitaient cependant à une classification globale par lame, sans segmentation fine des patterns. Des modèles plus avancés, comme celui de Pan et al. (*Nature Cancer*, 2024) (33) basé sur l'architecture ANORAK et intégrant un réseau de type Pyramid Pooling Cross-stream Attention Network, ont permis d'aller plus loin en réalisant une segmentation sémantique des patterns et un grading selon les critères IASLC. Néanmoins, les performances rapportées (Dice ≈ 0,60–0,65 ; kappa = 0,46) restent perfectibles, notamment sur certains patterns difficiles.
Parallèlement, Lami et al. (*Am J Pathol*, 2023) (43) ont montré qu'une approche basée sur des CNN standards pouvait atteindre un F1-score > 0,90 en classification et grading, mais sur des ensembles de données restreints et sans réelle validation multicentrique.

Dans ce contexte, notre travail se distingue par l'utilisation d'un modèle de fondation de type Vision Transformer, capable de tirer parti d'un apprentissage pré-entraîné sur un corpus massif et diversifié, tout en perfectionnant les résultats en combinant lissage et pondération. Cette approche nous a permis d'obtenir des indices de concordance pondérés supérieurs à ceux rapportés dans la majorité des études CNN, notamment après optimisation, tout en conservant une performance stable sur l'ensemble des patterns, y compris les plus rares.

Enfin, contrairement à plusieurs études qui se limitent à l'évaluation de la précision brute, notre analyse inclut une étude qualitative des situations d'erreur, mettant en évidence les contextes morphologiques spécifiques responsables des discordances. Cette démarche, rarement détaillée dans la littérature, apporte un éclairage important sur les conditions d'intégration clinique de l'IA en pathologie thoracique.

CONCLUSION ET PERSPECTIVES

Cette étude avait pour objectif d'évaluer les performances d'un modèle de fondation de type Vision Transformer pour la classification automatisée des patterns histologiques et le grading tumoral des ADK pulmonaires. L'analyse a montré une bonne concordance globale avec l'expertise d'un pathologiste senior, en particulier après application d'un processus d'optimisation combinant lissage et pondération. Les discordances observées concernaient principalement des contextes morphologiques complexes, et, dans la majorité des cas, elles n'avaient pas d'impact sur le grading final. L'intégration d'une analyse qualitative des erreurs a permis d'identifier des situations spécifiques où des améliorations ciblées du modèle pourraient encore accroître sa robustesse.

L'utilisation de modèles de fondation en pathologie ouvre la voie à une standardisation et une reproductibilité accrue dans l'évaluation des ADK pulmonaires. Leur capacité à analyser lames entières (WSI) et à segmenter finement les patterns permettrait de réduire la variabilité inter-observateur, d'accélérer le processus diagnostique et d'optimiser le grading tumoral, paramètre clé dans la stratification pronostique et la prise de décision thérapeutique. Ces modèles pourraient également faciliter la mise en place de systèmes de quantification morphologique à grande échelle, utiles pour les études multicentriques et les essais cliniques.

Plusieurs axes de travail sont envisagés pour renforcer l'intégration clinique de ce type de modèle :

- Evaluation de la performance du modèle pour le grading sur l'ensemble des lames tumorales pour chaque cas, afin de corréler le grade au pronostic (survie globale, survie sans récidive).
- Validation du modèle de classification sur des cohortes indépendantes afin de confirmer la robustesse et la transférabilité du modèle dans des contextes techniques et démographiques variés. Une validation à court terme est envisagée sur les lames d'ADK de la cohorte TCGA, ainsi que sur une cohorte du CHU de Saint Etienne (Pr F Forest).

Enfin, dans le cadre d'un projet collaboratif avec une équipe INRIA (Dr Xavier Descombes), nous sommes en train d'évaluer, sur cette même cohorte, la performance d'une approche originale d'apprentissage par renforcement (« *reinforcement learning »)* pour la reconnaissance des patterns histologiques et le grading tumoral (50).

En conclusion, les résultats préliminaires obtenus dans ce travail suggèrent que le modèle de fondation « H-optimus-0 » présente de meilleures performances pour le grading de l'ADK pulmonaire que les autres modèles publiés jusqu'alors, et pourrait représenter une avancée vers une pathologie augmentée, alliant expertise humaine et puissance analytique de l'IA.

BIBLIOGRAPHIE

1. WHO Classification of Tumours Editorial Board. WHO Classification of Tumours. Thoracic tumours. 5th edition. Vol. 5. International Agency for Research on Cancer; 2021.

2. Hegedűs F, Zombori-Tóth N, Kiss S, Lantos T, Zombori T. Prognostic impact of the IASLC grading system of lung adenocarcinoma: a systematic review and meta-analysis. Histopathology. juill 2024;85(1):51-61.

3. Van Der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. Nat Med. mai 2021;27(5):775-84.

4. Tan KS, Reiner A, Emoto K, Eguchi T, Takahashi Y, Aly RG, et al. Novel Insights Into the International Association for the Study of Lung Cancer Grading System for Lung Adenocarcinoma. Mod Pathol Off J U S Can Acad Pathol Inc. juill 2024;37(7):100520.

5. Zhang Y, Zhang Y, Hu Y, Zhang S, Zhu M, Hu B, et al. Validation of the novel International Association for the Study of Lung Cancer grading system and prognostic value of filigree micropapillary and discohesive growth pattern in invasive pulmonary adenocarcinoma. Lung Cancer. janv 2023;175:79-87.

6. Rami-Porta R, Asamura H, Travis WD, Rusch VW. Lung cancer — major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin. 2017;67(2):138-55.

7. Moreira AL, Ocampo PSS, Xia Y, Zhong H, Russell PA, Minami Y, et al. A Grading System for Invasive Pulmonary Adenocarcinoma: A Proposal From the International Association for the Study of Lung Cancer Pathology Committee. J Thorac Oncol Off Publ Int Assoc Study Lung Cancer. oct 2020;15(10):1599-610.

8. Borczuk AC. Updates in grading and invasion assessment in lung adenocarcinoma. Mod Pathol. janv 2022;35:28-35.

9. Rokutan-Kurata M, Yoshizawa A, Ueno K, Nakajima N, Terada K, Hamaji M, et al. Validation Study of the International Association for the Study of Lung Cancer Histologic Grading System of Invasive Lung Adenocarcinoma. J Thorac Oncol. oct 2021;16(10):1753-8.

10. Yoshizawa A, Motoi N, Riely GJ, Sima CS, Gerald WL, Kris MG, et al. Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases. Mod Pathol Off J U S Can Acad Pathol Inc. mai 2011;24(5):653-64.

11. Borczuk AC. Prognostic considerations of the new World Health Organization classification of lung adenocarcinoma. Eur Respir Rev. déc 2016;25(142):364-71.

12. Emoto K, Eguchi T, Tan KS, Takahashi Y, Aly RG, Rekhtman N, et al. Expansion of the Concept of Micropapillary Adenocarcinoma to Include a Newly Recognized Filigree Pattern as Well as the Classical Pattern Based on 1468 Stage I Lung Adenocarcinomas. J Thorac Oncol. nov 2019;14(11):1948-61.

13. Thunnissen E, Beasley MB, Borczuk AC, Brambilla E, Chirieac LR, Dacic S, et al. Reproducibility of histopathological subtypes and invasion in pulmonary adenocarcinoma. Mod Pathol Off J U S Can Acad Pathol Inc. déc 2012;25(12):1574-83.

14. Thunnissen E, Beasley MB, Borczuk AC, Brambilla E, Chirieac LR, Dacic S, et al. Reproducibility of histopathological subtypes and invasion in pulmonary adenocarcinoma. An international interobserver study. Mod Pathol Off J U S Can Acad Pathol Inc. déc 2012;25(12):1574-83.

15. Blaauwgeers H, Flieder D, Warth A, Harms A, Monkhorst K, Witte B, et al. A Prospective Study of Loose Tissue Fragments in Non-Small Cell Lung Cancer Resection Specimens: An Alternative View to « Spread Through Air Spaces ». Am J Surg Pathol. sept 2017;41(9):1226-30.

16. Trejo Bittar HE, Jerome JA, Hartman D, Pantanowitz L, Mehrad M, Dacic S. Prognostic significance of microscopic size in peripherally located scar-associated clinical stage I lung carcinomas. Lung Cancer Amst Neth. mai 2020;143:12-8.

17. Thunnissen E, Flieder D. Morphologic Logic: « Filigree » and « Classical » Micropapillary Pattern Are Orientation-Dependent Views of the Same Lesion. J Thorac Oncol Off Publ Int Assoc Study Lung Cancer. juill 2020;15(7):e120-1.

18. Emoto K, Tan KS, Rekhtman N, Adusumilli PS, Travis WD. The Newly Described Filigree Pattern Is an Expansion of the Micropapillary Adenocarcinoma Concept Rather Than a Proposed New Subtype. J Thorac Oncol Off Publ Int Assoc Study Lung Cancer. juill 2020;15(7):e121-4.

19. Warth A, Stenzinger A, Von Brünneck AC, Goeppert B, Cortis J, Petersen I, et al. Interobserver variability in the application of the novel IASLC/ATS/ERS classification for pulmonary adenocarcinomas. Eur Respir J. nov 2012;40(5):1221-7.

20. Lipkova J, Chen RJ, Chen B, Lu MY, Barbieri M, Shao D, et al. Artificial intelligence for multimodal data integration in oncology. Cancer Cell. 10 oct 2022;40(10):1095-110.

21. Ozyoruk KB, Can S, Darbaz B, Başak K, Demir D, Gokceler GI, et al. A deep-learning model for transforming the style of tissue images from cryosectioned to formalin-fixed and paraffin-embedded. Nat Biomed Eng. 23 déc 2022;6(12):1407-19.

22. Lu MY, Chen B, Williamson DFK, Chen RJ, Zhao M, Chow AK, et al. A multimodal generative AI copilot for human pathology. Nature. 10 oct 2024;634(8033):466-73.

23. Medathati NVK, Neumann H, Masson GS, Kornprobst P. Bio-inspired computer vision: Towards a synergistic approach of artificial and biological vision. Comput Vis Image Underst. sept 2016;150:1-30.

24. Hoeser T, Kuenzer C. Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review-Part I: Evolution and Recent Trends. Remote Sens. janv 2020;12(10):1667.

25. Yamamoto Y, Tsuzuki T, Akatsuka J, Ueki M, Morikawa H, Numata Y, et al. Automated acquisition of explainable knowledge from unannotated histopathology images. Nat Commun. 18 déc 2019;10(1):5642.

26. Calderaro J, Kather JN. Artificial intelligence-based pathology for gastrointestinal and hepatobiliary cancers. Gut. 1 juin 2021;70(6):1183-93.

27. Perez-Lopez R, Ghaffari Laleh N, Mahmood F, Kather JN. A guide to artificial intelligence for cancer researchers. Nat Rev Cancer. juin 2024;24(6):427-41.

28. Prabhu S, Prasad K, Robels-Kelly A, Lu X. AI-based carcinoma detection and classification using histopathological images: A systematic review. Comput Biol Med. 1 mars 2022;142:105209.

29. Wang X, Zhao J, Marostica E, Yuan W, Jin J, Zhang J, et al. A pathology foundation model for cancer diagnosis and prognosis prediction. Nature. 24 oct 2024;634(8035):970-8.

30. Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep. 23 mai 2016;6(1):26286.

31. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and Mutation Prediction from Non-Small Cell Lung Cancer Histopathology Images using Deep Learning. Nat Med. oct 2018;24(10):1559-67.

32. Liu S, Wang X, Liu M, Zhu J. Towards better analysis of machine learning models: A visual analytics perspective. Vis Inform. mars 2017;1(1):48-56.

33. Pan X, AbdulJabbar K, Coelho-Lima J, Grapa AI, Zhang H, Cheung AHK, et al. The artificial intelligence-based model ANORAK improves histopathological grading of lung adenocarcinoma. Nat Cancer. 10 janv 2024;5(2):347-63.

34. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat Med. juill 2019;25(7):1054-6.

35. Kather JN, Heij LR, Grabsch HI, Loeffler C, Echle A, Muti HS, et al. Pan-cancer image-based detection of clinically actionable genetic alterations. Nat Cancer. 27 juill 2020;1(8):789-99.

36. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Werneck Krauss Silva V, Busam KJ, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med. août 2019;25(8):1301-9.

37. Ding H, Feng Y, Huang X, Xu J, Zhang T, Liang Y, et al. Deep learning-based classification and spatial prognosis risk score on whole-slide images of lung adenocarcinoma. Histopathology. août 2023;83(2):211-28.

38. Fu Y, Jung AW, Torne RV, Gonzalez S, Vöhringer H, Shmatko A, et al. Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. Nat Cancer. 27 juill 2020;1(8):800-10.

39. Kaczmarzyk JR, O'Callaghan A, Inglis F, Gat S, Kurc T, Gupta R, et al. Open and reusable deep learning for pathology with WSInfer and QuPath. Npj Precis Oncol. 10 janv 2024;8(1):9.

40. Vorontsov E, Bozkurt A, Casson A, Shaikovski G, Zelechowski M, Severson K, et al. A foundation model for clinical-grade computational pathology and rare cancers detection. Nat Med. oct 2024;30(10):2924-35.

41. Chen RJ, Ding T, Lu MY, Williamson DFK, Jaume G, Song AH, et al. Towards a general-purpose foundation model for computational pathology. Nat Med. mars 2024;30(3):850-62.

42. Lu MY, Chen B, Williamson DFK, Chen RJ, Liang I, Ding T, et al. A visual-language foundation model for computational pathology. Nat Med. mars 2024;30(3):863-74.

43. Lami K, Ota N, Yamaoka S, Bychkov A, Matsumoto K, Uegami W, et al. Standardized Classification of Lung Adenocarcinoma Subtypes and Improvement of Grading Assessment Through Deep Learning. Am J Pathol. déc 2023;193(12):2066-79.

44. Miura E, Emoto K, Abe T, Hashiguchi A, Hishida T, Asakura K, et al. Establishment of artificial intelligence model for precise histological subtyping of lung adenocarcinoma and its application to quantitative and spatial analysis. Jpn J Clin Oncol. 4 sept 2024;54(9):1009-23.

45. Zhao Y, He S, Zhao D, Ju M, Zhen C, Dong Y, et al. Deep learning-based diagnosis of histopathological patterns for invasive non-mucinous lung adenocarcinoma using semantic segmentation. BMJ Open. juill 2023;13(7):e069181.

46. Wei JW, Tafe LJ, Linnik YA, Vaickus LJ, Tomita N, Hassanpour S. Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks. Sci Rep. 4 mars 2019;9(1):3358.

47. Komura D, Ishikawa S. Machine Learning Methods for Histopathological Image Analysis. Comput Struct Biotechnol J. 2018;16:34-42.

48. Baxi V, Edwards R, Montalto M, Saha S. Digital pathology and artificial intelligence in translational medicine and clinical practice. Mod Pathol. janv 2022;35(1):23-32.

49. Jaume G, Doucet P, Song AH, Lu MY, Almagro-Pérez C, Wagner SJ, et al. HEST-1k: A Dataset for Spatial Transcriptomics and Histology Image Analysis [Internet]. arXiv; 2024 [cité 15 juill 2025]. Disponible sur: http://arxiv.org/abs/2406.16192

50. Mohamad Mohamad FP, Maxime Gassier NP. Investigating Reinforcement Learning for Histopathological Image Analysis.

Table des figures

Figure 1 : Courbes de survie sans récidive selon la classification internationale IASLC et le système de classification basé sur le pattern prédominant……16

Figure 2 : Survie sans récidive (E) et survie globale (F) des patients après chirurgie selon le grade histologique de la tumeur……18

Figure 3 : Proposition de raisonnement devant une lésion suspecte d'adénocarcinome pulmonaire……..24

Figure 4 : Vue d'ensemble et détaillée d'un réseau neuronal convolutif (CNN)……..27

Figure 5 : Illustration de l'étape de « pooling » d'un CNN……28

Figure 6 : Modèle informatique proposé pour la cartographie de précision et les analyses d'hétérogénéité spatiale……30

Figure 7 : Architecture du modèle UNI et couverture des tâches cliniques……..31

Figure 8 : Performances de différents modèles de fondation (UNI, ResNet, REMEDIS, CTransPath) pour 15 tâches……33

Figure 9 : Tableau comparatif des performances de modèles IA récents dans la caractérisation des patterns tumoraux des ADK pulmonaires……34

Figure 10 : Tableau comparatif des convolutional neural network (CNN) classiques et des modèles de fondation……37

Figure 11 : Diagramme de flux présentant la sélection des patients……39

Figure 12 : Exemples d'annotations réalisées sur lames numérisées (logiciel QuPATH)……41

Figure 13 : Tableau comparatif des performances des modèles de fondation les plus récents……42

Figure 14 : Evaluation par le modèle de fondation « H-Optimus-0 » de la proportion des patterns sur une lame numérisée d'adénocarcinome pulmonaire……..……45

Figure 15 : Protocole de classification du grading des adénocarcinomes pulmonaires, basé sur le modèle de fondation « H-Optimus-0 »……46

Figure 16 : Concordance inter-observateurs (pathologiste senior versus pathologiste junior) sur le pattern prédominant (A) et sur le grading (B)……48

Figure 17 : Performances du modèle « brut » pour la classification des patterns (A) et le grading (B) des adénocarcinomes pulmonaires……49

Figure 18 : Performances du modèle « lissé » pour la classification des patterns (A) et le grading (B) des adénocarcinomes pulmonaires……50

Figure 19 : Performances du modèle « pondéré » pour la classification des patterns (A) et le grading (B) des adénocarcinomes pulmonaires……51

Figure 20 : Performances du modèle « lissé » et « pondéré » pour la classification des patterns (A) et le grading (B) des adénocarcinomes pulmonaires……..52

Figure 21 : Comparaison appareillée des proportions par pattern sur la cohorte test (validation). ……53

Figure 22 : Comparaison appareillée des proportions par pattern sur la cohorte d'entrainement……54

Figure 23 : Heat-map d'un adénocarcinome comportant un pattern lépidique……55

Figure 24 : Heat-map d'un adénocarcinome comportant un pattern acinaire……..56

Figure 25 : Heat-map d'un adénocarcinome comportant un pattern papillaire et un foyer de stroma inflammatoire……..56

Figure 26 : Heat-map d'un adénocarcinome comportant un pattern micropapillaire…57

Figure 27 : Heat-map d'un adénocarcinome comportant un pattern solide……..58

Figure 28 : Heat-map d'un adénocarcinome comportant un pattern cribriforme……..58

Figure 29 : Heat-map d'un adénocarcinome comportant un pattern « complex glandular »……..59

Figure 30 : Exemples de discordances sur les patterns……..60

Figure 31 : Exemples de discordance sur les grades (grade 3 versus grade 1)……..61

Figure 32 : Exemple n°1 des résultats obtenus via le modèle « H-optimis-0 »…….62

Figure 33 : Exemple n°2 des résultats obtenus via le modèle « H-optimis-0 »…….63

Figure 34 : Exemple n°3 des résultats obtenus via le modèle « H-optimis-0 »…….63

Table des tableaux

Table 1.08 (1) : Système de grading OMS des adénocarcinomes pulmonaires (selon la prédominance architecturale et la proportion de patterns de haut grade……18

Tableau 1 : Caractéristiques anatomo-cliniques des patients……47

Titre de la thèse : APPORT D'UN MODELE DE FONDATION POUR LE GRADING HISTOLOGIQUE DES ADENOCARCINOMES PULMONAIRES

Résumé :
Introduction : L'adénocarcinome (ADK) pulmonaire est la tumeur maligne la plus fréquente du poumon et présente un taux élevé de récidive post-opératoire. Le grade tumoral, basé sur l'évaluation par le Pathologiste de la proportion des différents patterns histologiques sur la pièce opératoire, est un critère histopronostique majeur. Cependant, le grading est une tâche complexe et chronophage, et présente une variabilité inter-observateur importante, incitant au développement d'outils d'intelligence artificielle. Récemment, les modèles de fondation ont montré de très bonnes performances dans diverses applications en Pathologie digitale. Les objectifs principaux de ce travail étaient d'évaluer la performance d'un modèle de fondation (*H-optimus-0*) pour la classification des patterns et le grading des ADK pulmonaires.
Matériel et Méthodes : Cette étude rétrospective monocentrique a inclus 122 ADK pulmonaires réséqués à l'hôpital Bichat, répartis en deux cohortes (entraînement n=36 ; validation n=86). Pour chaque cas, une lame HES représentative de la tumeur a été sélectionnée puis scannée. La proportion des différents patterns (lépidique, papillaire, micropapillaire, acinaire, cribriforme, solide, glandulaire complexe) et le grade (selon OMS) de chaque lame ont été évalués indépendemment par deux pathologistes (junior/senior). Les différents patterns ont été annotés sur la cohorte d'entrainement. Le modèle *H-optimus-0* a été appliqué sur les deux cohortes. Afin d'optimiser la performance du modèle, un lissage et une pondération des données ont été appliqués. Les prédictions du modèle ont été comparées à l'évaluation faite par le Pathologiste senior (PS), à l'aide de matrices de covariance et indice Kappa pondéré.
Résultats : Concernant la classification des patterns, les taux de concordance entre le modèle et le PS étaient de 0%, 8.3%, 16.7%, 28.6%, 42.9% ; 68% et 73.7% pour les patterns cribriforme, acinaire, micropapillaire, glandulaire complexe, papillaire, solide et lépidique, respectivement. Les taux de concordance entre le modèle et le PS concernant le grade étaient de 64.7% pour le grade 1, 43.8% pour le grade 2 et 98.1% pour le grade 3, avec un Kappa pondéré de 0.76. Seuls 4 cas présentaient une discordance de 2 grades (grade 1 vs grade 3) entre le modèle et le PS.
Conclusion : Ce travail montre que la performance du modèle de fondation *H-optimus-0* pour la classification des patterns est très variable selon le pattern, mais est satisfaisante pour le grading, avec une très bonne prédiction des grades 3. Ces résultats préliminaires sont à confirmer dans des cohortes indépendantes. A terme, l'intégration de cet outil à la pratique du Pathologiste pourrait améliorer la reproductibilité du grading des ADK pulmonaires, et ainsi améliorer l'évaluation du pronostic post-opératoire.

Spécialité : DES Anatomie et Cytologie Pathologiques
Mots clés français : adénocarcinome pulmonaire, classification histologique, intelligence artificielle, modèle de fondation, transformeur visuel, analyse d'images en microscopie, coupe histologique virtuelle, segmentation d'images médicales, pathologie numérique, concordance interobservateur.

Forme ou Genre :
fMeSH : Dissertation universitaire

Université Paris Cité
UFR de médecine
15 Rue de l'École de Médecine
75006 Paris